Charles University

Faculty of Science

Doctoral study programme: Physical chemistry

Mgr. Anna Dubánková

Structural and functional study of viral RNA polymerases

Strukturní a funkční studie virových RNA polymeráz

Doctoral thesis

Supervisor: Mgr. et Mgr. Evžen Bouřa, PhD.

Prague, 2019

Acknowledgements

First of all, I would like to thank my supervisor Evzen Boura for scientific and patient approach and willingness to help. I am really grateful to him for introducing me to the magical world of structural biology.

My special thanks also belong to all lab members for being more friends than colleagues and for creating friendly atmosphere with so much fun. I will always remember our coffee breaks and brainstormings in Kavárna.

Last but not the least, I would like to thank my family for supporting me spiritually throughout my studies and my life in general.

Thanks for all your encouragement!

Prohlašuji, že jsem nepředložila práci ani její podstatnou část k získání jiného nebo stejného akademického titulu. Práci jsem zpracovala samostatně a uvedla všechny použité informační zdroje a literaturu.

In Prague

……

Mgr. Anna Dubánková

Abstract

Viral RNA-dependent RNA polymerases (RdRps) are enzymes essential for viral multiplication. The general function of RdRp is universal for all RNA viruses: to recognise viral RNA, bind it and synthesize the complementary RNA strand. This series of steps is absolutely crucial for viral infection.

It is important to mention that the non-infected cell is incapable of replicating any RNA. The host cell thus does not naturally express any RdRps. I chose RdRps for my research because these enzymes are key to viral replication and thus an excellent target for antivirals.

This study characterises polymerases from *Picornaviridae* and *Flaviviridae* families, in depth. Picornaviral replication takes place in viral-induced membrane structures called Replication Organelles (ROs), where the polymerase is localised to the membrane. In this study, we investigated the recruitment of picornaviral polymerase membrane. Subsequently, we focused on the activation of picornaviral RdRp induced by the insertion of the very first residue into the protein core.

Next, we focused on the flaviviral RdRps specifically from yellow fever virus (YFV) and Zika virus (ZIKV). This study reports the first structure of a full length YFV polymerase and a model of ZIKV polymerase in complex with RNA. The model of ZIKV RdRp in complex with RNA provides the information needed for ligand docking.

Abstrakt (in Czech)

Virové RNA dependentní RNA polymerázy (RdRp) jsou enzymy nezbytné pro množení RNA virů. Obecná funkce RdRp je pro všechny RNA viry stejná: RdRp rozpozná virovou RNA, uchopí ji a syntetizuje komplementární řetězec RNA. Série těchto kroků je pro virovou infekci naprosto nepostradatelná.

Je důležité si uvědomit, že neinfikovaná buňka RNA nereplikuje. Hostitelská buňka přirozeně neexprimuje žádné RdRp. Pro můj výzkum jsem si vybrala RdRp, protože jsou nezbytné pro virovou replicakci a tudíž excelentní cíl pro virovou terapii.

V této studii jsou do hloubky charakterizovány polymerázy z rodiny picornavirů a flavivirů. Picornavirová replikace je lokalizovaná ve viry indukovaných, membránových strukturách, které nazýváme replikační organely (RO). V těchto RO je virová polymeráza lokalizovaná na membráně. Ukázali jsme podmínky, ve kterých se polymeráza váže na membránu. Následně byla studována aktivace picornavirových polymeráz, která je indukována inzercí první aminokyseliny do středu proteinu.

V této studii jsou rovněž diskutovány RdRp z rodiny flavivirů, konkrétně z viru žluté zimnice (YFV) a Zika viru (ZIKV). Studie prezentuje první strukturu polymerázy z YFV (v plné délce) a model ZIKV polymerázy v komplexu s RNA. Model ZIKV poskytuje strukturní informaci o aktivním místě potřebnou k “dockingu” ligandů.

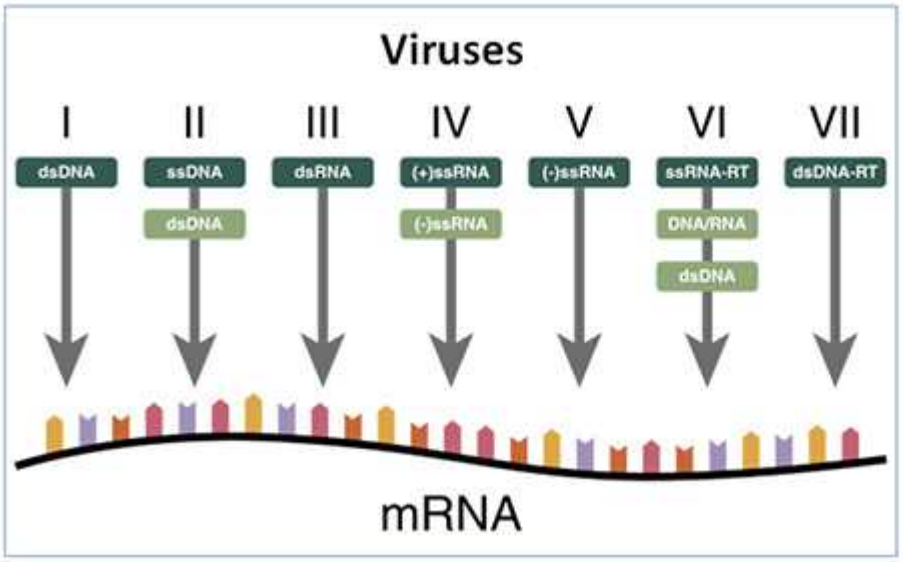
Table of contents

Introduction

Plus single-stranded (+RNA) viruses

The Baltimore classification sorts viruses into families depending on their type of genome and related manner of replication. This classification assumes that viruses with the same genome have a similar strategy of messenger RNA (mRNA) synthesis. In the study, Baltimore distinguishes seven viral classes and one of the discussed classes is class IV, +RNA viruses (picture 1). Positive sense RNA serves as both the genome and mRNA for these viruses.

+RNA viruses have developed various replication strategies but the most important replication tactic is the recognition of their genomic RNA by the host cell machinery as mRNA. The genome 5’ end can be protected by a cap in the case of Flaviviruses. Picornaviruses stabilise their genome by a covalently linked 3B protein, also known as viral protein genome linked (VPG) and the 5’ untranslated region containing an internal ribosomal entry site (IRES).



Picture 1: *Baltimore classification categorize viruses by type of genome and mechanism to generate mRNA[1].*

+RNA viral replication cycle

The replication cycle is initiated by the binding of the virus to a receptor on the host cell plasma membrane. Afterwards, the virus is internalized and the viral RNA genome is released into the cytoplasm. The +RNA genome contains one open reading frame which is translated by the host cell ribosomes into a single polyprotein. This polyprotein contains a complete set of viral proteins. Viral and host proteases cleave the polyprotein into individual proteins and stable precursors.

Proteins related to viral RNA replication are localised in viral induced membrane structures called ROs[2], which are described below. RdRps replicate RNA into complementary minus ssRNA, forming dsRNA as an intermediate. The minus ssRNA serves as a template to produce a larger amount of +RNA[3]. The +RNA is encapsidated and new viruses are released from the host cell.

Replication organelles (ROs)

+RNA viral RdRp are low efficiency enzymes in solution. To increase the efficiency, viruses replicate at a specific replication site, termed replication organelles (ROs)[4]. Localization of RNA replication on ROs brings several advantages:

Viral replication factors and cellular proteins are reduced from three-dimensional diffusion in the cytosol to two-dimensional space on the surface of the membrane. RO compartments are enriched with the required metabolites such as nucleotide triphosphates (NTPs), which increases replication efficiency by several orders of magnitude[5,65].

ROs enable separation of RNA replication from translation because all +RNA ROs originate from ribosome-free membranes[7]. The compartmentalization avoids interference of replication and translation machineries.

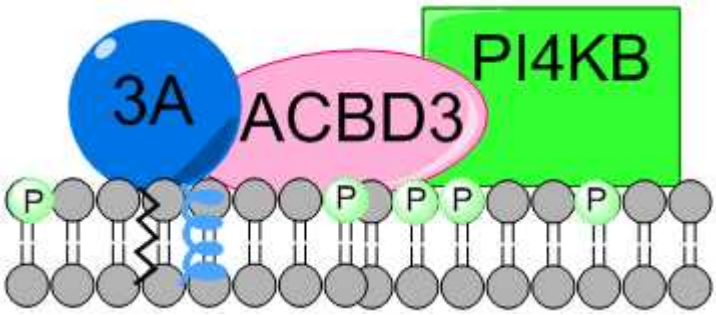
The intermediate in RNA replication is double-stranded RNA (dsRNA) which is consequently unwound by a viral helicase. The host cell innate immunity system recognises dsRNA and induces an immune response[8]. ROs minimise exposure of the viral replication complexes to the innate immunity cytoplasmic sensors.

Membrane rearrangement of ROs

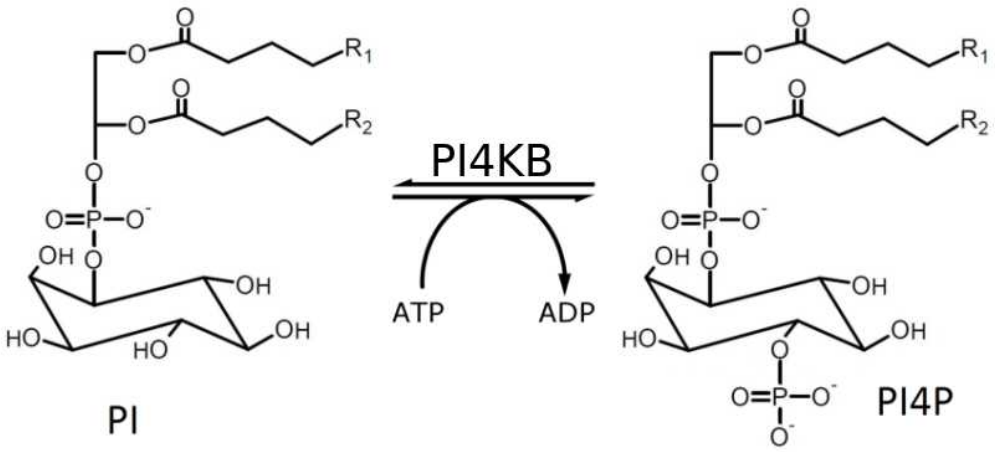
+RNA viruses reorganize and affect host cell endoplasmic reticulum (ER) to Golgi transport pathways. They recruit several cellular host factors to the ROs by direct protein−protein interaction. It was shown that inhibition of the ER to Golgi transport abolished enteroviral replication[9]. This signaling pathway is well described in the kobuviral genera from the *Picornaviridae* family.

A critical component of picornaviral ROs is viral membrane binding

protein 3A. Kobuviral 3A interacts with a host acyl-CoA-binding domain-containing protein-3 (ACBD3)[10,11]. The ACBD3 is a Golgi-resident protein involved in several signaling events[12]. The ACBD3 recruits phosphatidylinositol 4-kinase type III β (PI4KB) to the membrane[10,11] (picture 2). Then PI4KB specifically recognises phosphatidylinositol (PI) and phosphorylates it at position 4 of the inositol ring to produce a phosphatidylinositol 4-phosphate (PI4P)-rich microenvironment (picture 3).



Picture 2: *3A–ACBD3–PI4KB protein complex: Viral membrane binding protein 3A binds host factor ACBD3 to hijak PI4KB to the membrane. The PI4KB creates hyperphosphorylated membrane surface.*



Picture 3: *PI4KB phosphorylates PI at position 4 of the inositol ring to create PI4P.*

PI4P microenvironment regulates RNA replication

Enteroviruses and Flaviviruses recruit PI4KB host cell enzyme on ROs and replicate their RNA on the PI4P enriched membranes. It was shown that selective inhibition of PI4KB suppress the viral multiplication[13,14]. There is a hypothesis that enteroviral soluble RdRp ($3D^{pol}$) specifically binds the PI4P lipid which promotes the recruitment and stabilization of RdRp on the membrane from the cytosolic pool[15]. It was shown that polymerase membrane assembly significantly increases its processivity[16].

PI4P membrane function

The PI4P lipid is not only a precursor for phosphatidylinositol 4,5-bisphosphate, it has various other functions, as well. So far several PI4P effectors that use PI4P as a Golgi membrane signal were characterized[17]. The PI4P also regulates autophagy[18] and ER exit site biogenesis[19].

The PI4P lipid is important in membrane geometry. It was shown that PI4P in physiologically relevant concentration induces curvature in model membranes [20,21].

The PI4P lipid can be exchanged for cholesterol[22] or PS (phosphatidylserine)[23] against their concentration gradient by lipid transport proteins. The energy from PI4P hydrolysis is used for this transport[24]. The production of PI4P was described to increase the cholesterol levels in EMCV-infected cells[25].

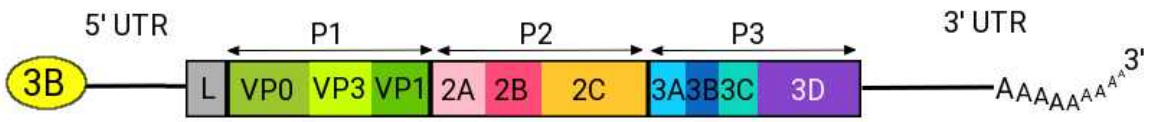
Picornaviruses

The *Picornaviridae* family belongs to a group of +RNA viruses. Picornaviruses include many important human pathogens such as poliovirus (PV), coxsackie virus B3 (CV), human rhinovirus (HRV) and others.

The appellation pico-RNA-viruses suggests their extremely small RNA

genome that has less than 10 kb. The genome is an infectious agens because itself is sufficient to infect the host cell.

The genomic 5' untranslated region contains IRES and a covalently attached viral protein 3B. The 3B protein serves as a primer for the RNA polymerase. The 3' end of the genome is polyadenylated[26]. The genome is translated into a single polyprotein which contains all picornaviral proteins (picture 4). The leader protein (L) protein followed by structural viral proteins zero to three (VP0 - VP3) and non-structural proteins (2A, 2B, 2C, 3A, 3B, 3C and $3D^{pol}$) are located at the N terminus of the polyprotein.



Picture 4: *Organisation of the picornaviral genome. The UTR with a covalently linked 3B protein is at the 5' end, followed by an open reading frame coding for the L protein, structural protein (P1) and non-structural proteins (P2-3). The 3' UTR is polyadenylated*[27].

Picornaviral proteome

The picornaviral polyprotein is autocatalytically cleaved into individual proteins or their stable precursors. First, the polyprotein pre-maturates into 3 proteins: P1, P2 and P3. The maturation continues until only single proteins remain. Each of these proteins has its specific function.

Structural proteins are pre-maturated as VP0, VP1 and VP3. Protein VP0 is consequently maturated into proteins VP2 and VP4[28]. Proteins VP1–VP4 assemble into the viral icosahedral capsid. The VP1 protein

forms a vertex with the capsid and is responsible for the host cell receptor attachment[29].

Protein P2 matures into 2A, 2B and 2C proteins. Protein 2A is a cysteine protease that cleaves between the P1 and P2 regions[30]. Protein 2B creates pores in the host cell ER[31]. Protein 2C is an NTPase and plays a role in viral RNA encapsidation[32].

Maturation of protein P3 leads to proteins 3A, 3B, 3C, 3D and their precursors with enzymatic activity. Protein 3AB is localised on replication organelles and is involved in RNA synthesis initiation[33]. Protein 3AB is a precursor for 3A and 3B. 3A is viral replication complex component and affects ER−Golgi transport[34]. Protein 3B is also known as VPG (viral protein genome-linked). The 3B protein serves as a primer protein for viral RNA replication. The absolutely conserved Tyr3 residue covalently binds to the 5' viral genome via its hydroxyl group[34,35]. The protein 3CD is a precursor of 3C and $3D^{pol}$. Protein 3CD and 3C have proteolytic activity, but recognize different cleavage sites[36,37]. The $3D^{pol}$ protein is the picornaviral RdRp. $3D^{pol}$ has two known activities: RNA replication and covalent attachment of a UMP molecule to the hydroxyl group of the 3B Tyr3[38].

The N terminus of some picornaviral polyprotein contains the L protein with proteolytic activity[39]. For better clarity, a list of non-structural picornaviral proteins including their function is shown in Table 1.

Table 1: *Overview of Picornaviral non-structural proteins function*

Protein	***Function***
L	Protease
2A	Protease
2B	Viroporin
2C	NTPase
3AB	RNA synthesis initiation
3A	Affects ER-Golgi transport
3B	Primer
3CD	Protease
3C	Protease
3Dpol	RdRp

Picornaviral polymerase $3D^{pol}$

Picornaviral polymerase is a single domain RdRp. The RdRp has a right-hand fold which is typical for polymerase. The RdRp consists of three subdomains called: Fingers, Thumb and Palm. The Fingers and Thumb subdomains create a template entry channel. The NTP entry channel is located on the boundaries of the Fingers, Thumb and Palm subdomains. The product of replication leaves RdRp through the dsRNA channel that is composed of Thumb and Palm domains[40]. The $3D^{pol}$ polymerase replicates

by an unusual primer-dependent mechanism, employing the 3B protein as its primer[41].

Proteolytic activation of Picornaviral RdRp ($3D^{pol}$)

Picornaviral polymerase $3D^{pol}$ is located at the extreme C terminal end of the polyprotein. $3D^{pol}$ is pre-maturated as 3CD. The 3CD protein has proteolytic activity and no polymerase activity[37]. The 3CD is cleaved into proteins 3C and $3D^{pol}$. The 3C protein is a protease (no polymerase activity) while the $3D^{pol}$ protein is an RdRp (no protease activity). It follows that the $3D^{pol}$ is activated by cleavage of the 3CD precursor. It was shown in the crystal structure of the 3CD form poliovirus[42] that the $3D^{pol}$ N terminus undergoes a conformational change after 3C cleavage. The N terminus of the $3D^{pol}$ in 3CD is a part of a 3C−3D linker, but after 3C cleavage the N terminus is buried in a surface pocket [43]. This mechanism was shown in all the available $3D^{pol}$ structures (from these viruses: PV[44], CV[45], encephalomyocarditis virus[46], foot and mouth disease virus[47], HRV[48] and enterovirus 71[49]).

The first residue of all the structures mentioned above is a glycine which interacts with residues in close proximity to the polymerase active site. The binding pocket is designed for the very first glycine residue. The interaction requires the small size and backbone torsional flexibility of this specific amino acid (glycine 1 is in fact in the Ramachandran disallowed region)[43].

Flaviviruses

Flaviviruses are enveloped +RNA viruses. This viral family also contains many important human pathogens including Dengue virus (DENV), hepatitis C virus (HCV), Japanese encephalitis virus (JEV), tick-borne encephalitis virus (TBE), West Nile virus (WNV), yellow fever virus (YFV),

Zika virus (ZIKV).

The RNA genome of the flaviviruses contains the 5′ cap. The +RNA genome can be translated into a single polyprotein which is consequently proteolytically maturated similarly as was described for picornaviruses.

Picture 5: *Gene localisation in the Flaviviral genome[50].*
The genome codes: C protein, a capsid protein; prM, a chaperone for the envelope protein E; E protein which mediates fusion between viral and cellular membranes; NS1 which is involved in virulence and replication[51]; NS2A which is involved in virion assembly[52]; NS2B, a cofactor of NS3; NS3, a protease; an NTPase and an RNA helicase; NS4A which regulates the ATPase activity of NS3; NS4B which induces the formation of ER-derived membrane[53]; NS5, a methyltransferase and an RdRp[54].

Flaviviral polymerase NS5

The flaviviral NS5 protein is composed of two domains: an N terminal methyltransferase (MTase) and an RdRp at the C terminus[55]. The MTase domain has two enzymatic activities: methylation of the guanine N7 and methylation of the ribose-2′-O position to create a viral RNA cap[56]. The capping increases RNA stability, prevents 5′–3′ exoribonuclease degradation and helps to avoid the activation of innate immunity[57]. The RdRp domain is responsible for replication of viral RNA and its general fold is very similar to picornaviral $3D^{pol}$. The main difference between $3D^{pol}$ and

RdRp NS5 is the presence of a priming loop. The priming loop is located above the interface of the three channels (Template entry, NTP entry and dsRNA channel) and allows the formation of a dinucleotide primer[58]. The priming loop is crucial for *de novo* initiation [40,59].

There is no host cell enzyme which could substitute the RdRp function. The NS5 protein is a major target for drug discovery because both domains of NS5 (MTase and RdRp) are absolutely necessary for viral multiplication: So far only available structures of full length Flaviviral NS5 enzyme were from three viruses: ZIKV[60], JEV[61] and DENV[62].

YFV - epidemiology

YFV is endemic in subtropical areas of Africa and South America[63]. This virus is spread by mosquitoes to infect primates. The massive YFV epidemic spread from South America and Europe in the 18th to 20th centuries. In the early 20th effective vaccines against YFV were developed. Although the vaccination led to a decline in the disease, there are still outbreaks periodically[63,64].

Yellow fever disease manifests itself as fever, epigastric pain, hepatitis, jaundice and haemorrhage and is fatal in 60 % of the cases[65]. A treatment specific for yellow fever infection is not available.

The development of a specific treatment is certainly advantageous to fully understand the infection mechanism and to structurally describe the possible targets for antiviral therapy. So far only a handful of structures of YFV proteins are available: helicase[66], MTase[67], non-structural protein 1 (involved in virulence and replication)[51] and capsid[68].

ZIKV - epidemiology

ZIKV belongs to the mosquito-borne branch of the *Flaviviridae* family. ZIKV infection typically manifests itself as very mild causing rash and headache in most people. But, the infection can have teratogenic effects in pregnant women. The infection can be transmitted to the fetus and can cause microcephaly (head malformation)[69]. In severe cases in adults, Zika infections can also result in Guillain–Barré syndrome[70].

In 2015 there was a widespread epidemic of the Zika disease when it spread from Brazil to other parts of South America. Several countries advised their citizens to postpone pregnancy until the viral impact on fetal development was known [71].

One of possible ZIKV target is the NS5 protein, but, unfortunately, only the structure of ZiKV NS5 in the apo form is available[72–74].

Aims of the study

This study is focused on the polymerases from two +RNA viral families: *Picornaviridae* and *Flaviviridae*. The first aim was to study the protein−membrane complex, 3A−ACBD3−PI4KB. The viral 3A protein hijacks the Golgi ACBD3 protein in order to recruit PI4KB to the membranes of ROs. This results in PI4KB catalysing the phosphorylation of PI to PI4P. Consequently, the effect of PI4P-rich membranes on $3D^{pol}$ membrane localisation was explored. The specific aims were:

1. To reconstitute the protein complex 3A-ACBD3-PI4KB on the membrane *in vitro* using artificial membranes and recombinant proteins.
2. To investigate how the $3D^{pol}$ enzyme interacts with the membrane.

The picornaviral polymerases $3D^{pol}$ have an unusual activation mechanism, where the first residue plays a key role in N terminus stabilisation. However, we noted that some picornaviral $3D^{pol}$ enzymes do not have the conserved first glycine residue. Therefore, we decided:

3. To structurally characterise $3D^{pol}$ enzymes with a unique N terminus, using protein crystallography.

4. To characterize the effect of the very first $3D^{pol}$ residue on polymerase activity.

The second part of this study is related to NS5 polymerases from the *Flaviviridae* family, specifically on the YFV and ZIKV genera. Structural characterisation of possible antiviral targets is a constitutive step in the

development of specific disease treatments. Therefore, our specific aims were:

5. To solve the structure of the full-length YFV NS5 protein containing the MTase and RdRp domains.
6. To create a biologically relevant model of Zika NS5 in complex with RNA.

Publications included in doctoral thesis

Dubankova, A.[#], Humpolickova, J.[#], Klima, M. & Boura, E. Negative charge and membrane-tethered viral 3B cooperate to recruit viral RNA dependent RNA-polymerase $3D^{pol}$. Sci. Rep. 7, 17309 (2017).

Dubankova, A., Horova, V., Klima, M. & Boura, E. Structures of kobuviral and siciniviral polymerases reveal conserved mechanism of picornaviral polymerase activation. J. Struct. Biol. (2019). doi:10.1016/j.jsb.2019.08.004

Dubankova, A. & Boura, E. Structure of the yellow fever NS5 protein reveals conserved drug targets shared among flaviviruses. Antiviral Research 169, 104536 (2019).

Šebera, J.[#], **Dubankova, A.**[#], Sychrovský, V., Ruzek, D., Boura, E. & Nencka, R. The structural model of Zika virus RNA-dependent RNA polymerase in complex with RNA for rational design of novel nucleotide inhibitors. Sci. Rep. 8, 11132 (2018).

# both authors contributed equally

Results overview

Negative charge and membrane-tethered viral 3B cooperate to recruit viral RNA-dependent RNA polymerase $3D^{pol}$.

Preview

PI4KB is a host factor essential for many +RNA viruses. It was shown that a selective inhibition of PI4KB has antiviral activity[13,14] but the exact role of PI4P in viral ROs is still poorly understood. One of the hypotheses suggested that $3D^{pol}$ polymerase is recruited to these hyperphosphorylated membranes by a specific interaction with PI4P[15]. In our publication "Negative charge and membrane-tethered viral 3B cooperate to recruit viral RNA-dependent RNA polymerase $3D^{pol}$." We showed using an *in vitro* reconstitution approach that not PI4P specifically but rather a negative charge in general recruits $3D^{pol}$ to the membrane.

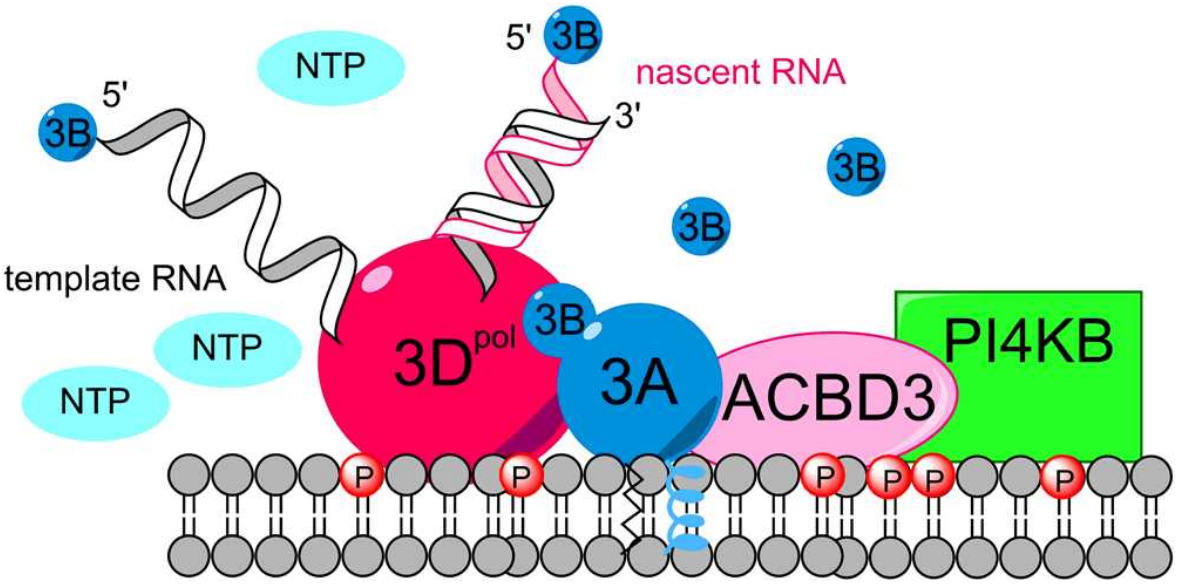
Summary

We reconstituted PI4KB membrane recruitment step-by-step *in vitro* that occurs in enterovirus- and kobuvirus-infected cells[75]. Giant unilamellar vesicles were used as artificial biomimetic membranes for the reconstitution of the 3A-ACBD3-PI4KB complex on their surface. We used recombinant proteins with fluorescent label for visualisation of proteins decorated membranes.

The 3A-ACBD3-PI4KB complex formation was well described by *in vivo* experiments[75]. Our *in vitro* experiments are in agreement with these conclusions. We demonstrated that the viral protein 3A recruits PI4KB to the membrane via ACBD3.

The PI4KB catalyses phosphorylation of PI to create PI4P. We observed that PI4P enrichment in the membrane does not lead to $3D^{pol}$

membrane recruitment. But, we observed $3D^{pol}$ membrane binding after addition of membrane-tethered 3B protein (mimicking 3AB intermediate). After a series of control experiments we observed that the $3D^{pol}$ polymerase is recruited to negatively charged membranes but not to neutral membranes in the presence of membrane-tethered 3B. A hypothetical model of replication complex is demonstrated in picture 6.



Picture 6: *Schematic model of the $3D^{pol}$-3AB-ACBD3-PI4KB protein complex*[76].

My contribution

I prepared expression constructs for $3D^{pol}$ proteins ($3D^{pol}$ from Aichi virus, PV, CV and Enterovirus 71). I recombinantly expressed and purified $3D^{pol}$ proteins. I performed liposome pull-down assays for $3D^{pol}$ membrane binds testing. I also expressed and purified the SidC protein (a PI4P biosensor). I verified a SidC−PI4P interaction using biochemical assays. I prepared the figures for publication and assisted in preparing the manuscript.

Structures of kobuviral and siciniviral polymerases reveal a conserved mechanism of picornaviral polymerase activation.

Preview

Picornaviral polymerases $3D^{pol}$ are activated by burial of their N terminus to a cavity near the protein active site. The stabilisation of the N terminus is mediated by the very first residue. So far all the structures of picornaviral polymerase available have a glycine as the first residue. In this publication we provided extensive alignment of picornaviral $3D^{pol}$ enzymes and we show that seven genera have a serine as the first residue. In the publication “Structures of kobuviral and siciniviral polymerases reveal conserved mechanism of picornaviral polymerase activation.” we showed $3D^{pol}$ structures containing an uncanonical N terminus and the importance of the first residue in $3D^{pol}$ for viral replication.

Summary

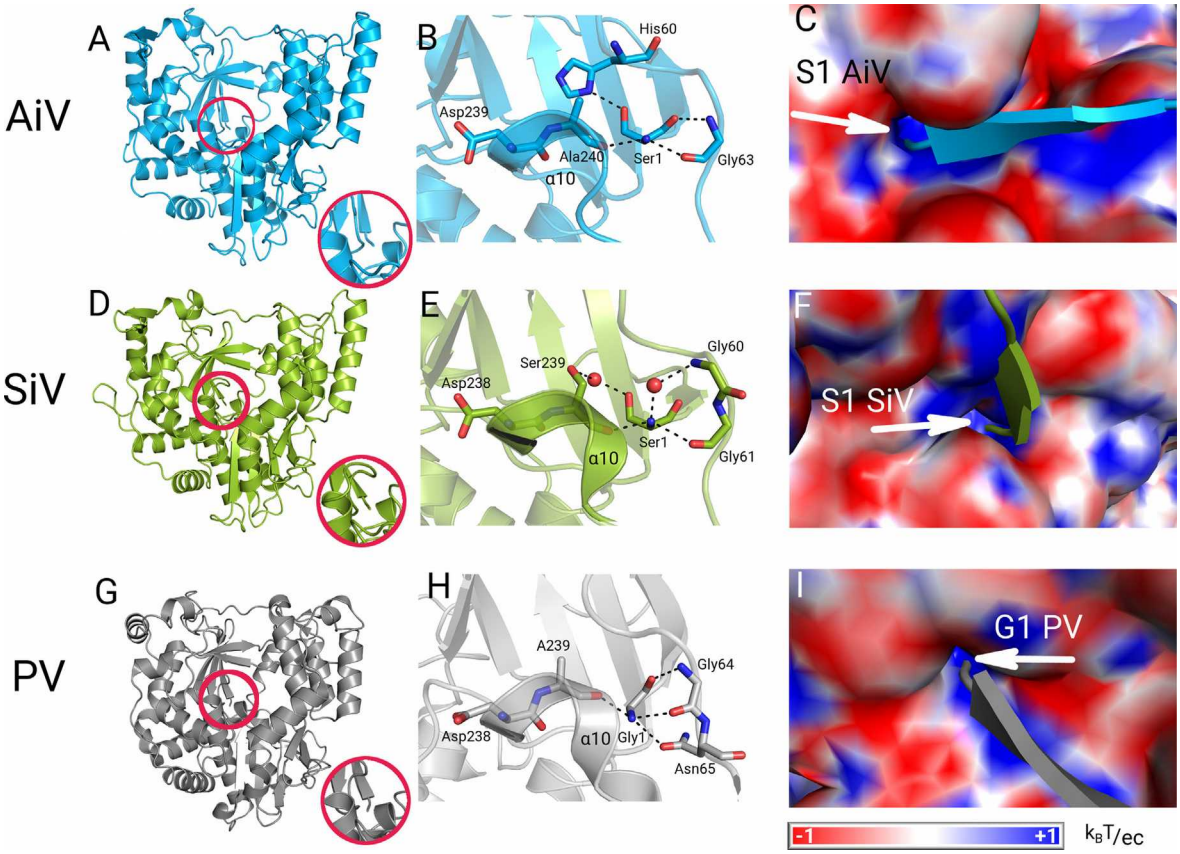
Sequential alignment of picornaviral $3D^{pol}$ enzymes uncovered seven genera with uncanonical residues as the first residue at the N-terminus (galliviruses, osciviruses, passeriviruse, sakobuviruses, saliviruses, kobuviruses and siciniviruses).

We solved crystal structures of kobuviral and siciniviral $3D^{pol}$ polymerases. Both have their N terminus buried in a surface pocket similar to glycine-beginning polymerases. The surface pocket of kobuviral and siciniviral $3D^{pol}$ is evolutionarily designed for interaction with the first serine residue. The hydroxyl group of Ser1 from kobuviral $3D^{pol}$ interacts with His60.

The hydroxyl group of the first serine from siciniviral $3D^{pol}$ creates a hydrogen bond and a water bridge with the Ser239. A detailed view of the

kobuviral and siciniviral polymerase in comparison to a PV polymerase is shown in the picture 7.

We tested the specificity of the N terminus binding pocket by a variety of mutations in the $3D^{pol}$ first residue. We checked mutation of the first serine to glycine residue which is conserved in the majority of $3D^{pol}$s. This mutation, as well as every other mutation tested, disrupted viral replication.



Picture 7: *Atypical fold of kobuviral (Aichi virus - AiV) and siciniviral (SiV) $3D^{pol}$ N terminus within $3D^{pol}$ enzymes[27].*

A*–*C *The first residue of AiV $3D^{pol}$ is locked into a positively charged cavity in proximity to the protein active site.* ***D*–*F*** *Corresponding view of the SiV $3D^{pol}$ structure.* ***G*–*I*** *The majority of picornaviruses has the very first glycine residue inserted in a charged cavity as illustrated by the PV $3D^{pol}$ structure[77].*

My contribution

I performed bioinformatic analysis of all picornaviral $3D^{pol}$ sequences. I prepared expression constructs for kobuviral as well as for siciniviral $3D^{pol}$ and performed mutagenesis experiments. I expressed and purified all recombinant proteins. I performed all crystallization experiments. I collected and analysed all crystallographic data, I prepared figures for the manuscript and assisted in preparing the manuscript.

Structure of the yellow fever NS5 protein reveals conserved drug targets shared among flaviviruses.

Preview

YFV NS5 is composed of two domains: MTase and RdRp. Both are important targets in drug design. So far the only available structure was the structure of YFV MTase[67]. The publication “Structure of the yellow fever NS5 protein reveals conserved drug targets shared among flaviviruses.” provides the first structure of a full-length YFV NS5 protein containing both the MTase and RdRp domains.

Summary

We solved the crystal structure of the full-length yellow fever NS5 enzyme at 3.1 Å resolution. We discovered regions that are structurally conserved in the *Flaviviridea* family. These regions are pivotal for structure-based drug design.

The priming loop is essential in the initiation of RNA replication. A residue responsible for a stacking interaction with the initiating nucleotide is localised in the priming loop. The YFV NS5 has this residue (Trp799) conserved with ZIKV NS5[78], but different to HCV[79]. YFV NS5 is stabilised by

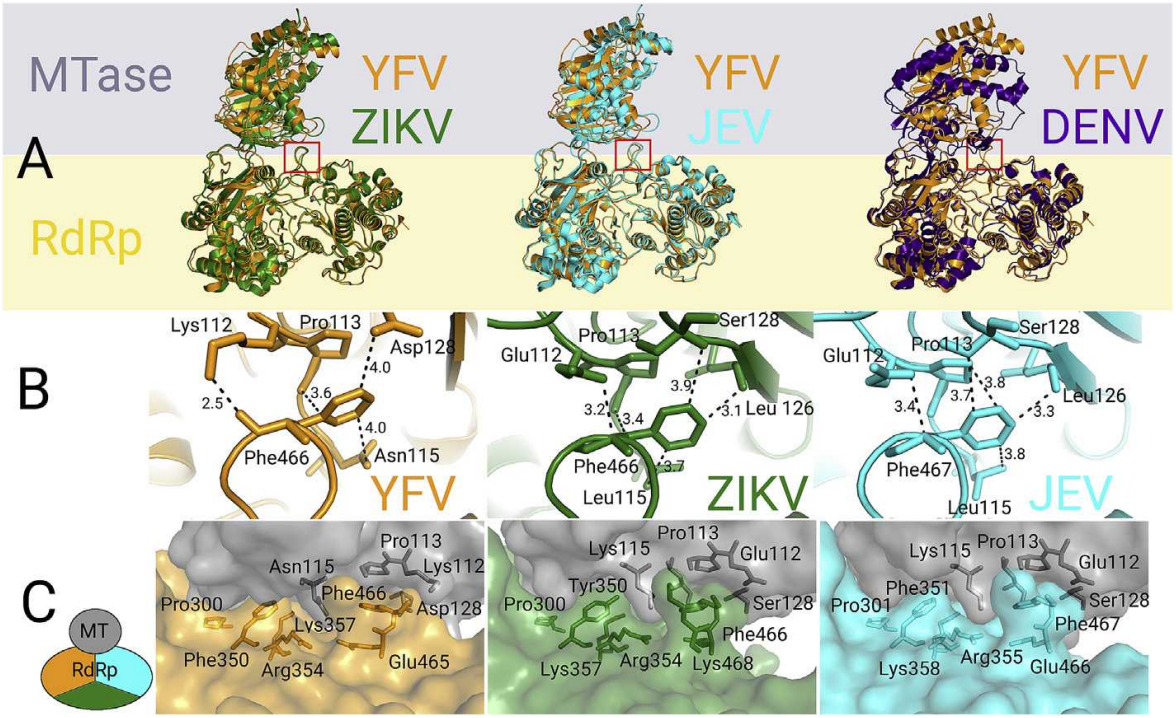
two zinc fingers. The first zinc ion site is coordinated by two cysteine (Cys448 and Cys451), histidine 443 and glutamate 439 residues. Except for glutamate 439 the zinc ion binding sites are absolutely conserved among NS5 proteins from YFV, ZIKV, DENV, and WNV.

The second zinc ion site is coordinated by two cysteine (Cys732 and Cys851) and two histidine (His716 and His718) residues. This zinc ion site is also highly conserved except for histidine 716 which is variable in ZIKV and WNV, but conserved in DENV.

The interdomain contact between MTase and RdRp is mediated by a highly conserved proline 113 (from the MTase domain) and phenylalanine 466 (from the RdRp domain) which is packed in a pocket on the MTase surface. The Phe466 residue is located in structural motif F. Both domains are stabilised by this interaction and we showed that the full-length YFV NS5 is more thermostable than the RdRp or MTase domain only.

A comparison of mutual domain orientations revealed that the YFV NS5 has the same orientation as ZIKV and JEV, but different than DENV NS5.

A detailed view of the MTase RdRp interface comparisons is displayed in picture 8.



Picture 8: *The MTase–RdRp interface.* [80].

***A** Structures of full length flaviviral NS5 proteins were superimposed on the structure of YF NS5 based on the RdRp domain. This revealed that the domain orientation is conserved in yellow fever, Zika and Japanese encephalitis virus RdRp but not in Dengue RdRp. **B** Conserved residues Phe466 and Pro113 define the MTase–RdRp interface. **C** Detail of the RdRp MTase interface. MTase is shown in gray. RdRp is colored depending on the viral origin (YFV–yellow, ZIKV–forest green, JEV–cyan).*

My contribution

I performed all experiments, analysed all data and assisted in preparing the manuscript preparation.

The structural model of Zika virus RNA-dependent RNA polymerase in complex with RNA for rational design of novel nucleotide inhibitors.

Preview

As mentioned, the flaviviral NS5 protein is an extensively studied target for drug design, especially in the case of ZIKV. So far all available structures of ZIKV NS5[73] are in the apo form . Our efforts to obtain a crystal structure of the ZIKV NS5 in complex with RNA were unsuccessful therefore we decided to prepare a sophisticated model that could be experimentally validated. In the publication "The structural model of Zika virus RNA-dependent RNA polymerase in complex with RNA for rational design of novel nucleotide inhibitors." we introduced a model of ZIKV polymerase in complex with RNA that proved to be very useful for docking studies of ZIKV RdRp inhibitors.

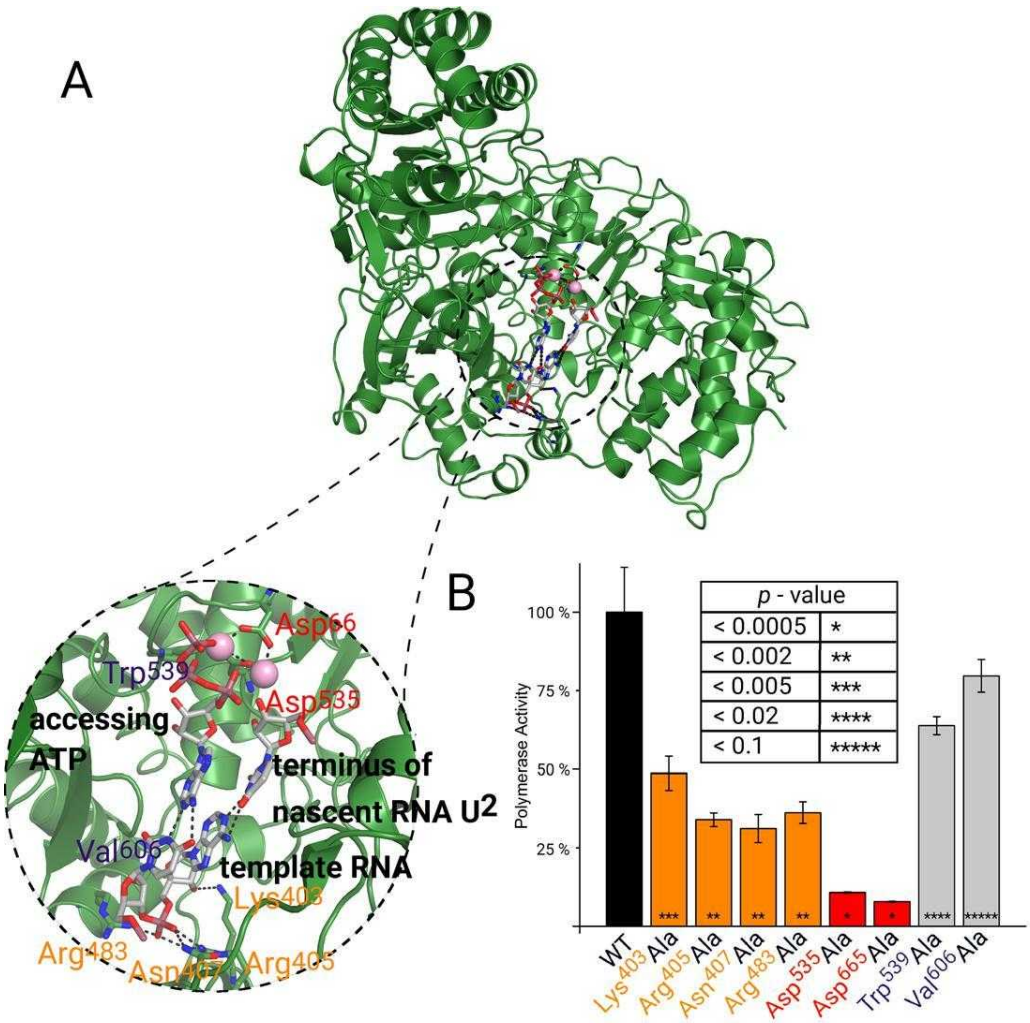
Summary

In this study we prepared a model of ZIKV NS5 in complex with a small fragment of RNA. The model was built on foundations of a crystal structure of the full length NS5 ZIKV[73] and the related HCV RdRp in a complex with ADP, Mn^{2+}, and a primer−template hairpin[58]. Our model proved to be essential for docking experiments necessary for rational inhibitor design. The calculated model contains ZIKV NS5, template RNA (represented by a dinucleotide) accessing ATP two magnesium ions and a terminus of the nascent RNA.

The validity of this model was verified by polymerase activity assays on single point-mutated ZIKV NS5. The mutations in Mg^{2+}-coordinating aspartates fully disrupt polymerase activity as predicted. Mutations in close

proximity to the polymerase active site had significantly lower activity compared to the wild type, see picture 9.

Consequently, we used the model for molecular docking experiments. First, we redocked the ATP molecule and compared the results with the original model. The position of the re-docked ATP nicely correlates with its original position. Consequently, we docked the NITD008 molecule which is a known ZIKV RdRp inhibitor[81]. This docking experiment revealed that NIT008 binds to the active site of the ZIKV RdRp with minimum distortion in comparison to the ATP molecule. Importantly, the same docking experiments performed on the ZIKV NS5 enzyme in the apo form always fail to reveal the binding mode of any nucleoside triphosphate.



Picture 9: *Enzymatic analysis of ZIKV NS5 model with RNA[78].*

A *ZIKV RdRp RNA complex model with a zoomed-in RNA binding site. The RNA template is a dinucleotide composed of an A – U. Nascent U[2] and ATP are present. In the enlarged picture the amino acids related to activity assay are highlighted.* ***B*** *Activity of ZIKV RdRp: Mutations in the red region disrupted polymerase activity, in the orange regions there was significantly decreased activity and in the gray regions slightly accentuated activity. Values in the box are Student t-test p-values.*

My contribution

I performed all biological and biochemical experiments, prepared the figures, and participated in the preparing the manuscript.

Discussion

Negative charge and membrane-tethered viral 3B cooperate to recruit viral RNA-dependent RNA polymerase $3D^{pol}$.

How the PI4KB kinase is hijacked was suggested based on *in vivo* experiments[75]. The *in vitro* approach we used confirmed this hypothesis and provided deeper molecular insight. We reconstituted the formation of the 3A-ACBD3-PI4KB complex on artificial membranes, where PI4KB phosphorylates PI to produce PI4P.

The role of PI4P on RO membranes is only poorly understood. One of the hypotheses suggested that the viral polymerase $3D^{pol}$ specifically interacts with PI4P and thus PI4P anchors the $3D^{pol}$ enzyme on the membrane[15]. Our results show that for $3D^{pol}$ membrane recruitment a membrane-tethered 3B protein is required together with negatively-charged lipids (e.g. PI4P or PS).

The PI4P role in ROs is probably more complex. For example, PI4P can arrange recruitment of host cell effectors to the replication complex or it can be exchanged for cholesterol by lipid transport proteins.

Structures of kobuviral and siciniviral polymerases reveal a conserved mechanism of picornaviral polymerase activation.

The majority of picornaviral polymerases $3D^{pol}$ have a glycine as their first residue, however, we found seven genera with an unusual N terminus starting with a serine residue. The $3D^{pol}$ gains its polymerization activity after cleavage of the 3CD precursor into 3C and $3D^{pol}$ proteins[37,82]. Afterwards, the first residue is buried in a surface pocket and activates the polymerase. We show that the N terminus is the least conserved region compared to other picornaviruses. In this study, we also show structural differences in picornaviral polymerases with the atypical N-terminus. We selected the kobuviral and siciniviral $3D^{pol}$ enzymes for our crystallization experiments. The structures revealed N-terminus stabilization by interaction of first residue with protein core.

In principle, the stabilization mechanism is the same as in other picornaviral generas[44,45,46,47,48,49.] The N-terminus binding pocket specifically recognises the first amino acid which leads to activation of the enzyme. However, the atomic details differ. The presence of an extra hydroxyl group in the first residue provides different N-terminal stabilization. The kobuviral Ser1 hydroxyl group creates a hydrogen bond with His60. The hydroxyl of the first serine from siciniviral $3D^{pol}$ creates a hydrogen bond and a water bridge with Ser239.

Congruently, our phylogenetic analysis based on the structures of available polymerases revealed that kobuviruses and siciniviruses are an evolutionarily distinct group within the *Picornaviridae* family.

Structure of the yellow fever NS5 protein reveals conserved drug targets shared among flaviviruses.

YFV is a grave infection spread by mosquitoes. Although an effective vaccination is available, there are still many outbreaks and an antiviral therapy is clearly needed. NS5 is an extensively studied protein. It consists of two domains, an MTase and RdRp domain. Both domains are *bona fide* targets for development of antiviral therapies.

In this study we presented a crystal structure of the full-length YFV NS5 enzyme for the very first time. Consequently, we compared YFV NS5 with ZIKV[60], JEV[61] and DENV NS5[62] structures which were already solved. YFV NS5 has conserved interface between MTase and RdRp with ZIKV and JEV but not with DENV NS5.

The F motif that is essential for the RdRp enzymatic activity is localized to the MTase RdRp interface and stabilized by this interface. Moreover, the interface stabilizes the MTase active site. Our structural analysis suggested that not only the RdRp activity is stimulated by the MTase domain [72], but also the MTase domain is stabilized by the RdRp domain.

A structural model of Zika virus RNA-dependent RNA polymerase in complex with RNA for rational design of novel nucleotide inhibitors.

The ZIKV epidemic in 2015 increased public awareness of health risks connected with ZIKV infection. For pregnant women ZIKV infection significantly increases the risk of fetal malformation.

ZIKV has been extensively studied since 2015. The ZIKV polymerase NS5 is a prime drug design target therefore, the structural information is crucial to effective drug design[58,73]. The structure was solved recently by several groups both from academia and industry[60,73].Unfortunately, all the structures solved are in the apo form and, therefore, not useful for docking experiments with NTP analogues because base pairing is paramount to an effective strategy. No crystal structure of flaviviral polymerase in complex with RNA is available. Fortunately, the structure of a sufficiently related HCV virus NS5 protein in complex with RNA[58] could be used together with the ZIKV NS5 apo structure[73] to devise a model of ZIKV polymerase in complex with RNA.

In this study, we used these two structures to calculate quantum mechanics / molecular mechanics (QM/MM) model of the ZIKV polymerase in complex with RNA. Our model shows template RNA, accessing ATP and the terminus of nascent RNA. We evaluated our model by biochemical assays and proved its usefulness in docking experiments.

Conclusions

This study contributed to our understanding of picornaviral RNA replication.

We confirmed establishment of a PI4K hijacking protein membrane complex using *in vitro* experiments, pure recombinant proteins and artificial membranes.

We showed that not PI4P, specifically, but, rather, a negative charge, in general, recruits $3D^{pol}$ to the membrane.

The structural study of picornaviral polymerases reveals a new perspective on $3D^{pol}$ activation. We solved structures of kobuviral and siciniviral polymerase $3D^{pol}$ with a unique N terminus and structurally described polymerase activation. We demonstrated the influence of the very first residue on polymerase activity.

Our results also contributed to potential polymerase-targeted rational drug design by providing YFV and ZIKV polymerase structures.

We solved the first structure of YFV NS5 polymerase. We compared the structure of YFV to other known NS5 structures to revealed structural similarities and differences. The YFV NS5 has the stacking residue of priming loop (Trp799) conserved with ZIKV NS5, but different to HCV.

The YFV NS5 is stabilised by two zinc fingers. The first zinc ion binding site is absolutely conserved among NS5 proteins except for the glutamate 439 residue. The second zinc ion binding site is also highly conserved except for histidine 716 residue which is various in ZIKV and WNV, but conserved in DENV.

A comparison of mutual domains orientation revealed that the YFV NS5 has the same orientation as ZIKV and JEV, but different than DENV

NS5.

We calculated a biologically-relevant model of ZIKV NS5 in complex with RNA-containing ZIKV NS5, template RNA (represented by a dinucleotide) accessing ATP two magnesium ions and a terminus of the nascent RNA. We evaluated the model using biological assays and were verified by ligand docking experiments.

List of results:

1. We reconstituted *in vitro* formation of the 3A-ACBD3-PI4KB protein complex on the surface of biomimetic membranes.
2. We observed that a picornaviral polymerase $3D^{pol}$ is recruited to the membrane surface in the presence of PI4P or other negatively-charged lipids and membrane-tethered 3B.
3. The sequence alignment of picornaviral $3D^{pol}$ polymerases revealed seven genera with a non-conserved first residue. We choose kobuviral and siciniviral $3D^{pol}$ enzymes for further structural analysis.
4. We solved the crystal structures of kobuviral and siciniviral $3D^{pol}$ polymerases. The structures revealed N-terminal stabilisation by the first residue. We also showed that the structures of kobuvirus and sicinivirus polymerases are evolutionarily distinct to other known picornaviral polymerases.
5. Using biological assays, we demonstrated the vitality of the correct first residue. We show that (from a tested set of mutations) only the original N-terminus activates the polymerase.
6. We solved the first crystal structure of YFV NS5. The structure revealed similarities and differences of YFV NS5 to others known flaviviral NS5 structures.
7. We calculated a biologically relevant model of ZIKV polymerase in complex with RNA for ligand docking studies.

Abbreviation

+RNA	Plus Single-stranded RNA
ACBD3	acyl-CoA-binding domain-containing protein-3
AiV	Aichi virus
CV	Coxsackie virus
DENV	Dengue virus
dsRNA	Double-stranded RNA
ER	Endoplasmic reticulum
HCV	Hepatitis C virus
HRV	Human rhinovirus
IRES	Internal ribosomal entry site
JEV	Japanese encephalitis virus
L	Leader protein
mRNA	Messenger RNA
MTase	Methyltransferase
NTP	Nucleotide triphosphate
PI	Phosphatidylinositol
PI4KB	Phosphatidylinositol 4-kinase III β
PI4P	Phosphatidylinositol 4-phosphate

PS	Phosphatidylserine
PV	Poliovirus
QM/MM	Quantum mechanics / Molecular mechanics
RdRp	RNA dependent RNA-polymerase
RO	Replication organelle
SiV	Sicini virus
TBE	Tick-borne encephalitis virus
UMP	Uracil monophosphate
VP	viral protein
VPG	viral protein genome-linked
WNV	West Nile virus
YFV	Yellow fever virus
ZIKV	Zika virus

Bibliography

1. 1 About Infectious Disease | Infectious Diseases. Available at: http://infectiousdiseases.edc.org/content/module/1/reading/2. (Accessed: 26th August 2019)
2. den Boon, J. A. & Ahlquist, P. Organelle-like membrane compartmentalization of positive-strand RNA virus replication factories. *Annu. Rev. Microbiol.* **64**, 241–256 (2010).
3. Knipe, D. M. *et al.* Fundamental Virology, 4th Edition; and Fields Virology, 4th Edition, Volumes I and II:Fundamental Virology, 4th Edition;Fields Virology, 4th Edition, Volumes I and II. *Clinical Infectious Diseases* **34**, 1029–1030 (2002).
4. Belov, G. A. *et al.* Complex dynamic development of poliovirus membranous replication complexes. *J. Virol.* **86**, 302–312 (2012).
5. Monine, M. I. & Haugh, J. M. Reactions on cell membranes: Comparison of continuum theory and Brownian dynamics simulations. *The Journal of Chemical Physics* **123**, 074908 (2005).
6. Ando, T. *et al.* Visualization and measurement of ATP levels in living cells replicating hepatitis C virus genome RNA. *PLoS Pathog.* **8**, e1002561 (2012).
7. Paul, D. & Bartenschlager, R. Architecture and biogenesis of plus-strand RNA virus replication factories. *World J Virol* **2**, 32–48 (2013).
8. Kato, H., Takahasi, K. & Fujita, T. RIG-I-like receptors: cytoplasmic sensors for non-self RNA. *Immunological Reviews* **243**, 91–98 (2011).
9. Irurzun, A., Perez, L. & Carrasco, L. Involvement of membrane traffic in the replication of poliovirus genomes: effects of brefeldin A. *Virology* **191**, 166–175 (1992).

10. Greninger, A. L., Knudsen, G. M., Betegon, M., Burlingame, A. L. & Derisi, J. L. The 3A protein from multiple picornaviruses utilizes the golgi adaptor protein ACBD3 to recruit PI4KIIIβ. *J. Virol.* **86**, 3605–3616 (2012).
11. Sasaki, J., Ishikawa, K., Arita, M. & Taniguchi, K. ACBD3-mediated recruitment of PI4KB to picornavirus RNA replication sites. *The EMBO Journal* **31**, 754–766 (2012).
12. Fan, J., Liu, J., Culty, M. & Papadopoulos, V. Acyl-coenzyme A binding domain containing 3 (ACBD3; PAP7; GCP60): an emerging signaling molecule. *Prog. Lipid Res.* **49**, 218–234 (2010).
13. Lamarche, M. J. *et al.* Anti-hepatitis C virus activity and toxicity of type III phosphatidylinositol-4-kinase beta inhibitors. *Antimicrob. Agents Chemother.* **56**, 5149–5156 (2012).
14. Rutaganira, F. U. *et al.* Design and Structural Characterization of Potent and Selective Inhibitors of Phosphatidylinositol 4 Kinase IIIβ. *J. Med. Chem.* **59**, 1830–1839 (2016).
15. Hsu, N.-Y. *et al.* Viral reorganization of the secretory pathway generates distinct organelles for RNA replication. *Cell* **141**, 799–811 (2010).
16. Peersen, O. B. Picornaviral polymerase structure, function, and fidelity modulation. *Virus Res.* **234**, 4–20 (2017).
17. D'Angelo, G., Vicinanza, M., Di Campli, A. & De Matteis, M. A. The multiple roles of PtdIns(4)P - not just the precursor of PtdIns(4,5)P2. *Journal of Cell Science* **121**, 1955–1963 (2008).
18. Yamashita, S.-I., Oku, M., Wasada, Y., Ano, Y. & Sakai, Y. PI4P-signaling pathway for the synthesis of a nascent membrane structure in selective autophagy. *J. Cell Biol.* **173**, 709–717 (2006).

19. Blumental-Perry, A. *et al.* Phosphatidylinositol 4-phosphate formation at ER exit sites regulates ER export. *Dev. Cell* **11**, 671–682 (2006).
20. McMahon, H. T. & Gallop, J. L. Membrane curvature and mechanisms of dynamic cell membrane remodelling. *Nature* **438**, 590–596 (2005).
21. Clayton, E. L., Minogue, S. & Waugh, M. G. Mammalian phosphatidylinositol 4-kinases as modulators of membrane trafficking and lipid signaling networks. *Prog. Lipid Res.* **52**, 294–304 (2013).
22. Tong, J., Yang, H., Yang, H., Eom, S. H. & Im, Y. J. Structure of Osh3 reveals a conserved mode of phosphoinositide binding in oxysterol-binding proteins. *Structure* **21**, 1203–1213 (2013).
23. Chung, J. *et al.* INTRACELLULAR TRANSPORT. PI4P/phosphatidylserine countertransport at ORP5- and ORP8-mediated ER-plasma membrane contacts. *Science* **349**, 428–432 (2015).
24. Moser von Filseck, J., Vanni, S., Mesmin, B., Antonny, B. & Drin, G. A phosphatidylinositol-4-phosphate powered exchange mechanism to create a lipid gradient between membranes. *Nat. Commun.* **6**, 6671 (2015).
25. Dorobantu, C. M. *et al.* Modulation of the Host Lipid Landscape to Promote RNA Virus Replication: The Picornavirus Encephalomyocarditis Virus Converges on the Pathway Used by Hepatitis C Virus. *PLoS Pathog.* **11**, e1005185 (2015).
26. Paul, A. V. & Wimmer, E. Initiation of protein-primed picornavirus RNA synthesis. *Virus Res.* **206**, 12–26 (2015).
27. Dubankova, A., Horova, V., Klima, M. & Boura, E. Structures of kobuviral and siciniviral polymerases reveal conserved mechanism of picornaviral polymerase activation. *J. Struct. Biol.* (2019). doi:10.1016/j.jsb.2019.08.004
28. Basavappa, R. *et al.* Role and mechanism of the maturation cleavage of VP0 in poliovirus assembly: structure of the empty capsid assembly intermediate at 2.9 A resolution. *Protein Sci.* **3**, 1651–1669 (1994).

29. Brandenburg, B. *et al.* Imaging poliovirus entry in live cells. *PLoS Biol.* **5**, e183 (2007).
30. Castelló, A., Izquierdo, J. M., Welnowska, E. & Carrasco, L. RNA nuclear export is blocked by poliovirus 2A protease and is concomitant with nucleoporin cleavage. *J. Cell Sci.* **122**, 3799–3809 (2009).
31. Sandoval, I. V. & Carrasco, L. Poliovirus infection and expression of the poliovirus protein 2B provoke the disassembly of the Golgi complex, the organelle target for the antipoliovirus drug Ro-090179. *J. Virol.* **71**, 4679–4693 (1997).
32. Rodríguez, P. L. & Carrasco, L. Poliovirus protein 2C has ATPase and GTPase activities. *J. Biol. Chem.* **268**, 8105–8110 (1993).
33. Gangaramani, D. R., Eden, E. L., Shah, M. & Destefano, J. J. The twenty-nine amino acid C-terminal cytoplasmic domain of poliovirus 3AB is critical for nucleic acid chaperone activity. *RNA Biol.* **7**, 820–829 (2010).
34. Belov, G. A. *et al.* Hijacking components of the cellular secretory pathway for replication of poliovirus RNA. *J. Virol.* **81**, 558–567 (2007).
35. Pelletier, J. & Sonenberg, N. Internal initiation of translation of eukaryotic mRNA directed by a sequence derived from poliovirus RNA. *Nature* **334**, 320–325 (1988).
36. Kean, K. M., Howell, M. T., Grünert, S., Girard, M. & Jackson, R. J. Substitution mutations at the putative catalytic triad of the poliovirus 3C protease have differential effects on cleavage at different sites. *Virology* **194**, 360–364 (1993).
37. Harris, K. S., Reddigari, S. R., Nicklin, M. J., Hämmerle, T. & Wimmer, E. Purification and characterization of poliovirus polypeptide 3CD, a proteinase and a precursor for RNA polymerase. *J. Virol.* **66**, 7481–7489 (1992).
38. Ferrer-Orta, C. *et al.* The structure of a protein primer-polymerase complex in the initiation of genome replication. *EMBO J.* **25**, 880–888 (2006).

39. Strebel, K. & Beck, E. A second protease of foot-and-mouth disease virus. *J. Virol.* **58**, 893–899 (1986).

40. Wang, C., Wang, C., Li, Q., Wang, Z. & Xie, W. Crystal Structure and Thermostability Characterization of Enterovirus D68 3D pol. *Journal of Virology* **91**, (2017).

41. Gohara, D. W., Arnold, J. J. & Cameron, C. E. Poliovirus RNA-dependent RNA polymerase (3Dpol): kinetic, thermodynamic, and structural analysis of ribonucleotide selection. *Biochemistry* **43**, 5149–5158 (2004).

42. Marcotte, L. L. *et al.* Crystal structure of poliovirus 3CD protein: virally encoded protease and precursor to the RNA-dependent RNA polymerase. *J. Virol.* **81**, 3583–3596 (2007).

43. Thompson, A. A. & Peersen, O. B. Structural basis for proteolysis-dependent activation of the poliovirus RNA-dependent RNA polymerase. *EMBO J.* **23**, 3462–3471 (2004).

44. Hansen, J. L., Long, A. M. & Schultz, S. C. Structure of the RNA-dependent RNA polymerase of poliovirus. *Structure* **5**, 1109–1122 (1997).

45. Campagnola, G., Weygandt, M., Scoggin, K. & Peersen, O. Crystal structure of coxsackievirus B3 3Dpol highlights the functional importance of residue 5 in picornavirus polymerases. *J. Virol.* **82**, 9458–9464 (2008).

46. Vives-Adrian, L. *et al.* The crystal structure of a cardiovirus RNA-dependent RNA polymerase reveals an unusual conformation of the polymerase active site. *J. Virol.* **88**, 5595–5607 (2014).

47. Ferrer-Orta, C. *et al.* Multifunctionality of a picornavirus polymerase domain: nuclear localization signal and nucleotide recognition. *J. Virol.* **89**, 6848–6859 (2015).

48. Love, R. A. *et al.* The crystal structure of the RNA-dependent RNA polymerase from human rhinovirus: a dual function target for common cold antiviral therapy. *Structure* **12**, 1533–1544 (2004).

49. Wu, Y. *et al.* Structures of EV71 RNA-dependent RNA polymerase in complex with substrate and analogue provide a drug target against the hand-foot-and-mouth disease pandemic in China. *Protein Cell* **1**, 491–500 (2010).
50. Rastogi, M., Sharma, N. & Singh, S. K. Flavivirus NS1: a multifaceted enigmatic viral protein. *Virol. J.* **13**, 131 (2016).
51. Wang, H. *et al.* Crystal structure of the C-terminal fragment of NS1 protein from yellow fever virus. *Sci. China Life Sci.* **60**, 1403–1406 (2017).
52. Xie, X., Zou, J., Puttikhunt, C., Yuan, Z. & Shi, P.-Y. Two distinct sets of NS2A molecules are responsible for dengue virus RNA synthesis and virion assembly. *J. Virol.* **89**, 1298–1313 (2015).
53. Muñoz-Jordán, J. L. *et al.* Inhibition of alpha/beta interferon signaling by the NS4B protein of flaviviruses. *J. Virol.* **79**, 8004–8013 (2005).
54. Issur, M. *et al.* The flavivirus NS5 protein is a true RNA guanylyltransferase that catalyzes a two-step reaction to form the RNA cap structure. *RNA* **15**, 2340–2350 (2009).
55. Selisko, B., Papageorgiou, N., Ferron, F. & Canard, B. Structural and Functional Basis of the Fidelity of Nucleotide Selection by Flavivirus RNA-Dependent RNA Polymerases. *Viruses* **10**, (2018).
56. Dong, H. *et al.* Biochemical and genetic characterization of dengue virus methyltransferase. *Virology* **405**, 568–578 (2010).
57. Hyde, J. L. & Diamond, M. S. Innate immune restriction and antagonism of viral RNA lacking 2′-O methylation. *Virology* **479-480**, 66–74 (2015).
58. Appleby, T. C. *et al.* Viral replication. Structural basis for RNA replication by the hepatitis C virus polymerase. *Science* **347**, 771–775 (2015).

59. Malet, H. *et al.* Crystal structure of the RNA polymerase domain of the West Nile virus non-structural protein 5. *J. Biol. Chem.* **282**, 10678–10689 (2007).

60. Wang, B. *et al.* The structure of Zika virus NS5 reveals a conserved domain conformation. *Nat. Commun.* **8**, 14763 (2017).

61. Lu, G. & Gong, P. Crystal Structure of the full-length Japanese encephalitis virus NS5 reveals a conserved methyltransferase-polymerase interface. *PLoS Pathog.* **9**, e1003549 (2013).

62. Zhao, Y. *et al.* A crystal structure of the Dengue virus NS5 protein reveals a novel inter-domain interface essential for protein flexibility and virus replication. *PLoS Pathog.* **11**, e1004682 (2015).

63. Barrett, A. D. T. & Higgs, S. Yellow fever: a disease that has yet to be conquered. *Annu. Rev. Entomol.* **52**, 209–229 (2007).

64. Monath, T. P. *et al.* Yellow fever vaccine supply: a possible solution. *Lancet* **387**, 1599–1600 (2016).

65. Tuboi, S. H., Costa, Z. G. A., da Costa Vasconcelos, P. F. & Hatch, D. Clinical and epidemiological characteristics of yellow fever in Brazil: analysis of reported cases 1998-2002. *Trans. R. Soc. Trop. Med. Hyg.* **101**, 169–175 (2007).

66. Wu, J., Bera, A. K., Kuhn, R. J. & Smith, J. L. Structure of the Flavivirus helicase: implications for catalytic activity, protein interactions, and proteolytic processing. *J. Virol.* **79**, 10268–10277 (2005).

67. Geiss, B. J. *et al.* Analysis of flavivirus NS5 methyltransferase cap binding. *J. Mol. Biol.* **385**, 1643–1654 (2009).

68. Zhang, Y. *et al.* Structures of immature flavivirus particles. *EMBO J.* **22**, 2604–2613 (2003).

69. Website, N. H. S. Zika virus. *nhs.uk* Available at: https://www.nhs.uk/conditions/zika/. (Accessed: 29th August 2019)

70. Méndez, N., Oviedo-Pastrana, M., Mattar, S., Caicedo-Castro, I. & Arrieta, G. Zika virus disease, microcephaly and Guillain-Barré syndrome in Colombia: epidemiological situation during 21 months of the Zika virus outbreak, 2015–2017. *Archives of Public Health* **75**, (2017).
71. Diniz, S. G. Zika virus and pregnancy: A perspective from Brazil. *Midwifery* **35**, 22–23 (2016).
72. Zhao, B. *et al.* Structure and function of the Zika virus full-length NS5 protein. *Nat. Commun.* **8**, 14762 (2017).
73. Upadhyay, A. K. *et al.* Crystal structure of full-length Zika virus NS5 protein reveals a conformation similar to Japanese encephalitis virus NS5. *Acta Crystallogr. Sect. F Struct. Biol. Cryst. Commun.* **73**, 116–122 (2017).
74. Godoy, A. S. *et al.* Crystal structure of Zika virus NS5 RNA-dependent RNA polymerase. *Nat. Commun.* **8**, 14764 (2017).
75. Ishikawa-Sasaki, K., Sasaki, J. & Taniguchi, K. A complex comprising phosphatidylinositol 4-kinase IIIβ, ACBD3, and Aichi virus proteins enhances phosphatidylinositol 4-phosphate synthesis and is critical for formation of the viral replication complex. *J. Virol.* **88**, 6586–6598 (2014).
76. Dubankova, A., Humpolickova, J., Klima, M. & Boura, E. Negative charge and membrane-tethered viral 3B cooperate to recruit viral RNA dependent RNA polymerase 3D. *Sci. Rep.* **7**, 17309 (2017).
77. Moustafa, I. M. *et al.* Structural dynamics as a contributor to error-prone replication by an RNA-dependent RNA polymerase. *J. Biol. Chem.* **289**, 36229–36248 (2014).
78. Šebera, J. *et al.* The structural model of Zika virus RNA-dependent RNA polymerase in complex with RNA for rational design of novel nucleotide inhibitors. *Sci. Rep.* **8**, 11132 (2018).
79. Butcher, S. J., Grimes, J. M., Makeyev, E. V., Bamford, D. H. & Stuart, D. I. A mechanism for initiating RNA-dependent RNA polymerization. *Nature* **410**, 235–240 (2001).

80. Dubankova, A. & Boura, E. Structure of the yellow fever NS5 protein reveals conserved drug targets shared among flaviviruses. *Antiviral Research* **169**, 104536 (2019).
81. Deng, Y.-Q. *et al.* Adenosine Analog NITD008 Is a Potent Inhibitor of Zika Virus. *Open Forum Infect Dis* **3**, ofw175 (2016).
82. Sun, D., Chen, S., Cheng, A. & Wang, M. Roles of the Picornaviral 3C Proteinase in the Viral Life Cycle and Host Cells. *Viruses* **8**, 82 (2016).

SCIENTIFIC REPORTS

OPEN

Negative charge and membrane-tethered viral 3B cooperate to recruit viral RNA dependent RNA polymerase 3D^{pol}

Received: 6 January 2017
Accepted: 29 November 2017
Published online: 11 December 2017

Anna Dubankova, Jana Humpolickova, Martin Klima & Evzen Boura

Most single stranded plus RNA viruses hijack phosphatidylinositol 4-kinases (PI4Ks) to generate membranes highly enriched in phosphatidylinositol 4-phosphate (PI4P). These membranous compartments known as webs, replication factories or replication organelles are essential for viral replication because they provide protection from the innate intracellular immune response while serving as platforms for viral replication. Using purified recombinant proteins and biomimetic model membranes we show that the nonstructural viral 3A protein is sufficient to promote membrane hyperphosphorylation given the proper intracellular cofactors (PI4KB and ACBD3). However, our bio-mimetic *in vitro* reconstitution assay revealed that rather than the presence of PI4P specifically, negative charge alone is sufficient for the recruitment of 3D^{pol} enzymes to the surface of the lipid bilayer. Additionally, we show that membrane tethered viral 3B protein (also known as Vpg) works in combination with the negative charge to increase the efficiency of membrane recruitment of 3D^{pol}.

Space in the capsids of small viruses is limited and small viruses do not encode every enzymatic activity required for their replication. Single stranded plus RNA (+RNA) viruses replicate at replication organelles (also known as replication factories or membranous webs) which provide an optimal replication environment and also protection from innate immunity[1]. Membranous webs are highly enriched in the signaling lipid PI4P (phosphatidylinositol 4-phosphate), yet +RNA viruses do not encode phosphatidylinositol 4-kinases (PI4Ks). Instead, they hijack a human enzyme, either PI4KA or PI4KB (also called PI4K IIIα or PI4K IIIβ)[2–7]. The other two human PI4K isoforms, PI4K2A and PI4K2B (also known as PI4K IIα or PI4K IIβ), are palmitoylated proteins[8], and this posttranslational modification possibly renders them a more difficult target for viruses to recruit. It is also possible that the different subcellular localization of different PI4K enzymes is the reason why PI4KB and PI4KA are hijacked by viruses and PI4K2A and PI4K2B are not[9]. However, although the subcellular localization of PI4K2A and PI4KB is similar (both are mainly Golgi localized), only PI4KB has been reported to be hijacked by viruses. In addition, although the subcellular localization of PI4KA and PI4KB is different (plasma membrane vs Golgi) both enzymes are used by HCV (the preference depends on the HCV genotype).

PI4Ks have been characterized extensively because they are essential host factors for many + RNA viruses[9]. Crystal structures of all PI4K isoforms except PI4KA are available[10–12], their complexes with binding partners Rab11, ACBD3 and 14-3-3 were structurally characterized[10,13–15] and potent and extremely selective inhibitors that exert antiviral activity have been developed[16–19]. PI4Ks are hijacked by viruses either directly or indirectly. Hepatitis C virus (HCV) uses a direct mechanism: its nonstructural NS5A protein directly binds and recruits PI4KA[3]. Similarly, the nonstructural 3A protein from the encephalomyocarditis virus (EMCV, genus Cardiovirus) interacts directly with PI4KA[20]. In contrast, most picornaviruses use an indirect mechanism with several variations among different members of the *Picornaviridae* family. ACBD3 (acyl-CoA-binding domain-containing protein-3) was recently shown to form a strong complex with PI4KB, activating its lipid kinase activity[13,21]. Nonstructural 3A proteins from several enteroviruses including poliovirus (PV) and coxsackievirus B3 (CVB3) can interact with both guanine nucleotide exchange factor-1 (GBF1)[22,23] and ACBD3 simultaneously[5,24]. Distinct, nonstructural 3A proteins from kobuviruses such as Aichi virus use nearly all their residues to interact with

Institute of Organic Chemistry and Biochemistry of the Czech Academy of Sciences, Prague, Czech Republic. Anna Dubankova and Jana Humpolickova contributed equally to this work. Correspondence and requests for materials should be addressed to E.B. (email: boura@uochb.cas.cz)

ACBD3[5,24,25]. In addition, genetic ablation of ACBD3 prevents the recruitment of PI4KB to Aichi virus replication sites[21]. It is important to mention that the picornavirus proteins arise from proteolytical processing of the viral polyprotein in such a way that various stable intermediates such as 3AB or 3CD are generated and that only membrane anchored 3AB is a substrate for the viral 3CD and 3 C protease[26] and also that the membrane composition plays a regulatory role[27]. The poliovirus 3AB protein was previously suggested to anchor the RNA replication complex to the membrane[28] through the C-terminal hydrophobic part of the 3A protein[29] and the 3AB fusion protein was also previously shown to activate the poliovirus $3D^{pol}$ enzyme[30]. The 3B protein (also known as VPg from viral protein genome linked) arises from the 3AB precursor by cleavage by the proteinase $3CD^{pro}$ and serves as a primer for the $3D^{pol}$ enzyme[31].

The reason membranous webs are highly enriched in PI4P is still poorly understood. Picornaviral polymerases are active on their own (without any protein co-factor although a primer, which is the 3B protein *in vivo* but can be both 3B or nucleic acid *in vitro* must be present). However, upon assembly of the replication complex their processivity is believed to significantly increase, more details can be found in recent review by OB Peersen[32]. Another intriguing feature of PI4P is that it can be exchanged for other lipids such as cholesterol[33–35] or phosphatidylserine (PS)[36] against concentration gradients because PI4P hydrolysis at the target membrane generates energy[37]. Indeed, production of PI4P to modify the cholesterol content of membranous webs was demonstrated for the encephalomyocarditis virus (EMCV)[20]. Another possible function of PI4P could be direct recruitment of viral effector proteins to the replication sites. PI4P binding proteins are well described in several pathogenic bacteria such as the SidC and SidM proteins from Legionella[38]. $3D^{pol}$ could be such a viral factor. If $3D^{pol}$ could bind PI4P it would be recruited to replication sites in a PI4P-dependent manner. This mechanism was reported for poliovirus (PV), and PI4P-mediated $3D^{pol}$ recruitment was proposed as a mechanism for picornaviral and flaviviral replication[39]. However, whether PI4P hyper-production is sufficient to recruit the polymerase to the surface of the lipid bilayer is not clear.

Here we sought to directly test the hypothesis that $3D^{pol}$ RNA polymerase is recruited to hyper-phosphorylated membranes using a clearly defined *in vitro* system. We reconstituted the initial generation of membranous web using purified recombinant proteins of the human Aichi virus and the biomimetic giant unilamellar vesicle (GUV) system. GUVs are very large vesicles comparable in size to human cells making them ideal for confocal microscopy. GUVs can also be filled with a sucrose solution which makes them heavier than the surrounding buffer, and as a result they do not move but instead sit at the bottom of the chamber where they are imaged. Additionally, GUVs can be prepared from almost any lipid mixture such that their lipid composition resembles that of the specific organelle they are mimicking (here, we use a mixture resembling the Golgi and viral replication organelles). For these reasons GUVs are used to reconstitute and thereby, gain molecular insight into biologically important processes that involve membranes. For instance, the ESCRT (endosomal sorting complex required for transport) complex catalyzes membrane scission[40–42] and GUVs were used to understand this reaction *in vitro*[43]. GUVs were also used to elucidate the mechanism of assembly and ESCRT recruitment by HIV Gag[44,45], the clathrin assembly on the surface of GUVs[46], and the function of Endophilin-A2 in endocytosis[47]. Additional examples and practical applications can be found in a recent book on model membranes[48].

We chose the human Aichi virus as a model organism because the interaction of its nonstructural 3A protein with host ACBD3 is well described[49,50] and reported to be required for the recruitment of PI4KB to the viral replication sites *in vivo*[21]. We demonstrate that the 3A protein is sufficient to facilitate membrane PI-phosphorylation when the appropriate cellular cofactors (ACBD3 and PI4KB) are available. However, although PI4P production alone did not lead to efficient membrane recruitment of $3D^{pol}$, the situation changed when membrane-tethered 3B protein was present. We demonstrate that not only the negatively charged PI4P but another negatively charged lipid, such as PS, can cooperate with membrane-anchored 3B to recruit the $3D^{pol}$ enzyme.

Results

3A efficiently recruits ACBD3 to model membranes. In order to faciliate the recruitment of ACBD3 to model membranes, we engineered a biomimetic recombinant 3A viral protein. The 3A protein is myristoylated at the N-terminal glycine residue[24] and also contains a hydrophobic region that every algorithm tested (CCTOP, HMMTOP, MemBrain, Memsat, Octopus, Philius, Phobius, Pro, Prodiv, Scampi, ScampiMsa, and TMHMM) predicts as a transmembrane helix that could anchor 3A protein in the lipid bilayer (Fig. 1A upper panel). This is also supported by our all-atom molecular dynamics simulations of the 3A:GOLD domain complex on the surface of the lipid bilayer[49]. Alternatively, it was suggested for the poliovirus 3AB protein that this hydrophobic region is semi-buried in the lipid bilayer[51] as depicted in the lower panel of Fig. 1A. We fused CFP and a 6xHis tag to the N-terminus of the 3A protein and we replaced the transmembrane helix with another 6xHis tag. The resulting CFP-His_6-3A-His_6, is tethered by both its N-terminus and C-terminus to a Ni^{2+} containing membrane in a similar fashion to that of the wild type protein (Fig. 1B). A pilot experiment revealed that the CFP-His_6-3A-His_6 could be tethered to artificial membranes that contain a nickel cation-bound lipid such as the DGS-NTA(Ni) (Fig. 1C, upper panel) but not to control membranes (Fig. 1C, lower panel).

ACBD3 is a Golgi resident scaffolding protein[52] that has been reported to interact with 3A proteins from several +RNA viruses including the Aichi virus[5,21,24,25,53]. We sought to directly test this interaction *in vitro* and in the context of a membrane. Recombinant ACBD3 labeled with rhodamine was only weakly localized to the membrane at 250 nM concentration (Fig. 1D, upper panel). However, when membranes harbored viral 3A protein, ACBD3 was efficiently recruited to these membranes (Fig. 1D, lower panel).

3A, ACBD3, and PI4KB form a protein complex on model membranes. PI4KB is a lipid kinase that must be recruited to membranes to function properly. The interaction of ACBD3 with PI4KB was previously found to have a dissociation constant of 320 nM, and ACBD3 can recruit PI4KB to membranes when ACBD3 is artificially tethered to the membrane surface[13]. We therefore sought to test if 3A recruited ACBD3

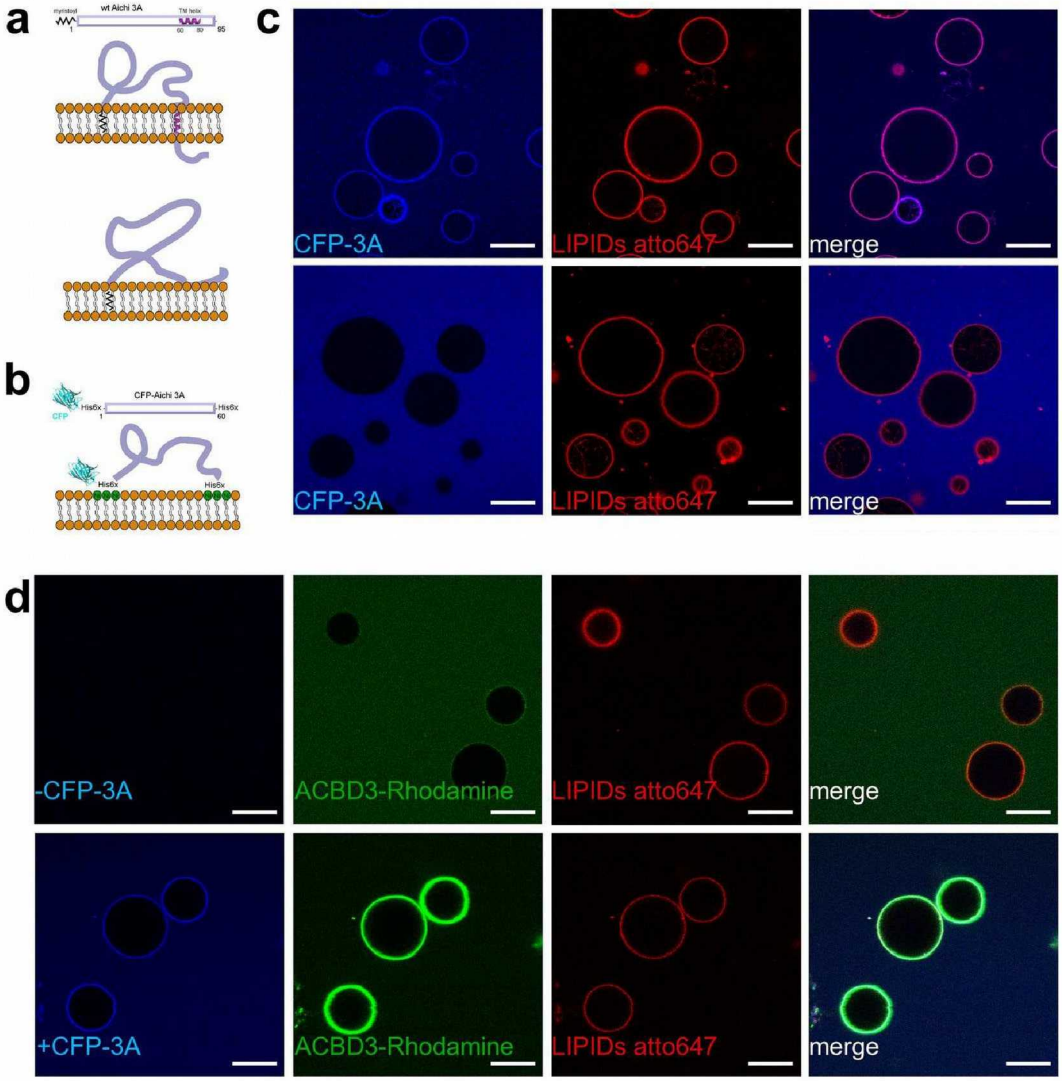


Figure 1. Aichi virus 3A protein on the membrane. (**A**) Schematic representation of the wild type 3A protein. The two possible topologies of the wild type 3A protein are shown. Upper panel – the C-terminal hydrophobic stretch is depicted as a transmembrane helix. Lower panel – the C-terminal hydrophobic stretch is depicted as semi-buried in the lipid bilayer. (**B**) Schematic representation of the mCerulean - 3A fusion protein (named CFP-3A). CFP-3A contains two 6xHis tags; one between the CFP and its N-terminus and one at the C-terminus. These two His tags can be used to attach the CFP-3A protein to a membrane containing DGS-NTA(Ni) (a lipid that has Ni^{2+} bound to its headgroup). (**C**) CFP-3A bound to GUVs. Upper panel: 250 nM CFP-3A was added to GUVs containing 5% of DGS-NTA(Ni) and ATTO647N-DOPE (0.1 mol %). Lower panel: as above but DGS-NTA(Ni) was replaced by POPC. The CFP-3A signal is in blue and the ATTO647 signal in red. Representative image of three independent experiments. Scale bar = 20 µm. (**D**) Aichi virus 3A protein efficiently recruits ACBD3 to the membrane. Rhodamine labeled ACBD3 (250 nM) was incubated with GUVs without 3A protein (upper panel) or with 3A protein (250 nM, lower panel). The CFP-3A signal is in blue, the ACBD3-rhodamine signal is in green, and ATTO647 labeled lipids in in red. Representative image of three independent experiments. Scale bar = 20 µm.

was capable of further recruiting PI4KB to form a larger 3A:ACBD3:PI4KB complex on the membrane. PI4KB is efficiently recruited to model membranes decorated with CFP-3A and ACBD3 (Fig. 2A,B). Next we aimed to elucidate if 3A:ACBD3:PI4KB forms oligomers on the surface of the lipid bilayer. We used fluorescent recovery after photobleaching (FRAP) in which large protein clusters would show significantly slower fluorescence recovery than individual protein complexes[54]. We compared the FRAP of the CFP-3A, CFP-3A:ACBD3, and CFP-3A:ACBD3:PI4KB (Fig. 2C) and found no statistically significant differences. Based on the fast FRAP we assume that the 3A:ACBD3:PI4KB protein complex forms a heterotrimer or smaller oligomers on the surface of the membrane but not large assemblies such as lattices that cannot exhibit fast FRAP (Fig. 2D).

PI4KB is highly active as part of the 3A:ACBD3:PI4KB protein complex. Only membrane-tethered ACBD3 activates the PI4KB enzyme[13]. It is likely that viruses use their 3A protein to tether the ACBD3 protein to

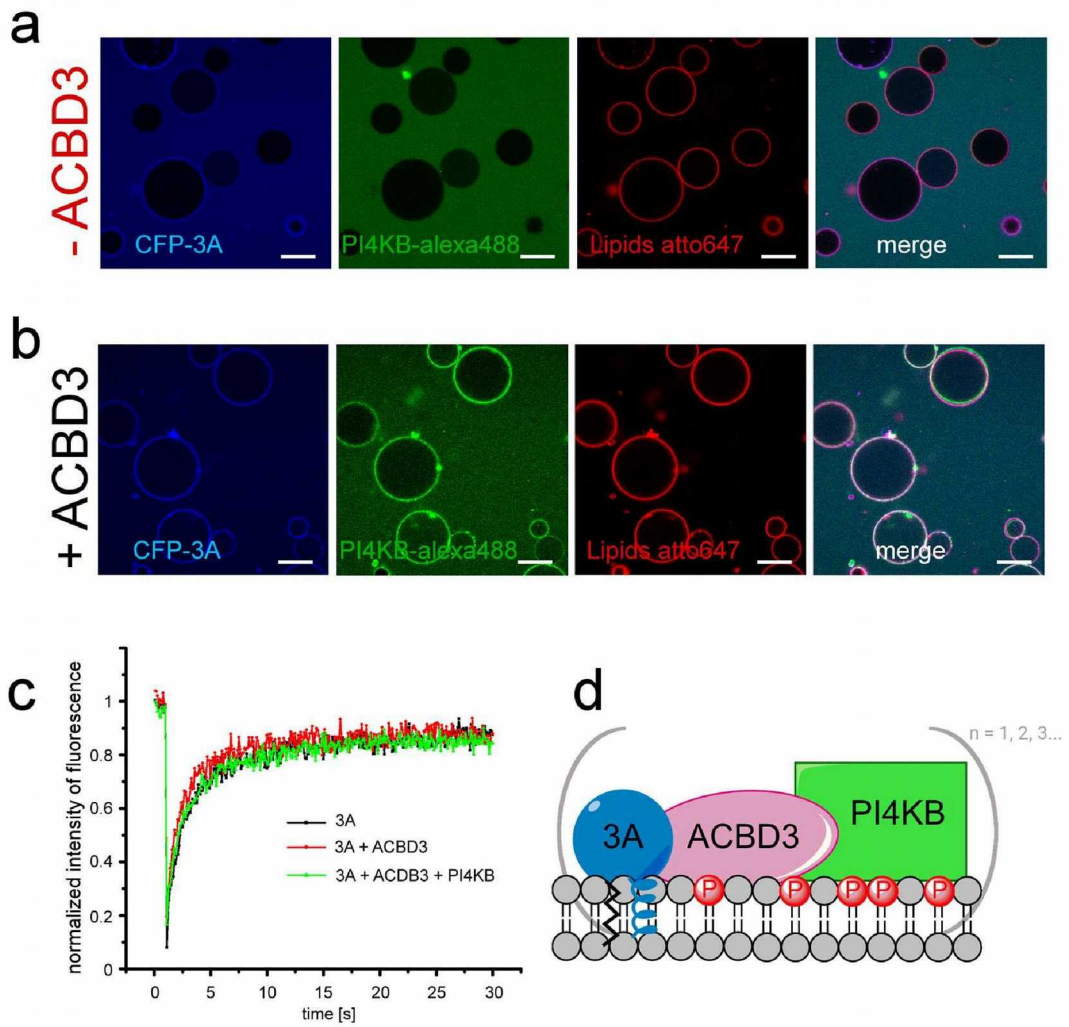


Figure 2. Aichi virus 3A protein recruits the PI4KB kinase via ACBD3. (**A**) 3A protein does not recruit PI4KB directly to GUVs. 250 nM CFP-3A and PI4KB labeled by Alexa488 were added to GUVs containing 5% of DGS-NTA(Ni) and ATTO647N-DOPE (0.1 mol %). The CFP-3A signal is in blue, the PI4KB-Alexa488 signal is in green and the ATTO647 signal is in red. Representative image of three independent experiments. Scale bar = 20 µm. (**B**) PI4KB is recruited to membranes when ACBD3 is present. CFP-3A, Alexa488 labeled PI4KB and unlabeled ACBD3 (250 nM each) were added to GUVs containing 5% DGS-NTA(Ni) and ATTO647N-DOPE (0.1 mol %). The CFP-3A signal is in blue, the PI4KB-Alexa488 signal in green and the ATTO647 signal is in red. Representative image of three independent experiments. Scale bar = 20 µm. (**C**) FRAP analysis of the 3A:ACBD3:PI4KB protein complex. A small cross-section of a GUV membrane was intensively bleached by a 405 nm laser and fluorescence recovery after photobleaching (FRAP) of CFP-3A was measured. (**D**) Schematic representation of the 3A:ACBD3:PI4KB protein complex.

the membrane in order to activate PI4KB[21]. The 3A protein is perfectly suited for this job because it is membrane tethered and tightly binds ACBD3. To test this hypothesis *in vitro* we used a SidC fluorescent PI4P biosensor (PI4P binding domain of SidC from Legionella pneumonia fused to mCherry)[38] to detect the relative amount of PI4P synthesized on the surface of GUVs (Fig. 3A). GUVs were first decorated with the viral CFP-3A protein and then incubated with PI4KB alone or with ACBD3 and PI4KB (Fig. 3B upper and middle panels). The control consisted of a mixture of all proteins but was devoid of ATP (Fig. 3B lower panel). The efficiency of membrane phosphorylation (the conversion of PI to PI4P) was determined from the biosensor's fluorescent signal on the surface of the membrane. Background levels of PI4KB lipid phosphorylation in the presence of only the 3A protein decorated GUVs were measured (Fig. 3B upper panel, Fig. 3C). However, the observed efficiency of the reaction in the presence of ACBD3 was far higher as judged by the bio-sensor recruitment (Fig. 3B middle panel). In the absence of ATP, no membrane phosphorylation was observed even when all the proteins were present (Fig. 3B lower panel), documenting the specificity of the system. As an additional control we performed the phosphorylation reaction in the presence of the specific PI4KB inhibitor MI364 (compound **23** in our recent publication[55]) and, as expected, we observed only limited recruitment of the mCherry-SidC biosensor to the GUVs (Fig. 3C). Furthermore, we performed an analogous experiment using unlabeled PI4KB and a fluorescent PI4KB inhibitor (compound **3** in our recent publication[56]). Again, we observed only limited recruitment of the mCherry-SidC biosensor. However,

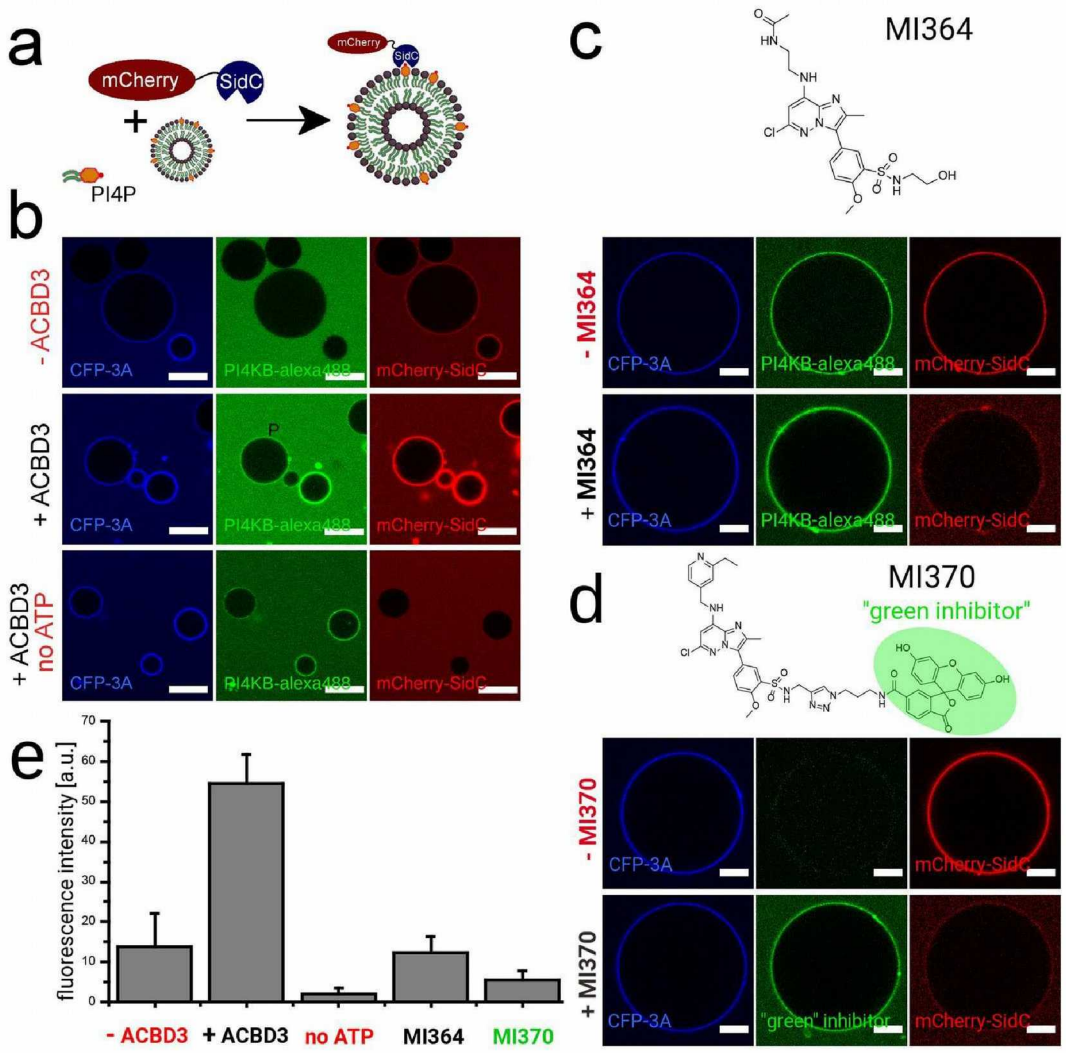


Figure 3. PI4KB is activated in the 3A:ACBD3:PI4KB protein complex. (**A**) Scheme of the experiment. A PI4P binding domain from the Legionella pneumonia SidC protein that binds to PI4P with nanomolar affinity was fused to mCherry and used as a fluorescent PI4P biosensor. (**B**) Production of PI4P on the surface of GUVs. Upper panel: GUVs decorated with the viral 3A-CFP protein (250 nM) were incubated with Alexa488 labeled PI4KB (250 nM) and SidC-mCherry (100 nM) and imaged using a confocal microscope. Middle panel: As above but also ACBD3 (250 nM) was also added. Lower panel: As the middle panel but no ATP was added. The CFP-3A signal is in blue, the PI4KB-Alexa488 signal in green and the mCherry-SidC signal is in red. Representative image of three independent experiments. Scale bar = 20 µm. (**C**) Inhibition by the compound MI364. On the top is the chemical structure of MI364. On the bottom is a representative image of GUVs phosphorylated by Alexa488 labeled PI4KB (upper panel: without MI364, lower panel: with 5 µM MI364). The CFP-3A signal is in blue, the PI4KB-Alexa488 signal in green and the mCherry-SidC signal is in red. Representative image of three independent experiments. Scale bar = 10 µm. (**D**) Inhibition by the fluorescent compound MI370. On the top is the chemical structure of MI370 (the fluorescent part in highlighted in green). On the bottom is a representative image of GUVs phosphorylated by unlabeled PI4KB (upper panel: without MI370, lower panel: with 5 µM MI370). The CFP-3A signal is in blue, the MI370 signal in green and the mCherry-SidC signal is in red. Representative image of three independent experiments. Scale bar = 10 µm. (**E**) Quantification of PI4P production. Quantification of the intensity of fluorescence of the SidC-mCherry PI4P biosensor on the surface of GUVs was. Standard deviations are based on three independent experiments.

in this setting the fluorescent inhibitor could be used to visualize unlabeled PI4KB on the surface of the GUVs (Fig. 3D). Quantification of the fluorescence intensity revealed that the apparent activity of the 3A:ACBD3:PI4KB protein complex was roughly 4-fold higher than that of the PI4KB kinase alone in this system (Fig. 3E).

PI4P and Aichi 3D^{pol} membrane recruitment. The reconstitution of ACBD3 recruitment to model membranes by viral 3A, and the subsequent recruitment of the lipid kinase PI4KB made possible the analysis of the membranes binding properties of 3D^{pol}. We used our *in vitro* GUV system to test PI4P mediated 3D^{pol} recruitment because this mechanism was suggested previously[39]. As shown above, the 3A:ACBD3:PI4KB complex

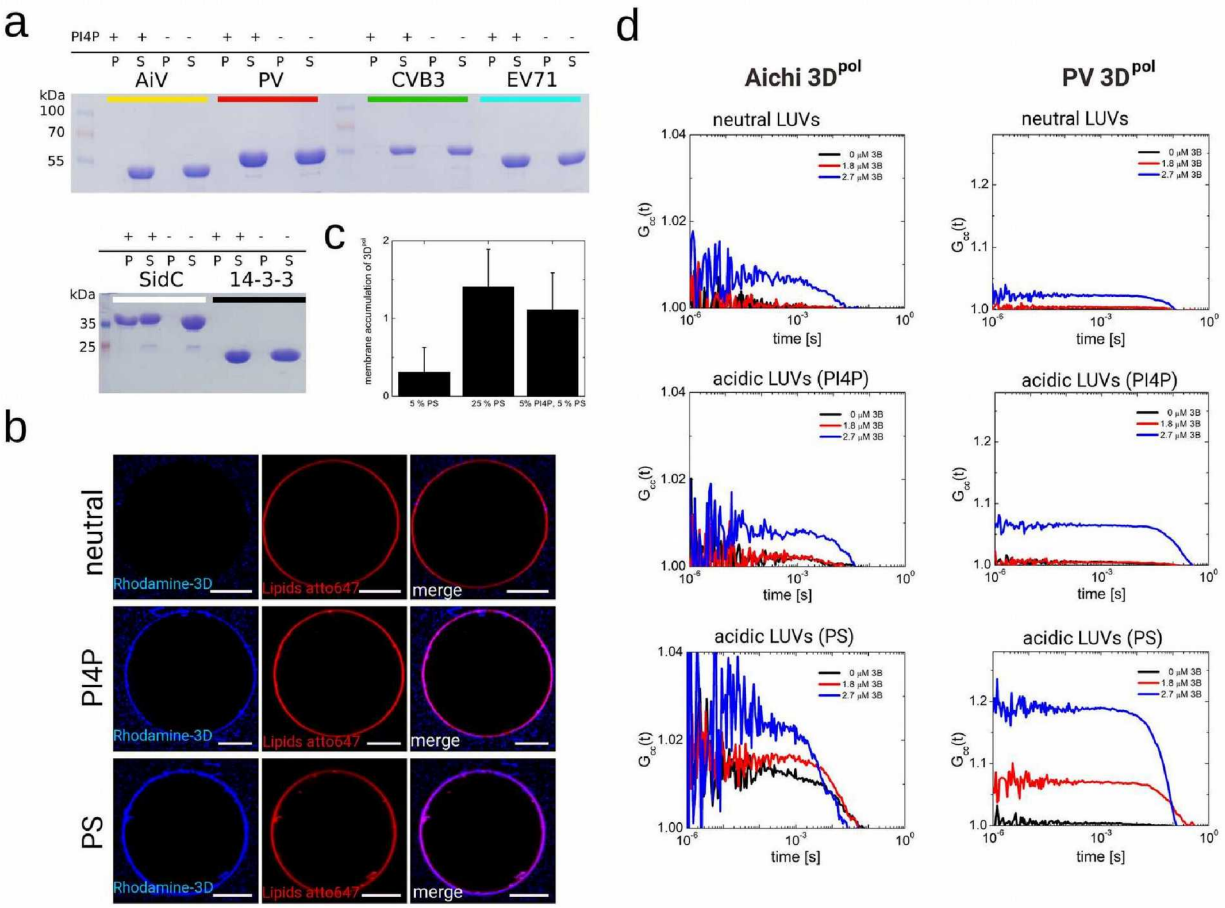


Figure 4. 3D^{pol} enzymes efficiently bind acidic membranes only in the presence of membrane-tethered 3B. (**A**) Recombinant 3D^{pol} enzymes from several +RNA viruses bind PI4P poorly in a liposome pulldown assay. 3D^{pol} from the Aichi virus (AiV), Polio virus (PV), Coxsackie virus B3 (CVB3) and Enterovirus 71 (EV71) show no binding to PI4P liposomes at 60 μM concentration. SidC was used as a positive control and 14-3-3ζ protein as a negative control. Liposomes were centrifuged and the pellets (denoted P) and supernatants (denoted S) were analyzed using SDS PAGE. For the full length gels see the SI Fig. 2. (**B**) GUV recruitment assay. GUVs of different composition were used. Upper panel – neutral membrane (5% PS), middle panel – PI4P acidic membrane (5% PS + 5% PI4P), lower panel – PS acidic membrane. Aichi 3D^{pol} fluorescence signal in blue (left), membrane in red (middle) and merged image (right). A typical result of three independent experiments. Scale bar = 20 μm. (**C**) Quantification of Aichi 3D^{pol} binding to membranes of different composition. Composition dependent membrane accumulation of CFP-3D^{pol}: excess of the fluorescence signal of CFP in the GUV pixels relative to the signal of unbound CFP-3D^{pol}. The error bars indicate a 95% confidence interval. (**D**) Cross-scorrelation curves of CFP – 3D^{pol} and LUVs. FCS curves representing temporal cross-correlation of the CFP-3D^{pol} (Aichi 3D^{pol} on the left panel and PV 3D^{pol} on the right panel) and ATTO647N-DOPE labelled LUVs at various membrane lipid compositions (neutral, enriched by PI4P, and enriched by PS; see SI Table 1). The curves show the cooperative effect of the 3B peptide on 3D^{pol} membrane recruitment. The concentration of the Aichi or PV 3B peptide (attached to the membrane surface by His-tag – 18:1 DGS-NTA(Ni) interaction) was 0 μM (black), 1.8 μM (red), and 2.7 μM (blue).

efficiently phosphorylates GUVs to make PI4P (Fig. 3). However, we did not observe 3D^{pol} membranes recruitment even though the mCherry-SidC nanomolar PI4P binder biosensor was recruited (SI Fig. 1A). Similarly, 3D^{pol} failed to bind GUVs prepared with synthetic PI4P (SI Fig. 1B). We reasoned that the 3D^{pol}:PI4P interaction might be weak, or that the Aichi 3D^{pol} is unique among 3D^{pol} enzymes in that it does not bind PI4P. Therefore, we tested several different 3D^{pol} enzymes (Aichi virus, PV, CVB3, EV71) in a liposome pulldown assay. However, no PI4P-mediated liposome binding of any of the polymerases was detected using this method (Fig. 4A).

Pulldown assays are not sensitive and a weak but biologically relevant interaction could be overlooked. Therefore, we performed a more sensitive experiment. GUVs were incubated with fluorescently (rhodamine) labeled Aichi 3D^{pol} at nanomolar concentration, which facilitates efficient binding due to the vast excess of lipids. In addition, the fluorescence image was taken using a low scanner speed (50 μs per pixel) to ensure a higher signal-to-noise ratio (Fig. 4B). We used GUVs with PI4P and two sets of controls: neutral GUVs in which PI4P was replaced with the neutral lipid PC and an acidic membrane where PI4P was replaced with the negatively charged PS lipid (we used four molecules of PS for every molecule of PI4P to create a highly acidic membrane

without the PI4P lipid). We observed that $3D^{pol}$ does not bind to neutral GUVs (Fig. 4B upper panel). However, $3D^{pol}$ did bind the PI4P-containing GUVs (Fig. 4B middle panel) suggesting that the $3D^{pol}$ enzyme is, indeed, capable of binding PI4P containing membranes albeit with low affinity. Surprisingly, the $3D^{pol}$ enzyme could also bind PS-containing acidic GUVs (Fig. 4B lower panel). Quantification of membrane binding revealed that $3D^{pol}$ enzyme binds acidic GUVs without a preference for specific acidic lipids (Fig. 4C).

However, the extent of $3D^{pol}$ recruitment by model acidic membranes seems too low to support efficient viral replication within the host cell. The only other reported binding partner of any $3D^{pol}$ enzyme (besides the RNA) is the small 3B protein that serves as a primer. A portion of 3B in the infected cell exists as a membrane tethered 3AB fusion protein. Thus, we used our *in vitro* system to test whether membrane-tethered 3B and the acidic membrane can cooperate to recruit $3D^{pol}$ enzymes. We could not use GUVs for this experiment because the reported dissociation constants between 3B and 3D are in the tens of micromolar range and it is not possible to use micromolar lipid concentrations in the case of GUVs. We instead used large unilamellar vesicles (LUVs). Aichi and poliovirus (PV) $3D^{pol}$ enzymes were incubated with LUVs in the presence of increasing concentrations of membrane-tethered 3B proteins, and the temporal fluorescence cross-correlation function between the CFP labeled $3D^{pol}$ and ATTO647N labeled LUVs was measured. In this analysis, the higher the cross-correlation the higher the binding (Fig. 4D, SI Fig. 4). We observed negligible binding of Aichi $3D^{pol}$ to neutral membranes at all 3B concentrations tested (0–2.7 μM) and small PV $3D^{pol}$ binding to neutral membranes at high membrane-tethered 3B concentrations. However, $3D^{pol}$ binding to acidic membranes (PI4P or PS) increased with increasing concentrations of membrane tethered 3B.

Discussion

PI4P production is essential for many + RNA viruses, however, + RNA viruses do not encode PI4K enzymes and must hijack human enzymes. Many viruses including the Aichi virus converged on hijacking the PI4KB enzyme through the interaction with the ACBD3 protein. We recently showed that ACBD3 is a sub-micromolar binder of PI4KB and that ACBD3 can recruit the PI4KB enzyme to any membrane both *in vitro* and *in vivo*[13]. Because the viral 3A protein is a transmembrane protein that tightly binds the GOLD domain of ACBD3[49], it can recruit ACBD3 and thus the PI4KB to any target membrane. Soon after formation of the 3A:ACBD3:PI4KB protein complex, the membranes become hyper-phosphorylated, which is a key step in the biogenesis of replication organelles - a complex process that is not fully understood. It is initiated after polyprotein synthesis followed by stepwise proteolytic cleavage by viral proteases. These cleavage events produce intermediate as well as final cleavage products that together with host factors initiate replication complex formation. One of them is the 3A:ACBD3:PI4KB protein complex.

Here we reconstituted *in vitro* the formation of the 3A:ACBD3:PI4KB protein complex and membrane phosphorylation. We first prepared fluorescently labeled Aichi 3A protein and tethered it to the surface of GUVs to mimic infected cells (Fig. 1). As expected, the membrane-tethered 3A protein was able to recruit ACBD3 and subsequently the PI4KB kinase to the model membranes. Thus, our *in vitro* reconstitution experiments show that the small 3A protein is sufficient to recruit and activate PI4KB through interaction with ACBD3. This is an elegant mechanism – a protein less than 100 amino acids is all that is needed for the membrane hyper-phosphorylation – a key step of replication organelles formation. Congruently, when our manuscript was under review another study was published that also shows that viral 3A protein can activate PI4KB through the ACBD3 protein[50].

The next key step in replication factory biogenesis is the recruitment of the viral $3D^{pol}$ enzyme. It was suggested that the $3D^{pol}$ enzyme binds directly to the PI4P lipid[39]. Therefore, we decided to reconstitute $3D^{pol}$ recruitment using our *in vitro* system. We did not observe $3D^{pol}$ recruitment to membranes under conditions in which the nanomolar PI4P reported was clearly recruited (SI Fig. 1). However, we did observe $3D^{pol}$ recruitment to PI4P rich membranes using a more sensitive technique (Fig. 4). Surprisingly, $3D^{pol}$ was also recruited to negative charged PS-containing membranes but not to neutral membranes (Fig. 4B). We repeated all experiments using the poliovirus $3D^{pol}$ enzyme to rule out the possibility that this mechanism is specific for the Aichi $3D^{pol}$ enzyme. Similar results were obtained, indicating that the 3B protein is crucial in recruitment of the $3D^{pol}$ enzymes. The PV 3B protein could somewhat recruit to neutral membranes while the Aichi 3B protein could not, but recruitment was more efficient in the context of negatively charged membranes. It is also worth mentioning, that we observed a somewhat lower cooperativity effect of the poliovirus 3B and $3D^{pol}$ proteins at the negatively charged membranes as documented by the lower amplitude of the cross-correlation function (SI Fig. 3). However, poliovirus $3D^{pol}$ is known to be able to form multimers[57,58] and the avidity effect probably more than compensates. Interestingly, the PS lipid was significantly better in recruitment of the Aichi $3D^{pol}$ enzyme while in the case of the PV $3D^{pol}$ PS performed only slightly better than PI4P (Fig. 4D) which further supports our hypothesis that PI4P is not specifically recognized by picornaviral $3D^{pol}$ enzymes.

Many human and animal +RNA viruses including poliovirus (PV), Coxsackievirus B3 (CVB3), rhinovirus, norovirus, and HCV target the ER, Golgi or TGN membranes highly enriched in PI4P and cholesterol[59]. However, the observed interaction of $3D^{pol}$ enzymes with a properly (negatively) charged cholesterol rich model membrane is rather weak. Importantly, $3D^{pol}$ is a part of a relatively stable precursor protein that could affect its membrane recruitment in infected cells. The 3 CD^{pro} precursor protein of poliovirus is stable enough to be analyzed by protein crystallography[60]. The putative function of $3D^{pol}$ in the 3 CD^{pro} protein is to modulate the proteolytic activity of the 3C protease domain and RNA binding of the $3CD^{pro}$ protein[32]. It is important to emphasize that the picornaviral $3CD^{pro}$ has no polymerase activity. It only becomes active upon proteolytic cleavage of the link between the 3C and 3D proteins[32].

As with most polymerases, $3D^{pol}$ requires a primer. While *in vitro* it can initiate RNA synthesis using an RNA or DNA primer *in vivo* the viral 3B protein is always used as a primer and the first step is uridylylation of its tyrosine side chain[61]. The viral 3B protein is partially present as a membrane-tethered 3AB fusion protein and it was previously suggested to serve as a tether for $3D^{pol}$ *in vivo*[28,62]. In our *in vitro* system, the membrane tethered 3B protein increased the efficiency of $3D^{pol}$ recruitment (Fig. 4D, SI Fig. 4). We conclude that negative charge and the

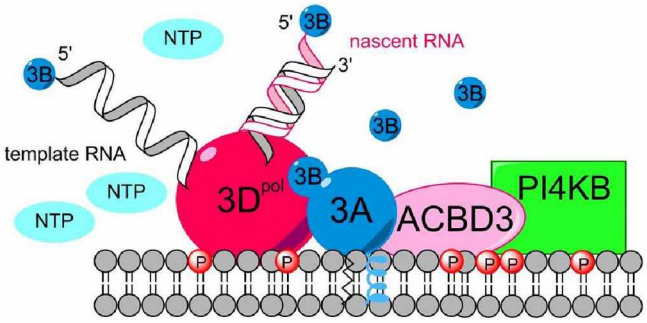


Figure 5. Schematic model of the 3D^{pol}:3AB:ACBD3:PI4KB protein complex.

membrane-tethered 3B protein work in combination to recruit 3D^{pol} enzymes. We thus propose that two protein complexes-3A:ACBD3:PI4KB (Fig. 2D) and 3D^{pol}:3AB:ACBD3:PI4KB (Fig. 5) - are in a dynamic equilibrium in infected cells. While both are PI-phosphorylation competent only the latter also participates in RNA synthesis.

Concluding Remarks. *In vitro* reconstitution is a powerful tool for testing biological hypotheses. The hypothesis that the non-structural 3A protein is the only viral protein needed to hyper-phosphorylate membranes when the cellular co-factors are present was proven in our *in vitro* system. However, our bio-mimetic GUV system reconstitution revealed that negative charged lipids and not necessarily PI4P are responsible for recruitment of the 3D^{pol} enzyme to the surface of the lipid bilayer. Moreover, our studies also revealed that the 3B protein works in combination with the negative charge to efficiently recruit the 3D^{pol} enzyme.

Materials and Methods

Cloning, Protein Expression and Purification. ACBD3, PI4KB (with deleted disordered loop 424–521 to increase stability and to facilitate bacterial expression), and mCherry-SidC (PI4P binding domains only, amino acid residues 613–742) were expressed as described previously[15]. AiV, EV 71, CVB3 3D^{pol} were cloned into the pRSFD vector with a His_6-GB1 purification/solubility tag at the N-terminus followed by a TEV protease cleavage site. PV 3D^{pol} was expressed from the plasmid pCG1. This plasmid contains a ubiquitin tag at the N-terminus and His_6 purification tag at the C-terminus of the PV 3D^{pol}. To remove the ubiquitin tag and generate PV 3D^{pol} with a native N-terminus, the PV 3D^{pol} was co-expressed with ubiquitin protease in *E. coli* and further purified as previously described[63]. CFP-3A protein was cloned into a pRSFD vector with a His_6 purification tag at the N-terminus followed by a TEV cleavage site and a mCerulean fluorescent protein (CFP) to produce final protein His_6-TEV-CFP-His-Aichi 3A (residues 1–58)-His_6. CFP-3D^{pol} proteins were cloned into pRSFD vector with a His_6 purification tag at the N-terminus followed by a TEV cleavage site and mCerulean fluorescent protein to produce the final protein His_6-TEV-CFP-3D^{pol}. These proteins were purified using standard protocols in our laboratory[64,65]. Briefly, the proteins were expressed in BL21 Stars cells grown in autoinduction media at 18 °C overnight. Cells were lysed in lysis buffer (50 mM Tris pH 8, 300 mM NaCl, 3 mM βME, 10% glycerol) and nickel affinity purification was carried out according to the manufacturer's instructions (Machery-Nagel). The His_6 and His_6-GB1 tags were removed by the TEV protease. The CFP-3A protein was prone to degradation, however, degradation products were removed by anion exchange chromatography. Finally, proteins were purified by size exclusion chromatography (SEC) in SEC buffer (20 mM Tris pH 8, 300 mM NaCl, 3 mM βME or 0.2 mM TCEP if the protein was meant to be labeled), concentrated to 3–10 mg/ml, aliquoted and stored at −80 °C until needed.

Protein labeling with fluorescent dyes. ACBD3, 3D^{pol} and PI4KB were labeled on native cysteine residues as described previously[13]. Briefly, a 3-fold molar excess of rhodamine or Alexa488 maleimide derivatives (Molecular Probes) was incubated with pure recombinant proteins at 4 °C overnight. The reaction was quenched by adding a large molar excess of βME and the proteins were purified again on SEC. Labeling efficiency was estimated spectroscopically to be 116% for PI4KB labeled with Alexa488, 250% for 3D^{pol} labeled with rhodamine and 181% for ACBD3 labeled with rhodamine.

Giant Unilamellar Vesicle Preparation and Imaging. Giant Unilamellar Vesicles (GUVs) were composed of POPC (54.9 mol %), POPS (5 or 10 mol %), PI4P (0 or 5 mol %) cholesterol (20 mol %), PI (10 mol %), DGS-NTA(Ni) (0 or 5 mol %) (Avanti Polar lipids), and ATTO647N-DOPE (0.1 mol %). Membrane composition for each experiment is also summarized in SI Table 1. When changing the concentration of the lipid mixture (e.g. to contain or not to contain PI4P or DGS-NTA(Ni)), charged lipids were always replaced with charged lipids (POPS for PI4P) and non-polar lipids with non-polar lipids (DGS-NTA(Ni) for POPC) to keep the net charge the same throughout all the experiments. GUVs were prepared by electroformation in a homemade teflon chamber as described previously[66]. Briefly, 50 μg of the lipid mixture in 10 μl volume was spread onto each ITO coated glass (5 × 5 cm) and dried under vacuum overnight. Later the glasses were placed in the teflon chamber separated by 1 mm thick spacers and 5 ml of 600 mM sucrose preheated to 60° was added and AC current (10 Hz, maximum amplitude 1 V) was applied for one hour in an incubator warmed to 60°. Then the chamber was allowed to cool down to room temperature and the GUVs were transferred with a glass pipette to a glass test tube. For imaging 100 μl of the GUVs were mixed with 100 μl of isotonic buffer (90 mM Tris pH 8, 20 mM $MgCl_2$, 40 mM Imidazole, 565 mM NaCl, 3 mM βME) in a BSA coated 4-chamber glass bottom dish (*In Vitro* Scientific). GUVs

were imaged using the filterless Zeiss LSM780 confocal system. To avoid bleed-through only mCerulean (CFP) tagged proteins (excitation 405 nm, emission 465–571 nm) and ATTO647-tagged lipids (excitation 633 nm, emission 645–759 nm) were excited simultaneously. Alexa488 (excitation 488, emission 508–578 nm), Rhodamine (excitation 561 nm, emission 565–585 nm) and mCherry (excitation 561 nm, emission 604–621 nm) fluorophores were excited separately. Images were processed using the ZEN 2012 software (Zeiss) and ImageJ[67].

Liposome (MLVs and LUVs) preparation. To prepare MLVs (multilamellar large vesicles) 1 mg of lipids composed of POPC (60 mol %), POPS (10 mol %), PI4P (10 mol %), and cholesterol (20 mol %) was mixed in chloroform. The chloroform was evaporated over a stream of dry nitrogen and the lipids were further dried under vacuum for at least three hours. Then the lipid film was rehydrated with 1 mL of liposome buffer (10 mM Tris pH = 7.4, 10 mM $MgCl_2$, 150 mM NaCl, 3 mM βME) and vortexed intensively. To obtain LUVs (large unilamellar vesicles) the turbid solution containing MLVs was extruded 21 times using 100 nm filters in the extruder (Avanti Polar Lipids Inc, Alabaster, AL, U.S.A.).

Liposome pulldown assay. Liposome pulldown assays were performed as described previously[68]. Briefly, MLVs with and without PI4P were prepared as described above. The liposomes were incubated with: AiV, PV, CVB3 and EV71 $3D^{pol}$ enzymes (final protein concentration 60 μM and final lipid concentration 0.5 mg/ml) in a total volume of 60 μl. SidC (KD towards PI4P = 70 nM[38]) was used as a positive control (15 μM) and 14-3-3ζ (30 μM) was used as a negative control. The reaction mixtures were incubated for 20 min on ice. The mixtures were then centrifuged (22000 g, 10 min). The supernatant was removed and the pellet re-suspended in 60 μl of liposome buffer (20 mM Tris pH 7.4, 150 mM NaCl, 3 mM βME) and analyzed by SDS-PAGE. Control liposomes with no PI4P were used to confirm that binding was due to PI4P.

Fluorescence cross-correlation spectroscopy (FCCS). The FCCS experiments were carried out at LSM 780 confocal microscope (Zeiss, Jena, Germany) equipped with an LSM upgrade kit (Picoquant, Berlin, Germany) enabling Time-Correlated Single Photon Counting (TCSPC) acquisition. CFP and ATTO647N was excited by a 405 nm laser at 20 MHz repetition frequency, and by a 633 nm continuous-wave laser, respectively. A 40x/1.2 water objective was used together with a pinhole in the detection plain. Behind the pinhole, light was guided to the detectors (tau-SPADs, Picoquant), in front of which the emission filters (482/35, and 650/50) were placed. The collected data were correlated using a home-written script in Matlab (Mathworks, Natick, MA) according to the algorithm described in[69]. To avoid detector crosstalk, the red channel fluorescence signal was split according to its TCSPC pattern (exponential for the signal generated by the pulsed laser and flat for the 633 nm continuous wave laser) into two contributions and only the signal assigned to the flat TCSPC profile was correlated. Details of data processing are given in Gregor and Enderlein[70].

50 μL of LUVs were mixed with 1 μL of CFP-$3D^{pol}$ (100 nM), 0–20 μL of 3B-His (18 μM), and the LUV buffer, so that the final volume was 100 μL, the concentration of CFP-$3D^{pol}$ was 1 nM, and the concentration of 3B-His ranged from 0 to 2.7 μM. The FCS experiment was carried out 15–20 minutes after mixing.

The lipid composition of LUVs is summarized in SI Table 1. Briefly: (POPC (64.99-x-y %), POPS (x%), PI4P (y%), PI (10%), cholesterol (20%), DGS-NTA(Ni) (5%), ATTO647N-DOPE (0.01%), where x = 0 and y = 0 for neutral membranes, x = 10 and y = 5 for acidic, PI4P-containing membranes, and x = 25% and y = 0% for acidic, highly charged membranes.

References

1. den Boon, J. A. & Ahlquist, P. Organelle-like membrane compartmentalization of positive-strand RNA virus replication factories. *Annual review of microbiology* **64**, 241–256, https://doi.org/10.1146/annurev.micro.112408.134012 (2010).
2. Berger, K. L. *et al.* Roles for endocytic trafficking and phosphatidylinositol 4-kinase III alpha in hepatitis C virus replication. *Proceedings of the National Academy of Sciences of the United States of America* **106**, 7577–7582, https://doi.org/10.1073/pnas.0902693106 (2009).
3. Berger, K. L., Kelly, S. M., Jordan, T. X., Tartell, M. A. & Randall, G. Hepatitis C virus stimulates the phosphatidylinositol 4-kinase III alpha-dependent phosphatidylinositol 4-phosphate production that is essential for its replication. *Journal of virology* **85**, 8870–8883, https://doi.org/10.1128/JVI.00059-11 (2011).
4. Arita, M. *et al.* Phosphatidylinositol 4-kinase III beta is a target of enviroxime-like compounds for antipoliovirus activity. *Journal of virology* **85**, 2364–2372, https://doi.org/10.1128/JVI.02249-10 (2011).
5. Sasaki, J., Ishikawa, K., Arita, M. & Taniguchi, K. ACBD3-mediated recruitment of PI4KB to picornavirus RNA replication sites. *The EMBO journal* **31**, 754–766, https://doi.org/10.1038/emboj.2011.429 (2012).
6. Reiss, S. *et al. Recruitment and activa*tion of a lipid kinase by hepatitis C virus NS5A is essential for integrity of the membranous replication compartment. *Cell host & microbe* **9**, 32–45, https://doi.org/10.1016/j.chom.2010.12.002 (2011).
7. Dornan, G. L., McPhail, J. A. & Burke, J. E. Type III phosphatidylinositol 4 kinases: structure, function, regulation, signalling and involvement in disease. *Biochemical Society transactions* **44**, 260–266, https://doi.org/10.1042/BST20150219 (2016).
8. Boura, E. & Nencka, R. Phosphatidylinositol 4-kinases: Function, structure, and inhibition. *Experimental cell research* **337**, 136–145, https://doi.org/10.1016/j.yexcr.2015.03.028 (2015).
9. Altan-Bonnet, N. & Balla, T. Phosphatidylinositol 4-kinases: hostages harnessed to build panviral replication platforms. *Trends in biochemical sciences* **37**, 293–302, https://doi.org/10.1016/j.tibs.2012.03.004 (2012).
10. Burke, J. E. *et al.* Structures of PI4KIIIbeta complexes show simultaneous recruitment of Rab11 and its effectors. *Science* **344**, 1035–1038, https://doi.org/10.1126/science.1253397 (2014).
11. Baumlova, A. *et al.* The crystal structure of the phosphatidylinositol 4-kinase IIalpha. *EMBO reports* **15**, 1085–1092, https://doi.org/10.15252/embr.201438841 (2014).
12. Klima, M. *et al.* The high-resolution crystal structure of phosphatidylinositol 4-kinase IIbeta and the crystal structure of phosphatidylinositol 4-kinase IIalpha containing a nucleoside analogue provide a structural basis for isoform-specific inhibitor design. *Acta crystallographica. Section D, Biological crystallography* **71**, 1555–1563, https://doi.org/10.1107/S1399004715009505 (2015).
13. Klima, M. *et al.* Structural insights and *in vitro* reconstitution of membrane targeting and activation of human PI4KB by the ACBD3 protein. *Scientific reports* **6**, 23641, https://doi.org/10.1038/srep23641 (2016).

14. Eisenreichova, A., Klima, M. & Boura, E. Crystal structures of a yeast 14-3-3 protein from Lachancea thermotolerans in the unliganded form and bound to a human lipid kinase PI4KB-derived peptide reveal high evolutionary conservation. *Acta Crystallogr F Struct Biol Commun* **72**, 799–803, https://doi.org/10.1107/S2053230X16015053 (2016).
15. Chalupska, D. *et al.* Structural analysis of phosphatidylinositol 4-kinase IIIbeta (PI4KB) - 14-3-3 protein complex reveals internal flexibility and explains 14-3-3 mediated protection from degradation *in vitro*. *Journal of structural biology*, https://doi.org/10.1016/j.jsb.2017.08.006 (2017).
16. Lamarche, M. J. *et al.* Anti-hepatitis C virus activity and toxicity of type III phosphatidylinositol-4-kinase beta inhibitors. *Antimicrobial agents and chemotherapy* **56**, 5149–5156, https://doi.org/10.1128/AAC.00946-12 (2012).
17. Raubo, P. *et al.* Discovery of potent, selective small molecule inhibitors of alpha-subtype of type III phosphatidylinositol-4-kinase (PI4KIIIalpha). *Bioorganic & medicinal chemistry letters* **25**, 3189–3193, https://doi.org/10.1016/j.bmcl.2015.05.093 (2015).
18. Mejdrova, I. *et al.* Highly Selective Phosphatidylinositol 4-Kinase IIIbeta Inhibitors and Structural Insight into Their Mode of Action. *Journal of medicinal chemistry* **58**, 3767–3793, https://doi.org/10.1021/acs.jmedchem.5b00499 (2015).
19. Rutaganira, F. U. *et al.* Design and Structural Characterization of Potent and Selective Inhibitors of Phosphatidylinositol 4 Kinase IIIbeta. *Journal of medicinal chemistry* **59**, 1830–1839, https://doi.org/10.1021/acs.jmedchem.5b01311 (2016).
20. Dorobantu, C. M. *et al.* Modulation of the Host Lipid Landscape to Promote RNA Virus Replication: The Picornavirus Encephalomyocarditis Virus Converges on the Pathway Used by Hepatitis C Virus. *PLoS pathogens* **11**, doi:ARTN e100518510.1371/journal.ppat.1005185 (2015).
21. Ishikawa-Sasaki, K., Sasaki, J. & Taniguchi, K. A complex comprising phosphatidylinositol 4-kinase IIIbeta, ACBD3, and Aichi virus proteins enhances phosphatidylinositol 4-phosphate synthesis and is critical for formation of the viral replication complex. *Journal of virology* **88**, 6586–6598, https://doi.org/10.1128/JVI.00208-14 (2014).
22. Wessels, E. *et al.* A viral protein that blocks Arf1-mediated COP-I assembly by inhibiting the guanine nucleotide exchange factor GBF1. *Developmental cell* **11**, 191–201, https://doi.org/10.1016/j.devcel.2006.06.005 (2006).
23. Wessels, E. *et al.* Molecular determinants of the interaction between coxsackievirus protein 3A and guanine nucleotide exchange factor GBF1. *Journal of virology* **81**, 5238–5245, https://doi.org/10.1128/JVI.02680-06 (2007).
24. Greninger, A. L., Knudsen, G. M., Betegon, M., Burlingame, A. L. & Derisi, J. L. The 3A protein from multiple picornaviruses utilizes the golgi adaptor protein ACBD3 to recruit PI4KIIIbeta. *Journal of virology* **86**, 3605–3616, https://doi.org/10.1128/JVI.06778-11 (2012).
25. Greninger, A. L., Knudsen, G. M., Betegon, M., Burlingame, A. L. & DeRisi, J. L. ACBD3 interaction with TBC1 domain 22 protein is differentially affected by enteroviral and kobuviral 3A protein binding. *mBio* **4**, e00098–00013, https://doi.org/10.1128/mBio.00098-13 (2013).
26. Lama, J., Paul, A. V., Harris, K. S. & Wimmer, E. Properties of purified recombinant poliovirus protein 3aB as substrate for viral proteinases and as co-factor for RNA polymerase 3Dpol. *The Journal of biological chemistry* **269**, 66–70 (1994).
27. Arita, M. Mechanism of Poliovirus Resistance to Host Phosphatidylinositol-4 Kinase III beta Inhibitor. *ACS Infect Dis* **2**, 140–148, https://doi.org/10.1021/acsinfecdis.5b00122 (2016).
28. Lyle, J. M. *et al.* Similar structural basis for membrane localization and protein priming by an RNA-dependent RNA polymerase. *The Journal of biological chemistry* **277**, 16324–16331, https://doi.org/10.1074/jbc.M112429200 (2002).
29. Fujita, K. *et al.* Membrane topography of the hydrophobic anchor sequence of poliovirus 3A and 3AB proteins and the functional effect of 3A/3AB membrane association upon RNA replication. *Biochemistry* **46**, 5185–5199, https://doi.org/10.1021/bi6024758 (2007).
30. Plotch, S. J. & Palant, O. Poliovirus protein 3AB forms a complex with and stimulates the activity of the viral RNA polymerase, 3Dpol. *Journal of virology* **69**, 7169–7179 (1995).
31. Paul, A. V., van Boom, J. H., Filippov, D. & Wimmer, E. Protein-primed RNA synthesis by purified poliovirus RNA polymerase. *Nature* **393**, 280–284, https://doi.org/10.1038/30529 (1998).
32. Peersen, O. B. Picornaviral polymerase structure, function, and fidelity modulation. *Virus research* **234**, 4–20, https://doi.org/10.1016/j.virusres.2017.01.026 (2017).
33. Tong, J., Yang, H., Yang, H., Eom, S. H. & Im, Y. J. Structure of Osh3 reveals a conserved mode of phosphoinositide binding in oxysterol-binding proteins. *Structure* **21**, 1203–1213, https://doi.org/10.1016/j.str.2013.05.007 (2013).
34. Arita, M. Phosphatidylinositol-4 kinase III beta and oxysterol-binding protein accumulate unesterified cholesterol on poliovirus-induced membrane structure. *Microbiol Immunol* **58**, 239–256, https://doi.org/10.1111/1348-0421.12144 (2014).
35. Roulin, P. S. *et al.* Rhinovirus uses a phosphatidylinositol 4-phosphate/cholesterol counter-current for the formation of replication compartments at the ER-Golgi interface. *Cell host & microbe* **16**, 677–690, https://doi.org/10.1016/j.chom.2014.10.003 (2014).
36. Chung, J. *et al.* Intracellular Transport. PI4P/phosphatidylserine countertransport at ORP5- and ORP8-mediated ER-plasma membrane contacts. *Science* **349**, 428–432, https://doi.org/10.1126/science.aab1370 (2015).
37. Moser von Filseck, J., Vanni, S., Mesmin, B., Antonny, B. & Drin, G. A phosphatidylinositol-4-phosphate powered exchange mechanism to create a lipid gradient between membranes. *Nature communications* **6**, 6671, https://doi.org/10.1038/ncomms7671 (2015).
38. Dolinsky, S. *et al.* The Legionella longbeachae Icm/Dot substrate SidC selectively binds phosphatidylinositol 4-phosphate with nanomolar affinity and promotes pathogen vacuole-endoplasmic reticulum interactions. *Infection and immunity* **82**, 4021–4033, https://doi.org/10.1128/IAI.01685-14 (2014).
39. Hsu, N. Y. *et al.* Viral reorganization of the secretory pathway generates distinct organelles for RNA replication. *Cell* **141**, 799–811, https://doi.org/10.1016/j.cell.2010.03.050 (2010).
40. Hurley, J. H., Boura, E., Carlson, L. A. & Rozycki, B. Membrane budding. *Cell* **143**, 875–887, https://doi.org/10.1016/j.cell.2010.11.030 (2010).
41. McCullough, J., Colf, L. A. & Sundquist, W. I. Membrane fission reactions of the mammalian ESCRT pathway. *Annual review of biochemistry* **82**, 663–692, https://doi.org/10.1146/annurev-biochem-072909-101058 (2013).
42. Rozycki, B., Boura, E., Hurley, J. H. & Hummer, G. Membrane-elasticity model of Coatless vesicle budding induced by ESCRT complexes. *PLoS computational biology* **8**, e1002736, https://doi.org/10.1371/journal.pcbi.1002736 (2012).
43. Wollert, T., Wunder, C., Lippincott-Schwartz, J. & Hurley, J. H. Membrane scission by the ESCRT-III complex. *Nature* **458**, 172–177, https://doi.org/10.1038/nature07836 (2009).
44. Carlson, L. A. & Hurley, J. H. *In vitro* reconstitution of the ordered assembly of the endosomal sorting complex required for transport at membrane-bound HIV-1 Gag clusters. *Proceedings of the National Academy of Sciences of the United States of America* **109**, 16928–16933, https://doi.org/10.1073/pnas.1211759109 (2012).
45. Carlson, L. A., Bai, Y., Keane, S. C., Doudna, J. A. & Hurley, J. H. Reconstitution of selective HIV-1 RNA packaging *in vitro* by membrane-bound Gag assemblies. *eLife* **5**, https://doi.org/10.7554/eLife.14663 (2016).
46. Saleem, M. *et al.* A balance between membrane elasticity and polymerization energy sets the shape of spherical clathrin coats. *Nature communications* **6**, 6249, https://doi.org/10.1038/ncomms7249 (2015).
47. Renard, H. F. *et al.* Endophilin-A2 functions in membrane scission in clathrin-independent endocytosis. *Nature* **517**, 493–496, https://doi.org/10.1038/nature14064 (2015).
48. Pabst, G., Kucerka, N., Nieh, M. P. & Katsaras, J. Liposomes, Lipid Bilayers and ModelMembranes From Basic Research to Application Preface. *Liposomes, Lipid Bilayers and Model Membranes: From Basic Research to Application*, Ix–X, https://doi.org/10.1201/b16617 (2014).
49. Klima, M. *et al.* Kobuviral Non-structural 3A Proteins Act as Molecular Harnesses to Hijack the Host ACBD3 Protein. *Structure* **25**, 219–230, https://doi.org/10.1016/j.str.2016.11.021 (2017).
50. McPhail, J. A., Ottosen, E. H., Jenkins, M. L. & Burke, J. E. The Molecular Basis of Aichi Virus 3A Protein Activation of Phosphatidylinositol 4 Kinase IIIbeta, PI4KB, through ACBD3. *Structure* **25**, 121–131, https://doi.org/10.1016/j.str.2016.11.016 (2017).

51. Wang, J., Ptacek, J. B., Kirkegaard, K. & Bullitt, E. Double-membraned liposomes sculpted by poliovirus 3AB protein. *The Journal of biological chemistry* **288**, 27287–27298, https://doi.org/10.1074/jbc.M113.498899 (2013).
52. Fan, J., Liu, J., Culty, M. & Papadopoulos, V. Acyl-coenzyme A binding domain containing 3 (ACBD3; PAP7; GCP60): an emerging signaling molecule. *Progress in lipid research* **49**, 218–234, https://doi.org/10.1016/j.plipres.2009.12.003 (2010).
53. Teoule, F. *et al.* The Golgi protein ACBD3, an interactor for poliovirus protein 3A, modulates poliovirus replication. *Journal of virology* **87**, 11031–11046, https://doi.org/10.1128/JVI.00304-13 (2013).
54. Wollert, T. & Hurley, J. H. Molecular mechanism of multivesicular body biogenesis by ESCRT complexes. *Nature* **464**, 864–869, https://doi.org/10.1038/nature08849 (2010).
55. Mejdrova, I. *et al.* Rational Design of Novel Highly Potent and Selective Phosphatidylinositol 4-Kinase IIIbeta (PI4KB) Inhibitors as Broad-Spectrum Antiviral Agents and Tools for Chemical Biology. *Journal of medicinal chemistry* **60**, 100–118, https://doi.org/10.1021/acs.jmedchem.6b01465 (2017).
56. Humpolickova, J., Mejdrova, I., Matousova, M., Nencka, R. & Boura, E. Fluorescent Inhibitors as Tools To Characterize Enzymes: Case Study of the Lipid Kinase Phosphatidylinositol 4-Kinase IIIbeta (PI4KB). *Journal of medicinal chemistry* **60**, 119–127, https://doi.org/10.1021/acs.jmedchem.6b01466 (2017).
57. Spagnolo, J. F., Rossignol, E., Bullitt, E. & Kirkegaard, K. Enzymatic and nonenzymatic functions of viral RNA-dependent RNA polymerases within oligomeric arrays. *Rna* **16**, 382–393, https://doi.org/10.1261/rna.1955410 (2010).
58. Pathak, H. B., Arnold, J. J., Wiegand, P. N., Hargittai, M. R. & Cameron, C. E. Picornavirus genome replication: assembly and organization of the VPg uridylylation ribonucleoprotein (initiation) complex. *The Journal of biological chemistry* **282**, 16202–16213, https://doi.org/10.1074/jbc.M610608200 (2007).
59. Altan-Bonnet, N. Lipid Tales of Viral Replication and Transmission. *Trends in cell biology* **27**, 201–213, https://doi.org/10.1016/j.tcb.2016.09.011 (2017).
60. Marcotte, L. L. *et al.* Crystal structure of poliovirus 3CD protein: virally encoded protease and precursor to the RNA-dependent RNA polymerase. *Journal of virology* **81**, 3583–3596, https://doi.org/10.1128/JVI.02306-06 (2007).
61. Paul, A. V., Rieder, E., Kim, D. W., van Boom, J. H. & Wimmer, E. Identification of an RNA hairpin in poliovirus RNA that serves as the primary template in the *in vitro* uridylylation of VPg. *Journal of virology* **74**, 10359–10370 (2000).
62. Hope, D. A., Diamond, S. E. & Kirkegaard, K. Genetic dissection of interaction between poliovirus 3D polymerase and viral protein 3AB. *Journal of virology* **71**, 9490–9498 (1997).
63. Gong, P., Kortus, M. G., Nix, J. C., Davis, R. E. & Peersen, O. B. Structures of coxsackievirus, rhinovirus, and poliovirus polymerase elongation complexes solved by engineering RNA mediated crystal contacts. *PloS one* **8**, e60272, https://doi.org/10.1371/journal.pone.0060272 (2013).
64. Nemecek, D. *et al.* Subunit folds and maturation pathway of a dsRNA virus capsid. *Structure* **21**, 1374–1383, https://doi.org/10.1016/j.str.2013.06.007 (2013).
65. Rezabkova, L. *et al.* 14-3-3 protein interacts with and affects the structure of RGS domain of regulator of G protein signaling 3 (RGS3). *Journal of structural biology* **170**, 451–461, https://doi.org/10.1016/j.jsb.2010.03.009 (2010).
66. Boura, E., Ivanov, V., Carlson, L. A., Mizuuchi, K. & Hurley, J. H. Endosomal sorting complex required for transport (ESCRT) complexes induce phase-separated microdomains in supported lipid bilayers. *The Journal of biological chemistry* **287**, 28144–28151, https://doi.org/10.1074/jbc.M112.378646 (2012).
67. Schindelin, J., Rueden, C. T., Hiner, M. C. & Eliceiri, K. W. The ImageJ ecosystem: An open platform for biomedical image analysis. *Molecular reproduction and development* **82**, 518–529, https://doi.org/10.1002/mrd.22489 (2015).
68. Boura, E. & Hurley, J. H. Structural basis for membrane targeting by the MVB12-associated beta-prism domain of the human ESCRT-I MVB12 subunit. *Proceedings of the National Academy of Sciences of the United States of America* **109**, 1901–1906, https://doi.org/10.1073/pnas.1117597109 (2012).
69. Wahl, M., Gregor, I., Patting, M. & Enderlein, J. Fast calculation of fluorescence correlation data with asynchronous time-correlated single-photon counting. *Optics express* **11**, 3583–3591 (2003).
70. Gregor, I. & Enderlein, J. Time-resolved methods in biophysics. 3. Fluorescence lifetime correlation spectroscopy. *Photochem Photobiol Sci* **6**, 13–18, https://doi.org/10.1039/b610310c (2007).

Acknowledgements

The work was supported by the Czech Science Foundation grant number 15–21030Y. The Academy of Sciences of the Czech Republic (RVO: 61388963) is also acknowledged. We are grateful to Prof. Joseph DeRisi (UCSF School of Medicine, San Francisco, CA) for sharing the synthetic Aichi virus genome, to Prof. Carolyn Machamer (Johns Hopkins University, Baltimore, MD) for sharing the ACBD3 encoding plasmid and to Prof. O. Peersen (Colorado State University) for sharing the pCG1 expression plasmid. We are also grateful to Dr. Edward Curtis for critical reading of the manuscript.

Author Contributions

A.D., J.H. and M.K. performed experiments. E.B. designed and supervised the project. E.B. wrote the manuscript, all authors contributed to data analysis and commented on the manuscript.

Additional Information

Supplementary information accompanies this paper at https://doi.org/10.1038/s41598-017-17621-6.

Competing Interests: The authors declare that they have no competing interests.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Journal of Structural Biology xxx (xxxx) xxx–xxx

Contents lists available at ScienceDirect

Journal of Structural Biology

journal homepage: www.elsevier.com/locate/yjsbi

Structures of kobuviral and siciniviral polymerases reveal conserved mechanism of picornaviral polymerase activation

Anna Dubankova, Vladimira Horova, Martin Klima, Evzen Boura*

Institute of Organic Chemistry and Biochemistry of the Czech Academy of Sciences, Flemingovo nam. 2., 166 10 Prague 6, Czech Republic

ARTICLE INFO

Keywords:
RNA
Polymerase
Crystal structure
Kobuvirus
Picornavirus

ABSTRACT

RNA-dependent RNA polymerase $3D^{pol}$ is a key enzyme for the replication of picornaviruses. The viral genome is translated into a single polyprotein that is subsequently proteolytically processed into matured products. The $3D^{pol}$ enzyme arises from a stable 3CD precursor that has high proteolytic activity but no polymerase activity. Upon cleavage of the precursor the newly established N-terminus of $3D^{pol}$ is liberated and inserts itself into a pocket on the surface of the $3D^{pol}$ enzyme. The essential residue for this mechanism is the very first glycine that is conserved among almost all picornaviruses. However, kobuviruses and siciniviruses have a serine residue instead. Intrigued by this anomaly we sought to solve the crystal structure of these $3D^{pol}$ enzymes. The structures revealed a unique fold of the $3D^{pol}$ N-termini but the very first serine residues were inserted into a charged pocket in a similar manner as the glycine residue in other picornaviruses. These structures revealed a common underlying mechanism of $3D^{pol}$ activation that lies in activation of the α10 helix containing a key catalytical residue Asp238 that forms a hydrogen bond with the 2′ hydroxyl group of the incoming NTP nucleotide.

1. Introduction

Positive-sense single stranded RNA viruses (+RNA viruses) are a large group of viruses (e.g. Picornavirales, Nidovirales, Tymovirales) that are comprised of many families. The *Picornaviridae* family includes many dangerous human pathogens such as poliovirus (PV), coxsackievirus, hepatitis A virus and some ostensibly less dangerous pathogens such as rhinoviruses (common cold) and kobuviruses (gastroenteritis). The picornaviral genome also serves as mRNA and is translated into a single polyprotein that is later processed into structural (VP0–VP4) and non-structural (2A, 2B, 2C, 3A, 3B, 3C, 3D) proteins, a leader (L) protein is also present in some picornaviruses (Fig. 1A) that together with many host factors ensure viral replication. The replication takes place at membranous replication organelles (ROs) that serve as a rendezvous place for viral proteins and perhaps also as a shelter from innate immunity (Harak and Lohmann, 2015; Nagy et al., 2016). ROs are usually multilamellar and enriched in specific lipids such as phosphatidylinositol 4-phosphate (PI4P) and cholesterol (Altan-Bonnet, 2017). It was suggested that PI4P directly recruits the viral polymerase to the ROs (Hsu et al., 2010) although recent results suggest that a negative charge in general rather than PI4P specifically is responsible for $3D^{pol}$ recruitment to the membrane (Dubankova et al., 2017).

A key enzyme for replication of RNA viruses is the RNA-dependent RNA polymerase (RdRp – termed $3D^{pol}$ in picornaviruses) because such an enzyme is not present in the eukaryotic cell and each RNA virus must encode for it. The mature $3D^{pol}$ enzyme arises from the cleavage of the stable 3CD precursor that itself arises from the single polyprotein. Interestingly, 3CD has no polymerase activity but does have high proteolytical activity (Harris et al., 1992; Sun et al., 2016). The prototypical $3D^{pol}$ enzyme is the PV polymerase that is best biochemically characterized (for a thorough recent review, see (Peersen, 2017)). $3D^{pol}$ has many unique features. Most importantly, it has low fidelity, which is responsible for the origin and distribution of quasispecies (mutant clouds established in infected organisms) (Andino and Domingo, 2015). Interestingly, it uses a protein primer – a short peptide (~22 AAs) called 3B or VPg (from viral protein genome linked) – that has a conserved tyrosine residue at position 3 and its hydroxyl group serves as the attachment point for the first nucleotide (Paul et al., 1998). Deep molecular insight into the structure of picornaviral $3D^{pol}$ enzymes was gained over the last decades as many were crystallized including the PV, coxsackie, FMDV, rhinovirus and enterovirus 71 polymerase (Campagnola et al., 2008; Ferrer-Orta et al., 2004; Hansen et al., 1997; Love et al., 2004; Thompson and Peersen, 2004; Wu et al., 2010) which were also crystallized in complex with RNA (Ferrer-Orta et al., 2004; Ferrer-Orta et al., 2007; Gong and Peersen, 2010; Gong et al., 2013). All the structures show the conserved right-hand fold typical for viral RNA

* Corresponding author.
E-mail address: boura@uochb.cas.cz (E. Boura).

https://doi.org/10.1016/j.jsb.2019.08.004
Received 28 June 2019; Received in revised form 8 August 2019; Accepted 9 August 2019

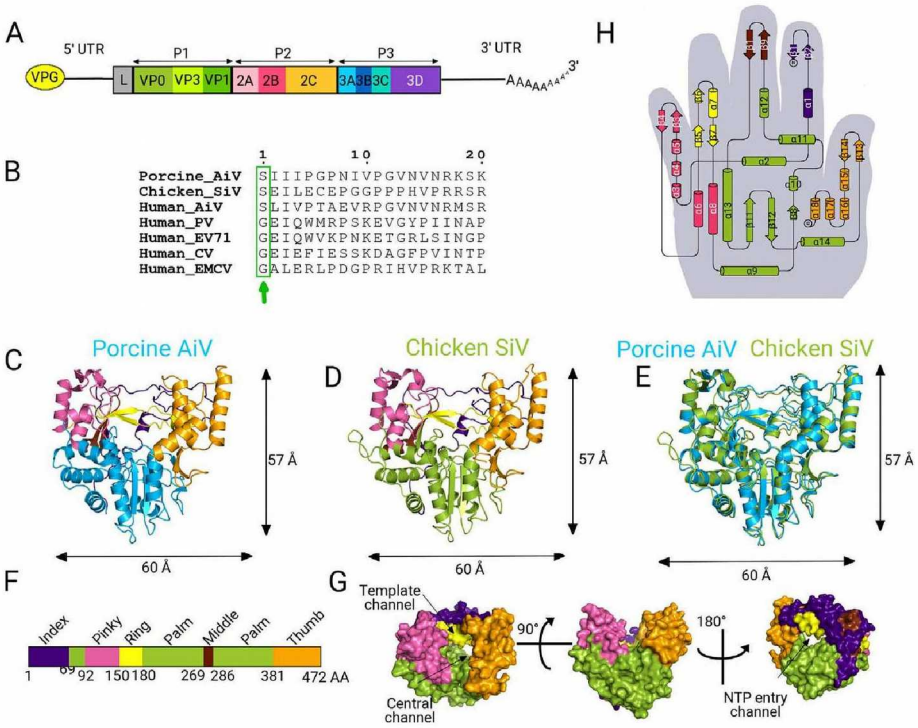


Fig. 1. Overall fold of Aichi 3D^{pol} enzyme. A) Aichi RNA genome organization: P1 region encodes structural proteins VP0, VP3 and VP1. While the P2 and P3 regions encode for non-structural proteins. The sequence encoding 3D^{pol} gene is located on the very 3′ end. B) Multiple alignment illustrates that kobuviruses (AiV) have conserved serine residue at the very N-termini of 3D^{pol} where all other picornaviruses have a glycine residue C and D) Structure of the porcine AiV 3D^{pol} with palm, thumb and fingers (index, middle, ring, pinky) domains. E) Superposition of the AiV and SiV 3D^{pol} enzymes reveals virtually identical fold. G) The location of the template, central and NTP entry channel within the 3D^{pol} AiV structure. H) The 3D^{pol} topology plot drawn on a silhouette of a right hand.

polymerases with the central RNA template channel. Interestingly, all the above mentioned RNA polymerases share the very first conserved glycine residues. Comparison of the crystal structure of the precursor protein 3CD with the mature 3D^{pol} protein (Marcotte et al., 2007) reveals that upon cleavage the N-terminus undergoes a conformational change which is essential for 3D^{pol} activation. For instance, a single amino acid residue in excess renders the 3D^{pol} enzyme inactive, which is important to take into account when producing the protein recombinantly (Gohara et al., 1999). Only seven genera (Kobuviruses, Galliviruses, Osciviruses, Passeriviruse, Sakobuviruses, Saliviruses and Siciniviruses) of the *Picornaviridae* family do not possess this conserved glycine residue and have a serine residue instead (SI Fig. 1).

Kobuviruses infect many mammalian species. The human kobuvirus was first isolated in the Japanese prefecture Aichi from a patient with acute gastroenteritis and was afterwards termed human Aichi virus (Yamashita et al., 1991). Aichi virus is a very common pathogen (more than 80% of adults have antibodies against it) and globally distributed (Kitajima and Gerba, 2015). Recently, Aichi virus also attracted scientific attention. It was crystallized revealing the mechanism of kobuvirus genome release (Sabin et al., 2016) and its replication was described as PI4KB (phosphatidylinositol 4-kinase B, one of four human enzymes that synthesize PI4P) dependent (Greninger et al., 2012; Sasaki et al., 2012). Congruently, inhibitors of PI4KB inhibit Aichi virus replication (Mejdrova et al., 2015; Mejdrova et al., 2017). Recently we solved the crystal structure of the Aichi 3A protein in complex with the GOLD domain of ACBD3 (Acyl-CoA-binding domain-containing protein 3, a Golgi resident protein and a nanomolar binder of PI4KB). The structure revealed that the 3A protein that is intrinsically disordered by itself, however, in the complex adopts a highly ordered structure and literally wraps around the GOLD domain of ACBD3 to hijack PI4KB to bring it to target membrane and activate it (Klima et al., 2017; Klima et al., 2016), a mechanism that is shared with enteroviruses (Horova et al., 2019). Confirming results were also reached using mass spectroscopy (McPhail et al., 2017). Still many features of kobuviruses including mechanism of replication remain unexplained. Sicinivirus is a new picornavirus genus that remains to be described in detail (Bullman et al., 2014). Here, we describe the atomic structures of the key enzyme – the RNA-dependent RNA polymerase 3D^{pol} of the Aichi virus (kobuvirus) and the Sicinivirus A.

2. Results

We sought to solve the crystal structure of the 3D^{pol} enzymes that have unique sequences at their N-termini – a serine residue where all other picornaviruses have a glycine residue (Fig. 1B, SI Fig. 1). We have screened several 3D^{pol} enzymes for the ability to form well diffracting crystals (Table 1) and discovered that the 3D^{pol} from the porcine Aichi virus and chicken Sicinivirus A form well diffracting crystals after the introduction of crystal-stabilizing mutations (detailed in M&M section).

The structures were solved by molecular replacement and revealed the classical right-hand fold consisting of fingers palm and thumb, with fingers subdivided in index, middle, ring and pinky (Fig. 1). The enzyme is comprised of 18 helices and 14 beta strands that make up the

Table 1
Statistics of crystallographic data collection and refinement.

Data collection		
Crystal	Sicinivirus 3D^{pol}	Aichi virus 3D^{pol}
Space group	$P4_22_12$	$P6_122$
Cell dimension	a = b = 92.5 Å, c = 161.0 Å	a = b = 170.5 Å, 231.3 Å
X-ray source	BESSY ID 14-1	BESSY ID 14-1
Resolution (Å)	46.62–2.30 (2.38–2.30)	48.13–2.64 (2.73–2.64)
Unique reflections	31,996 (3120)	59,076 (5790)
I/σ (I)	13.40 (2.10)	33.96 (4.98)
Wavelength (Å)	0.9184	0.9184
Multiplicity	14.0 (9.8)	21.9 (22.2)
Completeness (%)	99.93 (99.58)	99.96 (100.00)
R-merge, %	19.1 (100)	8.5 (67.8)
$CC_{1/2}$	0.997 (0.791)	1 (0.94)
Refinement		
R-work, %	19.18 (25.95)	19.10 (24.16)
R-free, %	22.46 (31.45)	22.19 (26.43)
RMS (bonds), °	0.009	0.004
RMS (angles), Å	1.11	0.76
Ramachandran (outliers/favored) (%)	0.85%/97.66%	0%/98.28%

Numbers in parentheses refer to the highest resolution shell.

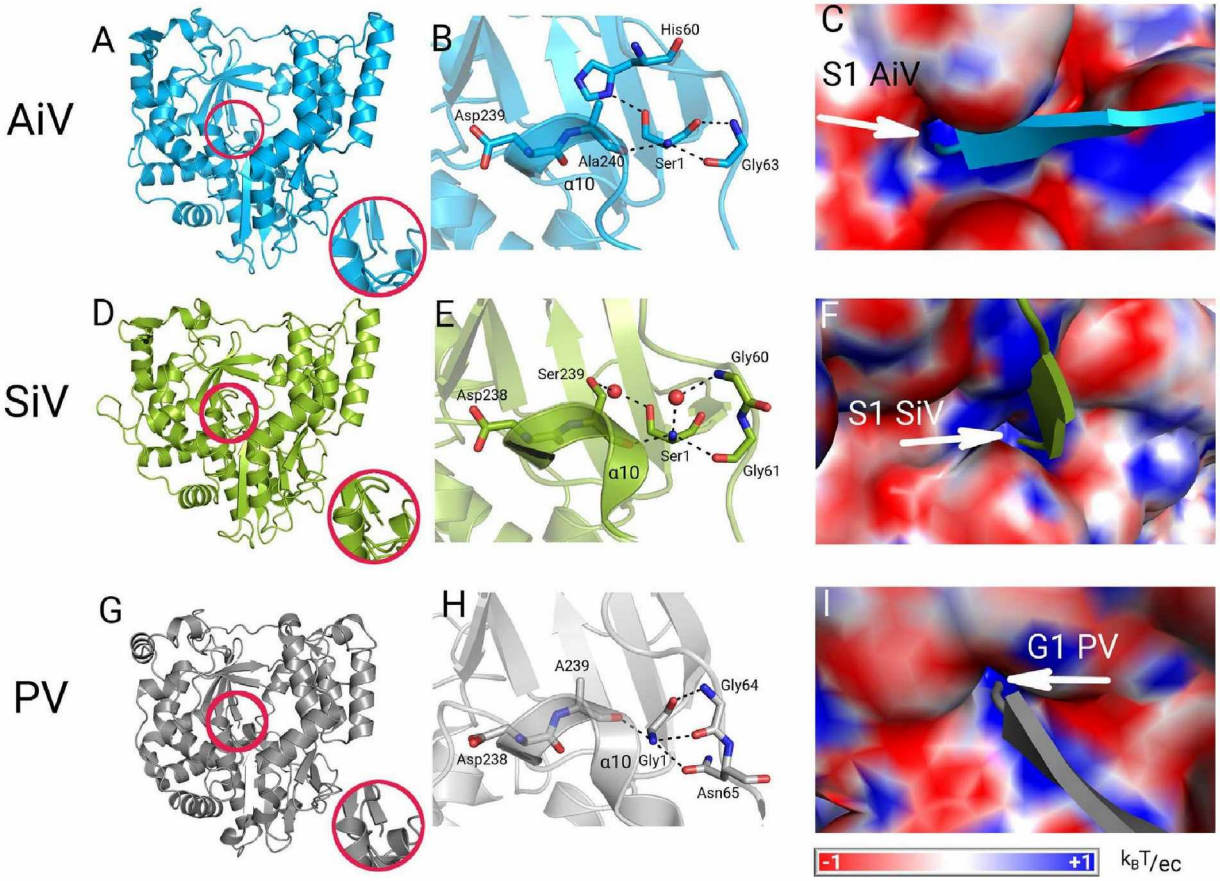


Fig. 2. Atypical fold of N termini of AiV and SiV 3D^{pol} among 3D^{pol} enzymes. A-C) The first residue of AiV 3D^{pol} is locked to a positively charged cavity leading to protein center by residues H40, G63 and A240. D-F) Corresponding view of the SiV 3D^{pol} structure. G-I) All other picornaviruses have the first glycine residue inserted in a charged cavity as illustrated by the PV 3D^{pol} structure.

six subdomains that we recognize within the polymerase family: thumb, index (α15-18, β13-14), middle (β9-10), ring (α7, β5-7), and pinky (α3-6, α8, β3-4) fingers and the palm (α2, α9-13, β8, β11-12) (Fig. 1H). We could also identify the three characteristic channels, the template binding, central and NTP entry channel (Fig. 1D) (Ferrer-Orta et al., 2006; Peersen, 2017). The template channel is surrounded by the palm domain and the ring and pinky fingers. It is further stabilized by the middle finger and wrapped around by the index finger. The central channel is located next to the template channel and is surrounded mainly by the palm domain and the ring finger. The third channel is the NTP entry channel. The structure of the NTP entry channel is created by the interaction of the ring finger and the palm domain. The index and thumb fingers are wrapped around the NTP channel and stabilize it (Fig. 1).

As mentioned, the first N-terminal residue (a glycine in all previously solved structures) of picornaviral polymerases inserts itself into a pocked within the enzyme (Fig. 2). The first residue is absolutely essential for the 3D^{pol} activity and is also well visible in our structures (SI Fig. 2) and its deletion or mutation leads to an inactive enzyme as was shown for the Coxsackie virus 3D^{pol} (Campagnola et al., 2008). In case of PV the first glycine residue forms a hydrogen bond with N65 (Thompson and Peersen, 2004). The first AiV serine residue forms hydrogen bonds with H60, G63 and A240 while the first SiV 3D^{pol} serine residue forms a water bridge with S239 (Fig. 2, for superposition with other picornaviral polymerases, please, see SI Fig. 3). We aimed to test the importance of the first serine residue using mutational analysis. Unfortunately, the AiV and SiV 3D^{pol} enzymes have only negligible activity *in vitro* under the same conditions where the PV polymerase is active (SI Fig. 4). Therefore, we devised a subgenomic replicon assay in cells. First, we replaced the P1 genomic fragment encoding the viral structural proteins by a sequence encoding the fluorescent protein mCherry generating the pUCSF-AiV/mCherry plasmid. Next, we introduced several mutations (Ser1Gly, Ser1Ala, +1Ser, +1Gly, +2Ser, and Δ3D^{pol}) and we tested the replication of the appropriate RNA in cell culture using FACS analysis (fluorescence-activated cell sorter, detailed in M&M) (Fig. 3). We observed efficient replication of the wild type RNA. The intensity of fluorescence in the case of the 3D^{pol} deletion mutant (Δ3D^{pol}) was two orders of magnitude lower and we consider this value as zero because the Δ3D^{pol} RNA cannot replicate and the whole signal comes from the translation of the original transfected RNA only. No replication was also observed for all the mutants with additional N-terminal residues (+1x Ser, +1x Gly, +2x Ser) and only very inefficient replication (~5% of wt) for Ser1Gly and Ser1Ala mutants. Ser1Ala mutant RNA also replicated inefficiently (~10% of wt). However, the 3CD precursor processing test revealed that the mutants + 1Ser and + 2Ser are rapidly degraded in cells (Fig. 3D) at the same conditions where the wild type 3CD is properly processed as expected and so are the mutants Ser1Gly, Ser1Ala, and + 1Gly. The mutations + 1Ser and + 2Ser are deleterious for the enzyme in cells and thus enzymatic activity of these mutants could not be tested.

These results demonstrate that a proper processing of the 3D^{pol} N-terminus is absolutely essential in the case of 3D^{pol} enzymes possessing a serine residue in the first position just as for the typical glycine-starting polymerases. Even mutation of the first serine residue to glycine mimicking all other picornaviruses is not tolerated.

3. Discussion

The main motivation of this study was to decipher the role of the very first 3D^{pol} serine residue that is present in the kobuvirus and

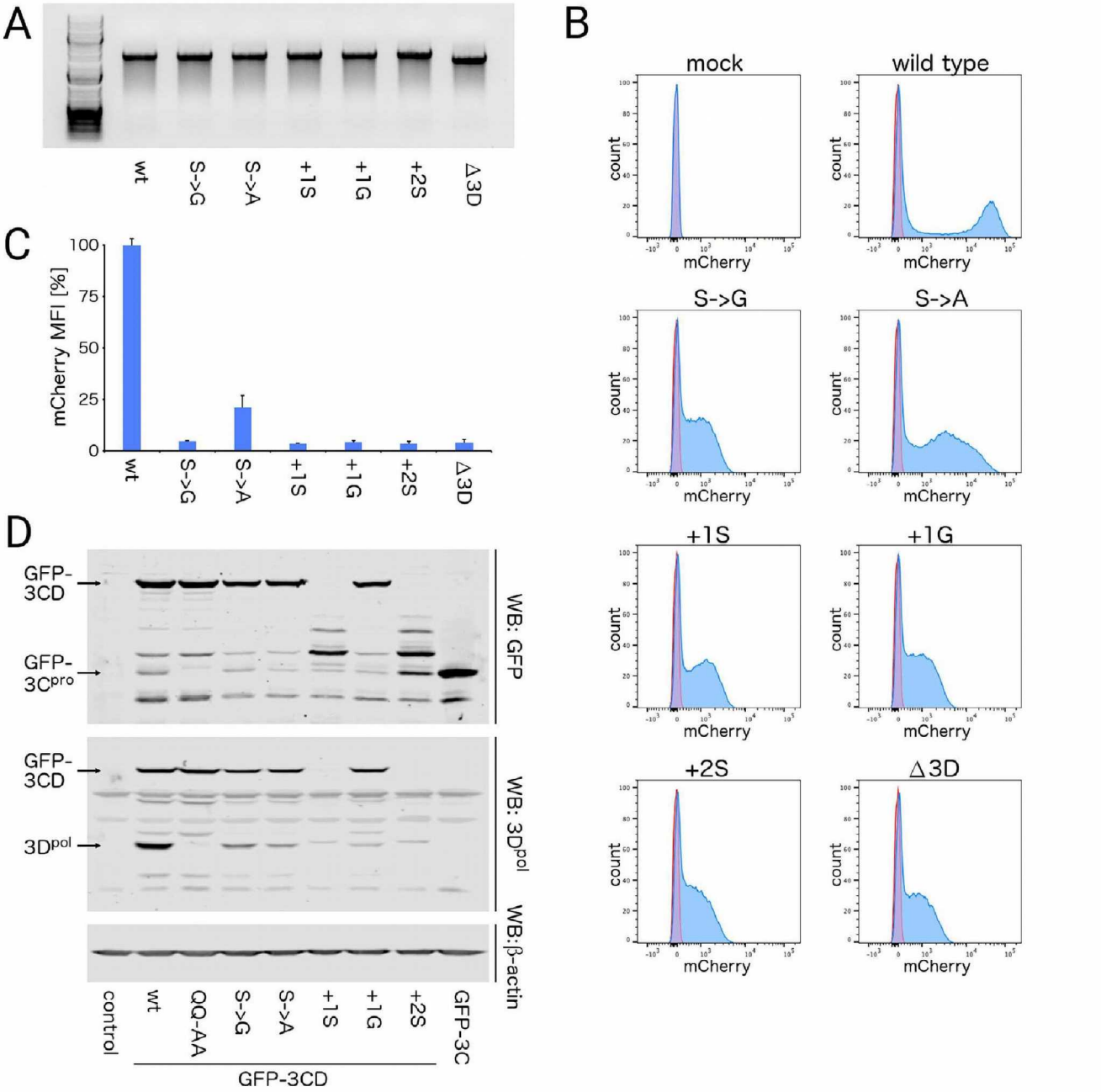


Fig. 3. FACS analysis of Aichi virus replication. A) Quality of RNA used for the viral replicon assay analyzed on the 1% agarose gel. B and C) Viral replicon assay. HeLa cells were transfected with the T7-amplified aichivirus subgenomic replicon wild-type RNA or its mutants within the $3D^{pol}$ region as indicated, and the reporter mCherry fluorescence was determined by flow cytometry. The data are presented as histograms of mCherry fluorescence from one representative experiment (B) or as the mean mCherry fluorescence ± standard error of the mean (SEM) from two independent experiments (C). D) AiV 3CD processing analysis. HEK293T cells were transfected with the EGFP-fused wild-type 3CD or its mutants as indicated, and the cell lysates were analyzed by immunoblotting using the anti-GFP and anti-AiV $3D^{pol}$ antibodies.

sicinivirus genera where most other members of the *Picornaviridae* family have a conserved glycine residue. Also all previously reported structures of picornaviral polymerases have the glycine residue in the first position. The $3D^{pol}$ N-terminus is liberated upon the cleavage of the stable 3CD precursor and is essential for the activation of the $3D^{pol}$ enzyme. At least two mechanisms how the $3D^{pol}$ N-terminus activates the polymerase activity were suggested. Campagnola et al. noticed a conserved hydrophobic residue (phenylalanine or tryptophan) at position 5 that is solvent exposed due to the unusual hydrogen bonding within the first two antiparallel β-sheets (Campagnola et al., 2008). This hydrophobic residue (Phe5 or Trp5) has a profound effect on the enzymatic activity because it is involved in the rate limiting NTP repositioning step through its interaction with the residue Tyr62. FMDV has an aspartate on the fifth position but this aspartate interacts with His64 which is structurally equivalent to Tyr62 (Ferrer-Orta et al., 2004). However, SiV $3D^{pol}$ has a glutamate residue and AiV $3D^{pol}$ a proline residue at position 5 (Fig. 1B). It is very unlikely that glutamate or proline and histidine could form a salt bridge in a similar fashion as aspartate and histidine as in the case of FMDV.

Based on the structure of the PV $3D^{pol}$ enzyme (Thompson and Peersen, 2004) it was also suggested that the function of the very first $3D^{pol}$ residue is to stabilize Asp238, a conserved residue located in a short α10 helix that forms a hydrogen bond with the 2′ hydroxyl group of incoming NTP nucleotide. This interaction is structurally conserved

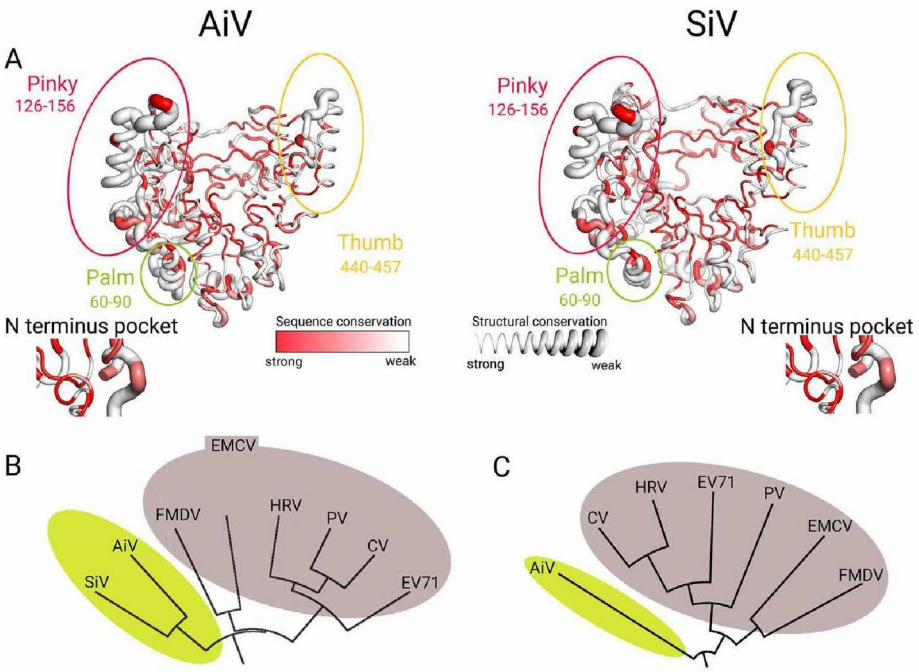


Fig. 4. Structural comparison of Aichi 3D^{pol} with other picornaviral 3D^{pol} enzymes. A) Sausage representation that illustrates the three-dimensional structure conservation of AiV and SiV 3D^{pol} according to thickness of the ribbon (the thickest is least conserved) and the conservation of primary sequence (least conserved regions are labeled in white). The radius of ribbons is proportional to the difference of Cα position between AiV or SiV 3D^{pol} and all homologous picornaviral polymerases of known structures. The major distinction was observed in the pinky and finger region and in helix α2. Structures of AiV and SiV polymerase are similar to each other's, but different to the others so far known 3D^{pol} structures of the *Picornaviridae* family. B) A structure based phylogenetic tree illustrating the apparent independent evolution of AiV and SiV polymerases. C) Analogous phylogenetic tree based on a structural alignment of capsid protomers from various picornaviruses is in agreement with the phylogenetic tree based on known 3D^{pol} structures.

in kobuviruses and sicinivirus. The first serine residue in AiV 3D^{pol} forms hydrogen bonds with His60, Gly63 and Ala240 to stabilize the α10 helix. The first serine residue of SiV forms a hydrogen bond and a water bridge with Ser239, a neighboring residue of Asp238 and thus also stabilizes the short α10 helix albeit using a different mechanism (Fig. 2). We conclude that activation of the α10 helix by the liberated N-terminus is a feature shared in all the structures of picornaviral polymerases and we suggest that this is the mechanism of 3D^{pol} activation upon proteolytical cleavage.

The N-termini of AiV and SiV 3D^{pol} are the least conserved regions when compared to other picornaviruses. We performed a structural comparison of the AiV and SiV 3D^{pol} enzymes with all the solved structures of picornaviral 3D^{pol} enzymes available in the PDB database using the ENDscript 2.0 (Robert and Gouet, 2014) (Fig. 4, low conserved regions are in thick ribbons). The analysis revealed that besides the N-termini also the helix α2, and the pinky and thumb fingers of the AiV and SiV 3D^{pol} are low conserved regions.

Next, we constructed a phylogenetic tree based on all the solved 3D^{pol} structures available in the PDB database by the lowest RMSD (root-mean-square deviation of atomic positions) method which illustrates that kobuviruses and sicinivirus are evolutionary distinct group within the *Picornaviridae* family (Fig. 4B). Very similar results were obtained when structures of viral capsid proteins were used (Fig. 4C) (Sabin et al., 2016) suggesting that the polymerase and the viral capsid started to evolve separately from other picornaviruses at the same time. This analysis sets the kobuvirus and sicinivirus genera apart from other picornaviruses, however, it also highlights the conserved mechanism of 3D^{pol} activation upon proteolytical processing.

4. Materials and methods

Protein expression and purification – Artificial codon optimized genes encoding the AiV and SiV 3D^{pol} enzymes (UniProtKB – B8R1T8_9PICO, A0A0P0QLC4_9PICO) were commercially synthesized (Invitrogen) and cloned into a pSUMO vector resulting into genes encoding for His_{8x}SUMO-3D^{pol} fusion proteins. To improve crystallization properties we introduced surface entropy reduction mutations (K57A, Q58A, R118L for SiV 3D^{pol} and E93,94A for AiV 3D^{pol}) (Goldschmidt et al., 2007). The protein was expressed and purified using our standard protocols for viral polymerases (Dubankova and Boura, 2019; Sebera et al., 2018). Briefly, it was expressed in the *E. coli* strain BL21 Star. After lysis the protein was immobilized on the NiNTA Agarose resin (Machery-Nagel) and extensively washed with the wash buffer (50 mM Tris pH = 8, 300 mM NaCl, 20 mM imidazole, 3 mM βME) and eluted with the wash buffer supplemented by 300 mM imidazole. The His_{8x}SUMO tag was cleaved by the Ulp1 protease and the protein was further purified by size exclusion chromatography. Finally, the protein was concentrated to 6 mg/ml and stored at −80 °C until needed.

Rabbit polyclonal antiserum against AiV 3D^{pol} was generated commercially (Exbio) using the recombinant AiV 3D^{pol} protein as immunogen. The polyclonal antibody to AiV 3D^{pol} was further purified by affinity chromatography using the HiTrap Protein A HP column (GE Healthcare) and then using the immunogen (AiV 3D^{pol})-coupled agarose beads AminoLink Coupling Resin (Thermo Fisher Scientific) according to the manufacturer's instructions.

Tissue cultures and transfections – Human HeLa and HEK293T were maintained in Dulbecco's modified Eagle's medium (Sigma) supplemented with 10% fetal calf serum (Gibco). RNA transfections of HeLa cells were carried out using the TransIT mRNA transfection kit (Mirus Bio). DNA transfections of HEK293T cells were performed using polyethylenimine (Sigma).

Viral subgenomic replicon assay – For viral subgenomic replicon assays, the pUCSF-AiV/mCherry plasmid was generated by replacing the capsid region within the pUCSF-AiV plasmid (kind gift of Prof. Joseph DeRisi, UCSF School of Medicine, San Francisco, USA) with the mCherry-encoding gene by restriction-free cloning. Further mutations within the 3D^{pol} encoding region were generated using the Q5 Site-Directed Mutagenesis Kit (New England BioLabs). All DNA constructs were verified by sequencing.

The pUCSF-AiV/mCherry wild-type and mutant plasmids were linearized by cleavage with HindIII-HF (Thermo Fisher Scientific) and purified using the mini spin columns (Epoch Life Science). Viral subgenomic replicon RNA was generated with TranscriptAid T7 High Yield Transcription Kit (Thermo Fisher Scientific) and purified using the RNeasy mini spin columns (Qiagen). For replicon assays, HeLa cells grown in 12-well plates were transfected with T7-amplified RNA using the TransIT mRNA transfection kit (Mirus Bio). At 9 h post transfection, the reporter mCherry fluorescence was determined by flow cytometry using BD LSR Fortessa (BD Biosciences) and the following optical configuration: 561 nm laser, 670/30 nm bandpass filter. Acquired data

were analyzed with the FlowJo software.

Western blotting and immunodetection – HEK293T cells grown in 6-well plates were transfected with the appropriate mutants of the EGFP-fused AiV 3CD protein using polyethylenimine (Sigma). Next day, the cells were harvested, washed twice with phosphate-buffered saline and lysed in the ice-cold lysis buffer (20 mM Tris pH 7.4, 100 mM NaCl, 50 mM NaF, 10 mM EDTA, 10% glycerol, 1% NP-40), supplemented with protease inhibitors (Complete protease inhibitor cocktail, Sigma). After solubilizing for 15 min on ice, the lysates were pre-cleared by centrifugation at 16,000g for 15 min. The resulting supernatants were mixed with the Laemmli sample buffer, boiled for 5 min, subjected to SDS-PAGE using the 15% polyacrylamide gels, Western-blotted onto nitrocellulose membranes (Thermo Fisher Scientific), and stained with the following antibodies: rabbit polyclonal antibody to GFP (Cell Signaling #2956), rabbit polyclonal antibody to AiV $3D^{pol}$ (home-made), and mouse monoclonal antibody to β-actin (Sigma #A5441). The images were acquired using the LI-COR Odyssey Infrared Imaging System.

Crystallization – Crystals grew in sitting drops in 1–2 days at 18 °C. Drops were set up from 50% of protein solution and 50% of precipitant. 12,5% PEG 1000, 12,5% PEG 3350, 12,5% (RS)-2-methyl-2,4-pentanediol, 0,2 M 0.2 M 1,6-hexanediol, 0.2 M 1-butanol, 0.2 M (RS)-1,2-propanediol, 0.2 M 2-propanol, 0.2 M 1,4-butanediol, 0.2 M 1,3-propanediol and 0,1 M MES/Imidazole pH 6.5 was used for the SiV $3D^{pol}$ and 1.7 M $(NH_4)_2SO_4$, 1.7% (v/v) PEG400, 15% (v/v) glycerol, 85 mM HEPES pH 7.5 for the AiV $3D^{pol}$.

Crystals were flash frozen in liquid nitrogen and diffraction data were collected at the 14.1 BESSY beamline in HZB synchrotron facility (Mueller et al., 2012).

Data analysis – Diffraction data were processed using XDS (Kabsch, 2010), Phenix (Adams et al., 2010) and CCP4 (Winn et al., 2011) packages. Molecular replacement was done with a polyalanine model based on the structure of EMCV $3D^{pol}$ (PDB ID: 4nz0) (Vives-Adrian et al., 2014). Structural figures were generated by PyMol (Schrodinger). The electrostatic surface was calculated using plugin APBS Tools (Baker et al., 2001). The structural alignment was done in Multiseq tool in VMD, STAMP (Roberts et al., 2006) and the phylogenetic tree was visualized in FigTree interaface (http://tree.bio.ed.ac.uk/software/figtree/).

5. Accession numbers

The crystal structures of the SiV and AiV $3D^{pol}$ enzyme were deposited in the PDB database under codes 6QWT and 6R1I.

Author contributions

AD, VH, and MK performed all experiments, analyzed data and prepared figures. EB designed the study, supervised the project and wrote the manuscript.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

We are grateful to Prof. Joseph DeRisi (UCSF School of Medicine, San Francisco, USA) for providing the pUCSF-AiV plasmid encoding the synthetically constructed complete aichivirus genome. We thank HZB for the allocation of synchrotron radiation beamtime. The work was supported from European Regional Development Fund; OP RDE; Project: "Chemical biology for drugging undruggable targets (ChemBioDrug)" (No. CZ.02.1.01/0.0/0.0/16_019/0000729) and by the Czech Science Foundation (grant number 17-07058Y). The Academy of Sciences of Czech Republic support (RVO:61388963) is also acknowledged.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jsb.2019.08.004.

References

Adams, P.D., Afonine, P.V., Bunkoczi, G., Chen, V.B., Davis, I.W., Echols, N., Headd, J.J., Hung, L.W., Kapral, G.J., Grosse-Kunstleve, R.W., McCoy, A.J., Moriarty, N.W., Oeffner, R., Read, R.J., Richardson, D.C., Richardson, J.S., Terwilliger, T.C., Zwart, P.H., 2010. PHENIX: a comprehensive Python-based system for macromolecular structure solution. Acta Crystallogr. Sect. D, Biol. Crystallogr. 66, 213–221.

Altan-Bonnet, N., 2017. Lipid tales of viral replication and transmission. Trends Cell Biol. 27, 201–213.

Andino, R., Domingo, E., 2015. Viral quasispecies. Virology 479–480, 46–51.

Baker, N.A., Sept, D., Joseph, S., Holst, M.J., McCammon, J.A., 2001. Electrostatics of nanosystems: application to microtubules and the ribosome. Proc. Natl. Acad. Sci. U.S.A. 98, 10037–10041.

Bullman, S., Kearney, K., O'Mahony, M., Kelly, L., Whyte, P., Fanning, S., Morgan, J.G., 2014. Identification and genetic characterization of a novel picornavirus from chickens. J. Gen. Virol. 95, 1094–1103.

Campagnola, G., Weygandt, M., Scoggin, K., Peersen, O., 2008. Crystal structure of coxsackievirus B3 3Dpol highlights the functional importance of residue 5 in picornavirus polymerases. J. Virol. 82, 9458–9464.

Dubankova, A., Boura, E., 2019. Structure of the yellow fever NS5 protein reveals conserved drug targets shared among flaviviruses. Antiviral Res. 169, 104536.

Dubankova, A., Humpolickova, J., Klima, M., Boura, E., 2017. Negative charge and membrane-tethered viral 3B cooperate to recruit viral RNA dependent RNA polymerase 3D (pol). Sci. Rep. 7, 17309.

Ferrer-Orta, C., Arias, A., Escarmis, C., Verdaguer, N., 2006. A comparison of viral RNA-dependent RNA polymerases. Curr. Opin. Struct. Biol. 16, 27–34.

Ferrer-Orta, C., Arias, A., Perez-Luque, R., Escarmis, C., Domingo, E., Verdaguer, N., 2004. Structure of foot-and-mouth disease virus RNA-dependent RNA polymerase and its complex with a template-primer RNA. J. Biol. Chem. 279, 47212–47221.

Ferrer-Orta, C., Arias, A., Perez-Luque, R., Escarmis, C., Domingo, E., Verdaguer, N., 2007. Sequential structures provide insights into the fidelity of RNA replication. Proc. Natl. Acad. Sci. U.S.A. 104, 9463–9468.

Gohara, D.W., Ha, C.S., Kumar, S., Ghosh, B., Arnold, J.J., Wisniewski, T.J., Cameron, C.E., 1999. Production of "authentic" poliovirus RNA-dependent RNA polymerase (3D(pol)) by ubiquitin-protease-mediated cleavage in Escherichia coli. Protein Expr. Purif. 17, 128–138.

Goldschmidt, L., Cooper, D.R., Derewenda, Z.S., Eisenberg, D., 2007. Toward rational protein crystallization: a Web server for the design of crystallizable protein variants. Protein Sci.: A Publ. Protein Soc. 16, 1569–1576.

Gong, P., Peersen, O.B., 2010. Structural basis for active site closure by the poliovirus RNA-dependent RNA polymerase. Proc. Natl. Acad. Sci. U.S.A. 107, 22505–22510.

Gong, P., Kortus, M.G., Nix, J.C., Davis, R.E., Peersen, O.B., 2013. Structures of coxsackievirus, rhinovirus, and poliovirus polymerase elongation complexes solved by engineering RNA mediated crystal contacts. PLoS ONE 8, e60272.

Greninger, A.L., Knudsen, G.M., Betegon, M., Burlingame, A.L., Derisi, J.L., 2012. The 3A protein from multiple picornaviruses utilizes the golgi adaptor protein ACBD3 to recruit PI4KIIIbeta. J. Virol. 86, 3605–3616.

Hansen, J.L., Long, A.M., Schultz, S.C., 1997. Structure of the RNA-dependent RNA polymerase of poliovirus. Structure 5, 1109–1122.

Harak, C., Lohmann, V., 2015. Ultrastructure of the replication sites of positive-strand RNA viruses. Virology 479, 418–433.

Harris, K.S., Reddigari, S.R., Nicklin, M.J., Hammerle, T., Wimmer, E., 1992. Purification and characterization of poliovirus polypeptide 3CD, a proteinase and a precursor for RNA polymerase. J. Virol. 66, 7481–7489.

Horova, V., Lyoo, H., Rozycki, B., Chalupska, D., Smola, M., Humpolickova, J., Strating, J., van Kuppeveld, F.J.M., Boura, E., Klima, M., 2019. Convergent evolution in the mechanisms of ACBD3 recruitment to picornavirus replication sites. PLoS Pathog. 15, e1007962.

Hsu, N.Y., Ilnytska, O., Belov, G., Santiana, M., Chen, Y.H., Takvorian, P.M., Pau, C., van der Schaar, H., Kaushik-Basu, N., Balla, T., Cameron, C.E., Ehrenfeld, E., van Kuppeveld, F.J., Altan-Bonnet, N., 2010. Viral reorganization of the secretory pathway generates distinct organelles for RNA replication. Cell 141, 799–811.

Kabsch, W., 2010. Xds. Acta Crystallogr. Sect. D, Biol. Crystallogr. 66, 125–132.

Kitajima, M., Gerba, C.P., 2015. Aichi virus 1: environmental occurrence and behavior. Pathogens 4, 256–268.

Klima, M., Chalupska, D., Rozycki, B., Humpolickova, J., Rezabkova, L., Silhan, J., Baumlova, A., Dubankova, A., Boura, E., 2017. Kobuviral non-structural 3A proteins act as molecular harnesses to hijack the host ACBD3 protein. Structure 25, 219–230.

Klima, M., Toth, D.J., Hexnerova, R., Baumlova, A., Chalupska, D., Tykvart, J., Rezabkova, L., Sengupta, N., Man, P., Dubankova, A., Humpolickova, J., Nencka, R., Veverka, V., Balla, T., Boura, E., 2016. Structural insights and in vitro reconstitution of membrane targeting and activation of human PI4KB by the ACBD3 protein. Sci. Rep. 6, 23641.

Love, R.A., Maegley, K.A., Yu, X., Ferre, R.A., Lingardo, L.K., Diehl, W., Parge, H.E., Dragovich, P.S., Fuhrman, S.A., 2004. The crystal structure of the RNA-dependent RNA polymerase from human rhinovirus: a dual function target for common cold antiviral therapy. Structure 12, 1533–1544.

Marcotte, L.L., Wass, A.B., Gohara, D.W., Pathak, H.B., Arnold, J.J., Filman, D.J., Cameron, C.E., Hogle, J.M., 2007. Crystal structure of poliovirus 3CD protein: virally encoded protease and precursor to the RNA-dependent RNA polymerase. J. Virol. 81, 3583–3596.

McPhail, J.A., Ottosen, E.H., Jenkins, M.L., Burke, J.E., 2017. The molecular basis of Aichi virus 3A protein activation of phosphatidylinositol 4 kinase IIIbeta, PI4KB, through ACBD3. Structure 25, 121–131.

Mejdrova, I., Chalupska, D., Kogler, M., Sala, M., Plackova, P., Baumlova, A., Hrebabecky, H., Prochazkova, E., Dejmek, M., Guillon, R., Strunin, D., Weber, J., Lee, G., Birkus, G., Mertlikova-Kaiserova, H., Boura, E., Nencka, R., 2015. Highly selective phosphatidylinositol 4-kinase IIIbeta inhibitors and structural insight into their mode of action. J. Med. Chem. 58, 3767–3793.

Mejdrova, I., Chalupska, D., Plackova, P., Muller, C., Sala, M., Klima, M., Baumlova, A., Hrebabecky, H., Prochazkova, E., Dejmek, M., Strunin, D., Weber, J., Lee, G., Matousova, M., Mertlikova-Kaiserova, H., Ziebuhr, J., Birkus, G., Boura, E., Nencka, R., 2017. Rational design of novel highly potent and selective phosphatidylinositol 4-kinase IIIbeta (PI4KB) inhibitors as broad-spectrum antiviral agents and tools for chemical biology. J. Med. Chem. 60, 100–118.

Mueller, U., Darowski, N., Fuchs, M.R., Forster, R., Hellmig, M., Paithankar, K.S., Puhringer, S., Steffien, M., Zocher, G., Weiss, M.S., 2012. Facilities for macromolecular crystallography at the Helmholtz-Zentrum Berlin. J. Synchrotron Radiat. 19, 442–449.

Nagy, P.D., Strating, J.R., van Kuppeveld, F.J., 2016. Building viral replication organelles: close encounters of the membrane types. PLoS Pathog. 12, e1005912.

Paul, A.V., van Boom, J.H., Filippov, D., Wimmer, E., 1998. Protein-primed RNA synthesis by purified poliovirus RNA polymerase. Nature 393, 280–284.

Peersen, O.B., 2017. Picornaviral polymerase structure, function, and fidelity modulation. Virus Res.

Robert, X., Gouet, P., 2014. Deciphering key features in protein structures with the new ENDscript server. Nucleic Acids Res. 42, W320–W324.

Roberts, E., Eargle, J., Wright, D., Luthey-Schulten, Z., 2006. MultiSeq: unifying sequence and structure data for evolutionary analysis. BMC Bioinf. 7, 382.

Sabin, C., Fuzik, T., Skubnik, K., Palkova, L., Lindberg, A.M., Plevka, P., 2016. Structure of Aichi virus 1 and its empty particle: clues towards kobuvirus genome release mechanism. J. Virol.

Sasaki, J., Ishikawa, K., Arita, M., Taniguchi, K., 2012. ACBD3-mediated recruitment of PI4KB to picornavirus RNA replication sites. EMBO J. 31, 754–766.

Sebera, J., Dubankova, A., Sychrovsky, V., Ruzek, D., Boura, E., Nencka, R., 2018. The structural model of Zika virus RNA-dependent RNA polymerase in complex with RNA for rational design of novel nucleotide inhibitors. Sci. Rep. 8, 11132.

Sun, D., Chen, S., Cheng, A., Wang, M., 2016. Roles of the picornaviral 3C proteinase in the viral life cycle and host cells. Viruses 8, 82.

Thompson, A.A., Peersen, O.B., 2004. Structural basis for proteolysis-dependent activation of the poliovirus RNA-dependent RNA polymerase. EMBO J. 23, 3462–3471.

Vives-Adrian, L., Lujan, C., Oliva, B., van der Linden, L., Selisko, B., Coutard, B., Canard, B., van Kuppeveld, F.J., Ferrer-Orta, C., Verdaguer, N., 2014. The crystal structure of a cardiovirus RNA-dependent RNA polymerase reveals an unusual conformation of the polymerase active site. J. Virol. 88, 5595–5607.

Winn, M.D., Ballard, C.C., Cowtan, K.D., Dodson, E.J., Emsley, P., Evans, P.R., Keegan, R.M., Krissinel, E.B., Leslie, A.G., McCoy, A., McNicholas, S.J., Murshudov, G.N., Pannu, N.S., Potterton, E.A., Powell, H.R., Read, R.J., Vagin, A., Wilson, K.S., 2011. Overview of the CCP4 suite and current developments. Acta Crystallogr. Sect. D, Biol. Crystallogr. 67, 235–242.

Wu, Y., Lou, Z., Miao, Y., Yu, Y., Dong, H., Peng, W., Bartlam, M., Li, X., Rao, Z., 2010. Structures of EV71 RNA-dependent RNA polymerase in complex with substrate and analogue provide a drug target against the hand-foot-and-mouth disease pandemic in China. Protein Cell 1, 491–500.

Yamashita, T., Kobayashi, S., Sakae, K., Nakata, S., Chiba, S., Ishihara, Y., Isomura, S., 1991. Isolation of cytopathic small round viruses with BS-C-1 cells from patients with gastroenteritis. J. Infect. Dis. 164, 954–957.

Antiviral Research 169 (2019) 104536

Contents lists available at ScienceDirect

Antiviral Research

journal homepage: www.elsevier.com/locate/antiviral

Structure of the yellow fever NS5 protein reveals conserved drug targets shared among flaviviruses

Anna Dubankova, Evzen Boura*

Institute of Organic Chemistry and Biochemistry of the Czech Academy of Sciences, Flemingovo Nam. 2, 166 10, Prague 6, Czech Republic

ARTICLE INFO

Keywords:
RNA
Polymerase
Crystal structure
Flavivirus
Yellow fever
Drug design

ABSTRACT

Yellow fever virus (YFV) is responsible for devastating outbreaks of Yellow fever (YF) in humans and is associated with high mortality rates. Recent large epidemics and epizootics and exponential increases in the numbers of YF cases in humans and non-human primates highlight the increasing threat YFV poses, despite the availability of an effective YFV vaccine. YFV is the first human virus discovered, but the structures of several of the viral proteins remain poorly understood. Here we report the structure of the full-length NS5 protein, a key enzyme for the replication of flaviviruses that contains both a methyltransferase domain and an RNA dependent RNA polymerase domain, at 3.1 Å resolution. The viral polymerase adopts right-hand fold, demonstrating the similarities of the Yellow fever, Dengue and Zika polymerases. Together this data suggests NS5 as a prime and ideal target for the design of pan-flavivirus inhibitors.

1. Introduction

Yellow fever virus (YFV) is a mosquito-borne virus (Epelboin et al., 2017), along with several other flaviviruses such as the Zika virus (ZIKV) and Dengue virus. The population of the mosquito *Aedes aegypti*, which is the major mosquito species responsible for transmission of these viruses, is growing due to global climatic changes (Monaghan et al., 2018). As a result, flaviviruses are becoming a global threat to mankind as documented by the recent Zika (Abushouk et al., 2016) and Yellow Fever (YF) outbreaks (Blake and Garcia-Blanco, 2014; Couto-Lima et al., 2017). The outbreaks are moreover associated with high mortality (YF) or cause microcephaly in infants and Guillain-Barré syndrome (ZIKV) in some patients (Krauer et al., 2017). Unfortunately, despite the availability of an effective YFV vaccine, large epidemics and epizootics of YF were recently reported in Brazil and Angola with exponential increases in the numbers of YF cases in humans and non-human primates (Barrett, 2018). These outbreaks indicate that both an effective antiviral vaccine and antiviral cure are important.

YFV and its transmission by mosquitoes was discovered in 1900 (Reed et al., 1900) and significant progress in understanding this virus has been made since then. For instance, the structure of the YFV capsid was solved by cryo-electron microscopy (Zhang et al., 2003) and the crystal structure of NS3 helicase is also available (Wu et al., 2005). Despite these advances, the structures of several of the viral proteins remain poorly understood (Douam and Ploss, 2018). Of these, the NS5 protein is a key enzyme for the replication of flaviviruses because it contains an N-terminal methyltransferase (MTase) domain and a C-terminal RNA dependent RNA polymerase (RdRp) domain (Fig. 1A) (Selisko et al., 2018). The MTase domain has two enzymatic activities: it is responsible for guanine-N7 and nucleoside-2′-O methylation (Coutard et al., 2017; Dong et al., 2010; Hercik et al., 2017a), both of which increase the stability of newly synthesized RNA by preventing degradation by 5′–3′ exoribonucleases. RNA methylation also helps the virus to avoid the innate immunity (Hyde and Diamond, 2015) as non-methylated RNA is recognized by the innate immunity factor Ifit1 (IFN-induced RNA-binding protein) and its translation inhibited. The RdRp domain is essential for replication of viral RNA because its enzymatic activity cannot be replaced by any host enzyme. Therefore, both domains are indispensable for flaviviral replication making the NS5 enzyme a prime target for drug discovery (Lim et al., 2015).

For some flaviviruses including Zika and Dengue the structure of the NS5 protein was reported (Duan et al., 2017; Godoy et al., 2017; Upadhyay et al., 2017; Zhao et al., 2015, 2017). In the case of Yellow fever, however, only the structure of the methyltransferase domain was described (Geiss et al., 2009). Here, we describe the crystal structure of the full-length Yellow fever NS5 enzyme at 3.1 Å resolution. It reveals the right-hand fold typical for viral polymerases and highlights conserved regions between Yellow fever, Dengue and Zika, thus providing critical information for structure-aided drug design.

* Corresponding author.
E-mail address: boura@uochb.cas.cz (E. Boura).

https://doi.org/10.1016/j.antiviral.2019.104536
Received 16 April 2019; Received in revised form 7 June 2019; Accepted 12 June 2019
Available online 13 June 2019

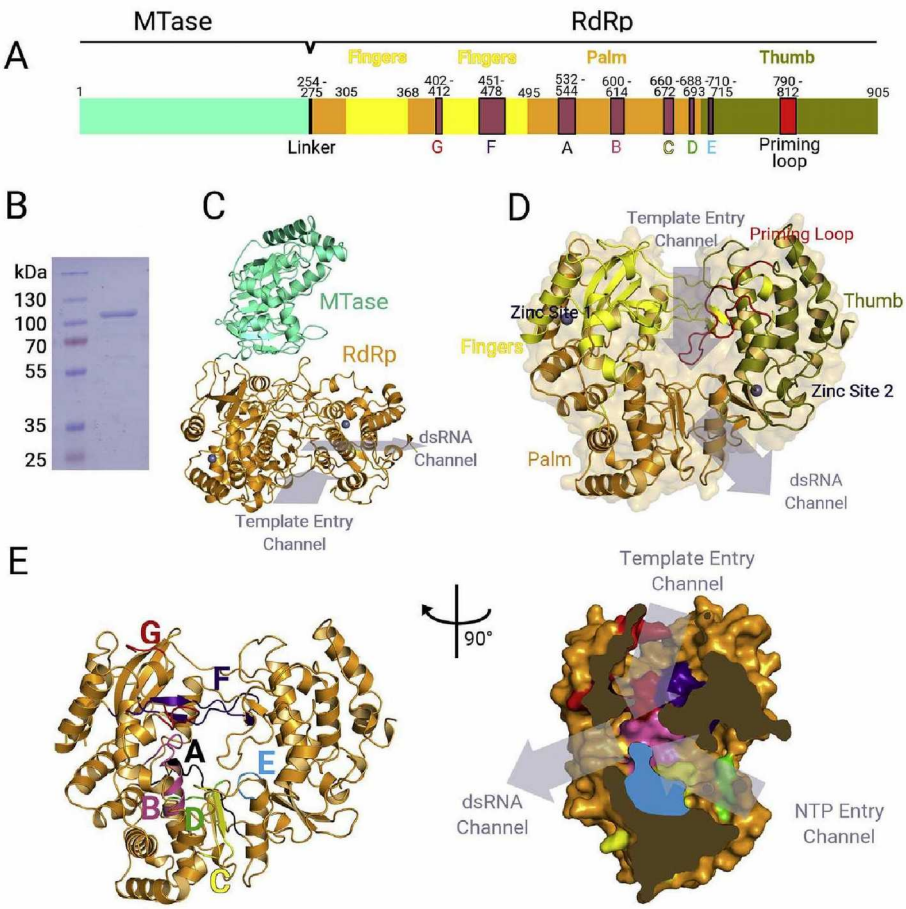


Fig. 1. Structure of the full-length YF NS5 protein.
A. YFV NS5 is composed of the MTase domain and the RdRp. Three subdomains and seven motifs (A-G) occur in the RdRp domain.
B. Recombinant YFV NS5 FL protein was used for all experiments (99% pure based on SDS PAGE analysis).
C. The structure of the YF FL NS5 protein in ribbon representation. The MTase domain is shown in cyan and the RdRp domain is shown in orange.
D. Structure of the RdRp domain. The template entry, dsRNA and NTP entry channels as well as the fingers, palm and thumb subdomains are highlighted.
E. Positions of the seven motifs of the RdRp domain. The template entry channel contains the F and G motifs, while the NTP entry channel contains the B, C, D and E motifs.

2. Results

We aimed to obtain the crystal structure of the Yellow fever polymerase, as understanding the structural conservation and differences across flaviviruses will help in rational drug design. We expressed the full-length NS5 protein from the Yellow fever strain 17D vaccine in *E. coli* and optimized the purification protocol to obtain highly pure protein (Fig. 1B). After crystal optimization, we obtained crystals that diffracted to ~3 Å and were able to collect a complete dataset at 3.1 Å resolution. The crystals contained two molecules of the NS5 enzyme in the asymmetric unit and we solved the structure by molecular replacement using the methyltransferase domain and a homology model of the RdRp domain (described in detail in M&M section). This made it possible to trace the entire polypeptide chain except for the very first five residues at the N-terminus, the very last eleven residues at the C-terminus, and a disordered loop between Ile410 and Glu418. The two molecules in the asymmetric unit are virtually identical and superimpose with an RMSD of 0.29 Å (SI Fig. 1). The final model was refined to R_{work} = 22.07% and R_{free} = 27.00% (Table 1).

The overall structure can be divided into an RdRp domain and a methyltransferase domain that is stacked on the fingers and F motif of the RdRp domain (Fig. 1C). The MTase domain consists of nine helices and nine β-strands that forms a classic α/β/α sandwich structure (Byszewska et al., 2014), which form the GTP and SAM binding sites. We clearly observed the electron density of a ligand in the SAM pocket (SI Fig. 2) and we modeled the SAH (S-adenosylhomocysteine) into the map. However, the observed electron density probably corresponds to a mixture of SAH and SAM (S-adenosylmethionine) since both these ligands are present in *E. coli*. and they co-purify with the recombinant protein. Because a high resolution structure of YF MTase domain is already available (Geiss et al., 2009) and we didn't observe any discrepancies (RMSD of 0.94 Å, SI Fig. 3) we will not describe the structure of the MTase domain in detail here.

The RdRp domain has a right-hand fold as expected (Fig. 1) with three channels (template entry, dsRNA and NTP entry channel) and can be divided into fingers, palm and thumb subdomains (Fig. 1C). The fingers subdomain is composed of two regions, the first of which is

Table 1
Statistics of crystallographic data collection and refinement.

Data collection	
Space group	$P2_12_12_1$
Cell dimension (Å)	a = 108.3 b = 142.1 , c = 152.2
X-ray source	BESSY ID 14-2
Resolution (Å)	47.8 - 3.0 (3.11 - 3.00)
Unique reflections	47,148 (4576)
I/σ (I)	9.91 (1.13)[a]
Wavelength (Å)	0.9184
Multiplicity	8.3 (8.5)
Completeness (%)	99.19 (97.65)
R-pim, %	8.1 (68.7)
$CC_{1/2}$	0.994 (0.419)
CC*	0.999 (0.769)
Refinement	
R-work, %	22.07 (32.22)
R-free, %	27.00 (33.90)
RMS(bonds), °	0.002
RMS(angles), Å	0.51
Ramachandran (outliers/favored) (%)	0.0%/97.8%

Numbers in parentheses refer to the highest resolution shell.
[a] I/σ (I) = 2 at 3.1 Å.

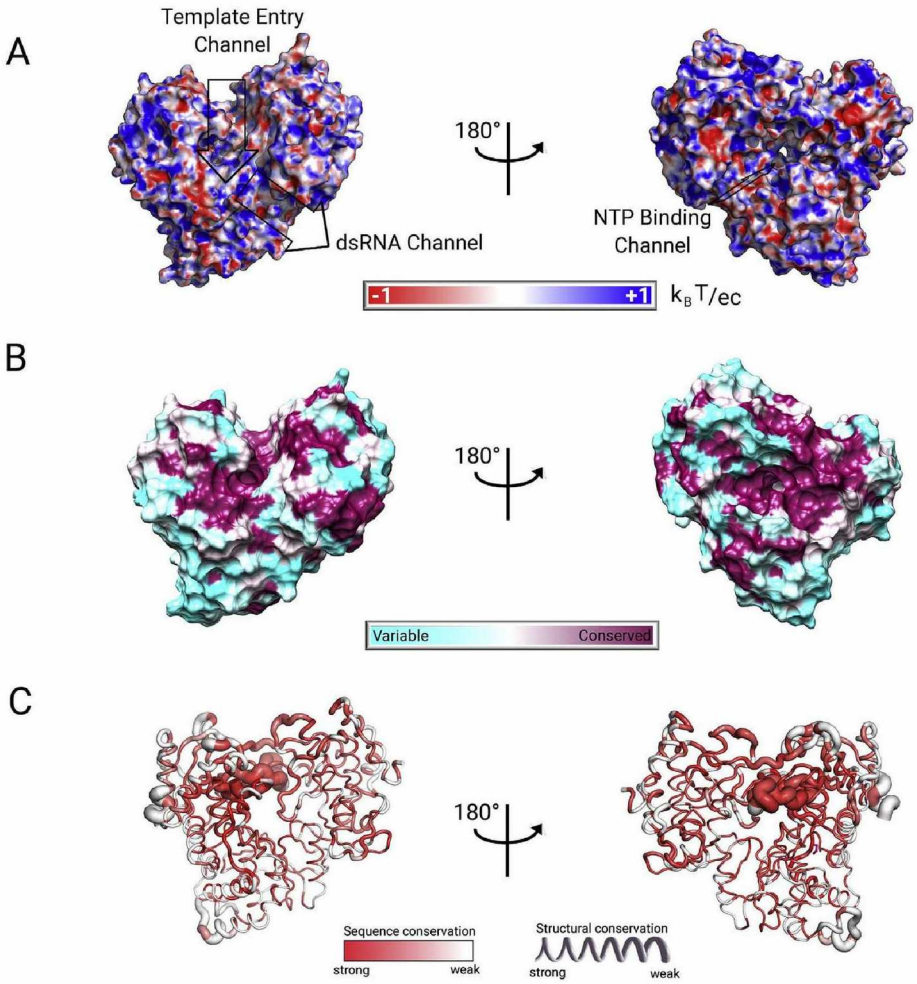


Fig. 2. Surface and structural analysis of YF RdRp.
A. The surface of the template entry channel is positively charged and, optimized to bind the RNA substrate.
B. Surface representation of the YF RdRp domain colored according to sequence conservation (cyan: not conserved, dark magenta: conserved). This shows the highly conserved protein core and variable periphery.
C. Sausage representation based on a combination of sequence and structural alignments of Flaviviral NS5 proteins reveals structural differences in regions in which the sequence is conserved. Structures used for this alignment are: WNV (2hcn, 2hfz), DENV(2j7u, 3vws, 4v0q, 4v0r, 5ccv, 5dto, 5iq6, 5k5m), JEV (4hdg, 4k6m, 4mtp), ZIKV (5m2x, 5tfr, 5tmh, 5u0b, 5u0c, 5u04, 5wz3).

located in helices α11 - 14 and sheets β10 - 12 (residue 305–368) and the second of which is located in helices α16 - 20 and sheets β13 - 14 (residues 402–495). In contrast, the palm subdomain is spread throughout the RdRp. It consists of three regions, helix α10 (residues 275–305), helices α15 – 16 (residues 368–402) and helices α21 – 31 together with beta sheets β16 – 20 (residues 495–720) and forms the core of the RdRp. The thumb subdomain is composed of the C terminus of RdRp (helices α32 – 40 and beta sheets β21 - 22, residues 722–905) and together with the palm subdomain form the dsRNA channel.

The seven catalytic motifs (A - G), that are present in every viral RdRp (Lu and Gong, 2017), could also be detected in the YF RdRp domain. Motif A is located in helix α23 and contains the catalytically important aspartate residues Asp535 and Asp541. Motif B contains highly conserved Ser603 and Gly604 residues. The sidechain of Ser603 interacts with the 2′-hydroxyl group of the ribose ring during active side closure (Appleby et al., 2015) while motifs C and E, located in β18 and 19 respectively in a loop between helix α31 and sheet β20, interact with the backbone of the RNA product. Motif D that is localized adjacent to helix α30 is important for the active site closure by the conserved lysine residue Lys359 (Yang et al., 2012). Motif F consists of three beta sheets (β13 – 15) and connecting loops, it is located at the top of the NTP entry channel and its primary function is to bind the NTP nucleotide. Importantly, it contains a conserved phenylalanine residue Phe466 at the interface of the RdRp and MTase domains. Motif G contacts the backbone of the template RNA in other non-flaviviral RdRp enzymes (Appleby et al., 2015; Gong et al., 2013; Gong and Peersen, 2010) is located in a loop located between α16 and 17 (Fig. 1A).

The template entry channel is surrounded by the fingers and thumb subdomains with the priming loop above. It also contains motifs G and F, which interact with the RNA template. It is followed by the dsRNA channel which forms a ~45° angle to the template entry channel. The dsRNA channel is formed mainly by the palm subdomain and motifs C and E. The NTP channel is localized between the template and dsRNA channels and contains the A, B and D motifs (Fig. 1E). Surface electrostatic potential analysis reveals that the channels are positively charged (Fig. 2) which is essential for RNA binding. The priming loop (Fig. 1D in red) is localized above the interface of the three channels. Its purpose is to properly position nucleotides for RNA polymerization (Ng et al., 2008). Based on structural analysis of the HCV polymerase it was proposed that a stacking interaction occurs between a residue in the priming loop (tyrosine in the case of HCV) and the initiating nucleotide (Butcher et al., 2001). In Zika this role was proposed to be played by the highly conserved tryptophan residue Trp797 (Zhao et al., 2017) which corresponds to Trp799 in YF. In a computational (QM/MM devised) model of Zika NS5 with RNA, the Trp797 is positioned above the base suggesting a T-shaped stacking interaction (Sebera et al., 2018).

Two zinc binding sites occur in the fingers and thumb subdomains of the RdRp (Fig. 3). These are also present in every flaviviral RdRp structure solved so far. The first zinc atom is coordinated by Cys448, Cys451, His443 and Glu439 and the second zinc binding site is formed by Cys732, Cys851, His716 and His718. A glutamate residue in a zinc binding site is somewhat unusual, however, this glutamate and other residues forming the first zinc binding site are absolutely conserved among NS5 proteins from YFV, ZIKV, DEN, and WNV (Fig. 3C). The second zinc binding site is partially conserved: Cys732, Cys851, His716 are conserved, while His718 is replaced by an asparagine residue in Zika and by a threonine residue in WNV (Fig. 3C). As a consequence, the zinc ion is coordinated by water molecule instead of a histidine at

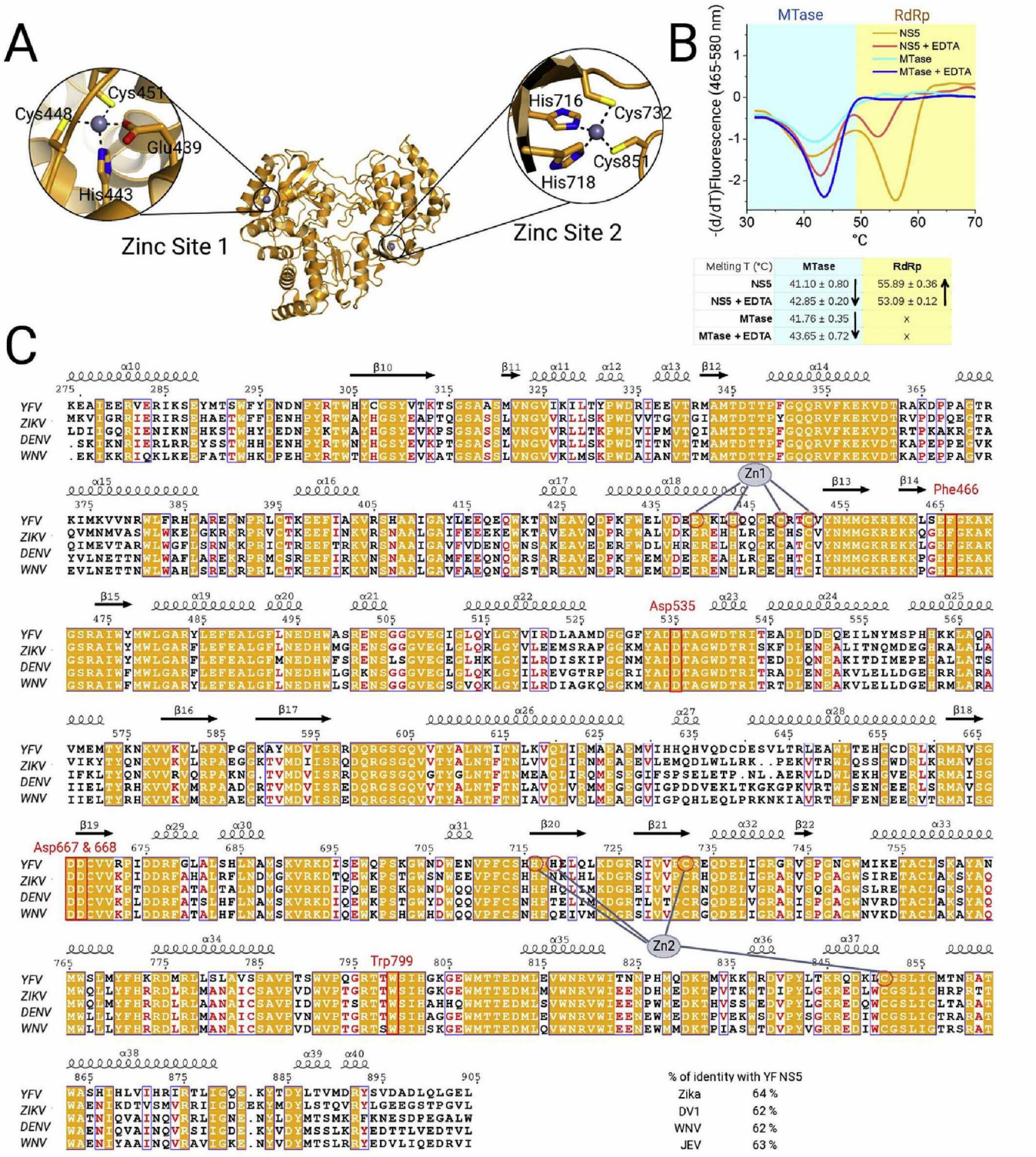


Fig. 3. Localization of the zinc binding sites in the RdRp domain.

A. Zoomed view of the two YF RdRp zinc binding sites. The first zinc atom is coordinated by Glu439, His443, Cys448 and Cys451, while the second one by His716, His718, Cys732 and Cys851.

B. The zinc biding sites stabilize the NS5 protein. Addition of the chelating agent EDTA decreases the thermal stability of the full length NS5 protein but not that of the MTase domain.

C. Sausage representation based on a combination of sequence and structural alignments of Flaviviral NS5 proteins reveals structural differences in regions in which the sequence is conserved. Structures used for this alignment were from WNV (2hcn, 2hfz), DENV(2j7u, 3vws, 4v0q, 4v0r, 5ccv, 5dto, 5iq6, 5k5m), JEV(4hdg, 4k6m, 4mtp), and ZIKV (5m2x, 5tfr, 5tmh, 5u0b, 5u0c, 5u04, 5wz3). In the lower right panel is shown percentage of individual mentioned sequences identity with yellow fever NS5 sequence.

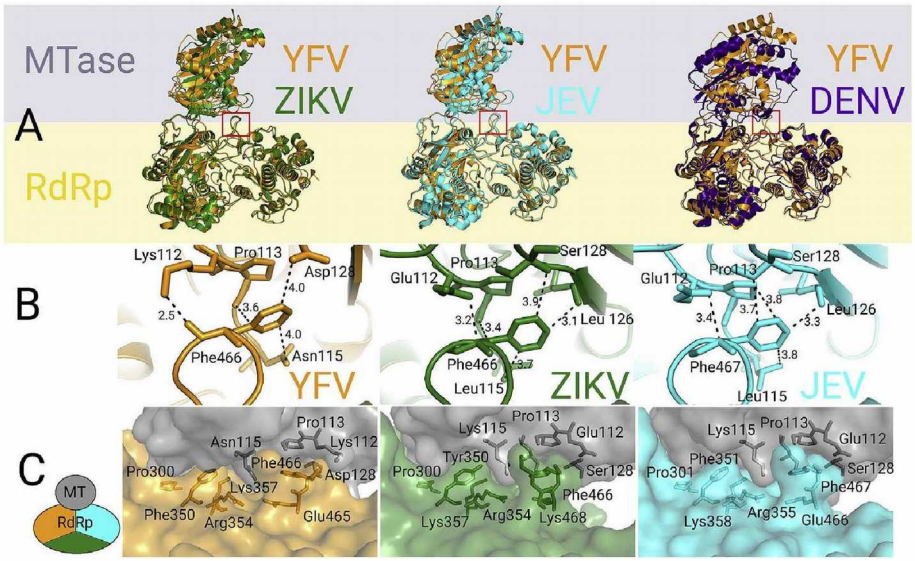


Fig. 4. The MTase:RdRp interface.
A. Structures of full length flaviviral NS5 proteins were superimposed on the structure of YF NS5 based on the RdRp domain. This revealed that the domain orientation is conserved in Yellow fever, Zika and Japanese encephalitis virus RdRp but not in Dengue RdRp.
B. The conserved residues Phe466 and Pro113 define the MTase:RdRp interface.
C. Detail of RdRp MTase interface. MTase is shown in gray color. The color of RdRp is colored depending on viral origin (YFV – yellow, ZIKV – forest green, JEV – cyan).

this position (Duan et al., 2017). The zinc binding sites stabilize the RdRp domain, as the melting temperature (T_m) of the RdRp domain drops from 55.9 °C and to 53.1 °C in the presence of the chelating agent EDTA (Fig. 3B).

For Zika NS5 enzyme it was previously shown that the MTase domain activates the RdRp domain (Zhao et al., 2017). The Zika MTase domain stabilizes motif F in the RdRp domain using a mechanism in which the conserved phenylalanine residue Phe466 stacks against a pocket on the surface of the MTase domain formed by Leu115, Leu126, Pro113 (Fig. 4, and Fig. Sl) (Fig. 4 for the electron density maps). This inter-domain interface is further stabilized by a hydrogen bond between the sidechain of Glu112 and the backbone of motif F (Rusanov et al., 2018). We superimposed the structure of the YF NS5 with other known full-length flaviviral NS5 proteins (Fig. 4A) and this comparison revealed that the MTase:RdRp interface is conserved among YF, Zika (PDB: 5tfr) (Upadhyay et al., 2017) and JEV (Japanese encephalitis virus, PDB:4k6m) (Lu and Gong, 2013). In contrast, the structure of Dengue NS5 (PDB: 4v0q) (Zhao et al., 2015) maintains an F motif that is disordered. Moreover, the residues of the MTase domain adjacent to the conserved residue Pro113, which form the patch which interacts with Phe466, are not conserved, although, the mode of binding is conserved (Fig. 4). In case of YF, Lys112 forms a hydrogen bond with the backbone of Phe466. However, the aromatic ring of Phe466 residue is always stacked against the pyrrolidine ring of Pro113, which is also seen in the structures of the full-length Zika and JEV NS5 enzymes, suggesting that the interaction of Phe466 with Pro113 is the key feature of the MTase:RdRp interface (Fig. 4B and C).

3. Discussion

YF killed hundreds of thousands of people in the 18th and 19th centuries and, according to the World Health Organization, it still causes 84,000–170,000 severe cases and 29,000–60,000 deaths annually despite extensive vaccination. The primary target for antiviral therapy is the NS5 enzyme because both of its domains - the RdRp and MTase - are together crucial for viral replication. As a result, inhibitors of NS5 enzyme are in continual development (Eyer et al., 2016; Hercik et al., 2017b; Stephen et al., 2016). Despite this, the structure of the full-length YF NS5 protein was until now unavailable despite the importance of YF for global human health. Structure-guided inhibitor design is being used with increasing frequency and the lack of a structure of the YF NS5 has made a significant impediment for rational drug design specific to YF. Here we filled this gap by solving the structure of YF NS5 enzyme at 3.1 Å resolution.

This structure reveals the MTase:RdRp interface, which is conserved in the structures of the YF, Zika, and JEV NS5 enzymes but not in the structure of Dengue NS5. The two key residues forming this interface are Pro113 in the MTase domain and Phe446 in the F motif of the RdRp domain. These residues are also conserved in Dengue NS5 enzyme, although the F motif in the structure of the full-length Dengue NS5 enzyme is disordered. The different orientation of the MTase domain in the case of the Dengue enzyme could be a crystal packing artifact, however, that seems unlikely because another structure of the Dengue NS5 enzyme with an inhibitor bound has the same orientation of the MTase and RdRp domain (Lim et al., 2016). It is well established, however, that by stabilizing the F motif, the MTase domain activates the RdRp domain. Interestingly, our structure reveals that the residue Glu111, a neighbor of Lys112 which is involved in F motif stabilization, also forms a hydrogen bond with the sugar hydroxyl group of SAH (Fig. 5). This suggests that not only the MTase domain but also the RdRp domain could influence the enzymatic activity of the MTase domain. Alternatively, it could mean that only ligand bound MTase domain can activate the RdRp domain. Unfortunately, this hypothesis is hard to test because full-length NS5 enzyme without ligand cannot be prepared in bacterial nor eukaryotic expression systems. Another suggested mechanism of an interplay between the MTase and RdRp polymerase is the GTP depended multimerization of the Zika full-length NS5 which could be important for the regulation of the RdRp domain by the MTase domain (Ferrero et al., 2019). Also the role of the membrane remains to be better defined as membranes were suggested to recruit viral polymerases (Dubankova et al., 2017; Hsu et al., 2010) and such a recruitment could initiate the polymerization.

Taken together, this structure revealed several conserved regions among flaviviral polymerases including the catalytic aspartate residues Asp535, Asp667 and Asp668 which coordinate magnesium cations during polymerization reaction, the conserved tryptophan Trp797 within the priming loop and the highly conserved first zinc binding site in the fingers subdomain but not the second zinc binding site (Fig. 6). Our structural analysis suggests that a compound targeting any of these sites would likely be active against multiple flaviviruses.

4. Materials and methods

Protein expression and purification – Artificial codon optimized gene encoding YFV NS5 enzyme (strain 17D vaccine, UniProtKB - P03314) was commercially synthesized (Invitrogen) and cloned into a pSUMO vector using Gibson assembly method. The resulting sequence encoded for a His_{8x}SUMO-NS5 fusion protein. The protein was expressed and

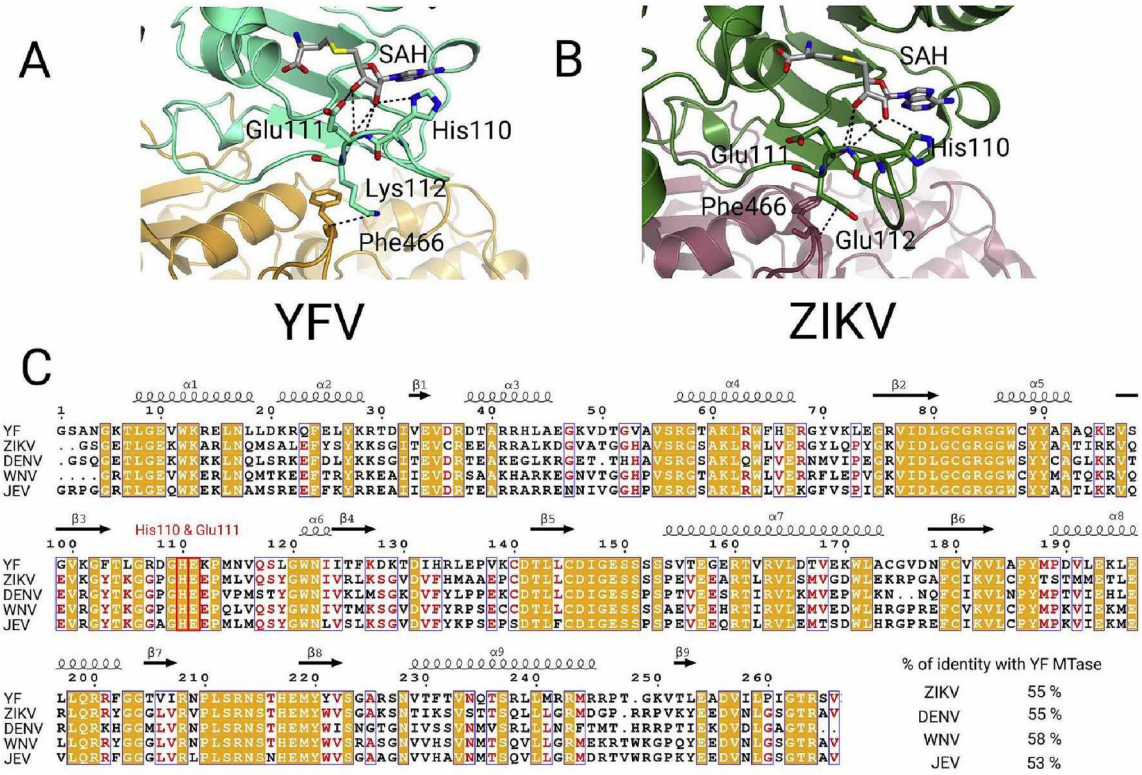


Fig. 5. Loop of the MTase domain contains both, residues coordinating the SAH ligand and residues stabilizing the F motif of the RdRp domain.
A. Detailed view showing the hydrogen bond of MTase localized residue Lys112 to the conserved residue Phe466 of the F motif and the neighboring residues Glu111 and His110 stabilizing the SAH ligand.
B. Residue Glu112 fulfills the same function as YFV Lys112 in other flaviviral polymerases.
C. Sequence alignment illustrating conservation of key residues except of Lys112 that is replaced by Glu112 in other flaviviruses, although, Glu112 fulfills the same structural function.

purified using our standard protocols with some modifications (Baumlova et al., 2014). Briefly, it was expressed in *E. coli* strain BL21 Star in autoinduction ZY media supplemented with 50 μM $ZnSO_4$. The bacteria were lysed using Emulsiflex C3 (Avestin) in the lysis buffer (50 mM Tris pH = 8, 500 mM NaCl, 20 mM imidazole, 3 mM βME). After lysis the protein was immobilized on NiNTA Agarose resin (Machery-Nagel) and extensively washed with the lysis buffer and eluted with the lysis buffer supplemented with 300 mM imidazole. The His_{8x}SUMO solubilisation tag was cleaved by the Ulp1 protease (1.0 mg per 20 mg of YF NS5) at 4 °C overnight. The YF NS5 was further purified by size exclusion chromatography in SEC buffer (20 mM Tris pH 8, 1 M NaCl, 3 mM βME, 10% glycerol). Finally, the protein was concentrated to 6 mg/ml and stored in −80 °C until needed.

Crystallization, data analysis and refinement – Crystals grew in sitting drops consisting of 300 nl of the protein and the well solution (80 mM Tris pH = 8.5, 1.92M $(NH_4)_2SO_4$, 20% (v/v) glycerol) in about a week. Upon harvest were the crystals flash frozen in liquid nitrogen and diffraction data were collected at the 14.2 BESSY beamline in HZB synchrotron facility (Mueller et al., 2012). The crystals diffracted to 3 Å and belonged to the orthorhombic $P2_12_12_1$ space group.

The diffraction data were collected from a single frozen crystal at 100 K and processed using XDS (Kabsch, 2010) within the XDSAPP GUI (Krug et al., 2012). The data were cut at 3 Å where the correlation coefficient $C_{1/2}$ = 0.4. The structure was solved by molecular replacement using two models, the YF methyltransferase domain (pdb code 3EVA) (Geiss et al., 2009) and a homology model of the RdRp domain based on the structure of Japanese encephalitis virus NS5 (pdb code 4K6M) (Lu and Gong, 2013). The structure was further build and refined manually using Phenix (Adams et al., 2010) and Coot (Emsley et al., 2010) to R_{work} = 22.07% and R_{free} = 27.03% and good geometry as summarized in Table 1. Structural figures were generated by PyMol (Schrodinger).

Thermal Shift Assay - The thermal shift assay was run in presence of 1 μM protein, 5000 x diluted SYPRO Orange Protein Gel Stain and 20 mM Tris pH 7.4, 40 mM NaCl, 10 mM βME, 20 mM $MgCl_2$ (and 100 mM EDTA). The mixture was continually heated from 25 to 95 °C in 20 min. The fluorescence signal was measured in a ROCHE LightCycler® 480.

Primary structure alignments were rendered and analyzed by ESPript 3.0 (Robert and Gouet, 2014).

Accession numbers

Atomic coordinates and structure factors have been deposited in the Protein Data Bank with accession codes 6QSN.

Author contributions

AD performed all experiments and wrote the manuscript. EB conceived and supervised the study, analyzed the data, and wrote the manuscript.

Conflicts of interest

The authors declare that they have no conflict of interest.

Funding

The work was supported from European Regional Development Fund; OP RDE; Project: "Chemical biology for drugging undruggable targets (ChemBioDrug)" (No. CZ.02.1.01/0.0/0.0/16_019/0000729) and by the Academy of Sciences Czech Republic (RVO: 61388963).

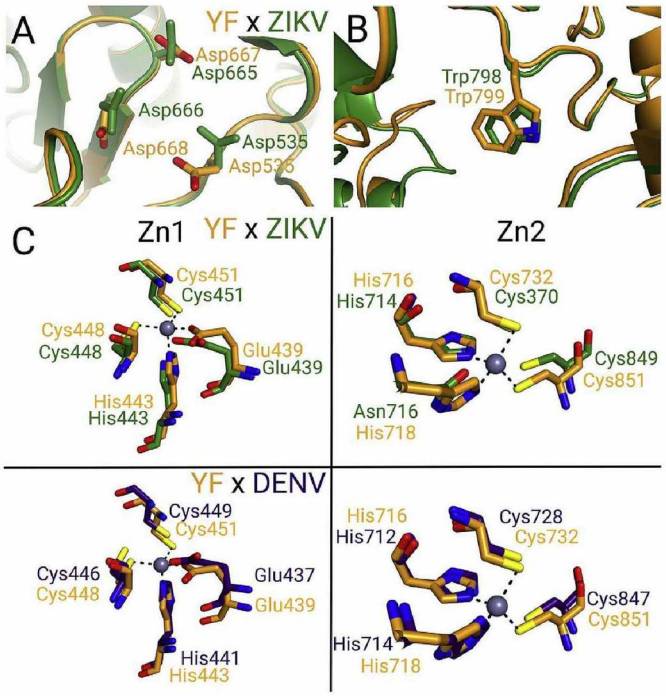


Fig. 6. Superposition of selected motifs of YF, ZIKV or DENV polymerases.
A. YF and ZIKV superposition illustrates the conservation of the Mg^{2+} binding Asp trinity.
B. Residue Trp799 in the priming loop is highly conserved.
C. Superposition of YF and ZIKV or DENV zinc fingers 1 and 2 (Zn1 and Zn2). Zn1 is fully conserved compared to zinc finger in position 2, where His 718 is conserved in YF and DENV, but not in ZIKV RdRp.

Acknowledgements

We thank HZB for the allocation of synchrotron radiation beamtime. We are grateful to Dr. Perta Krafcikova for the gift of the recombinant MTase domain and to David Murray, Tom Leonard and Edward Curtis for critical reading of the manuscript.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.antiviral.2019.104536.

References

Abushouk, A.I., Negida, A., Ahmed, H., 2016. An updated review of Zika virus. J. Clin. Virol.Off. Publ. Pan-Am. Soc. Clin. Virol. 84, 53–58.
Adams, P.D., Afonine, P.V., Bunkoczi, G., Chen, V.B., Davis, I.W., Echols, N., Headd, J.J., Hung, L.W., Kapral, G.J., Grosse-Kunstleve, R.W., McCoy, A.J., Moriarty, N.W., Oeffner, R., Read, R.J., Richardson, D.C., Richardson, J.S., Terwilliger, T.C., Zwart, P.H., 2010. PHENIX: a comprehensive Python-based system for macromolecular structure solution. Acta Crystallogr. Sect. D Biol. Crystallogr. 66, 213–221.
Appleby, T.C., Perry, J.K., Murakami, E., Barauskas, O., Feng, J., Cho, A., Fox third ed, D., Wetmore, D.R., McGrath, M.E., Ray, A.S., Sofia, M.J., Swaminathan, S., Edwards, T.E., 2015. Viral replication. Structural basis for RNA replication by the hepatitis C virus polymerase. Science 347, 771–775.
Barrett, A.D.T., 2018. The reemergence of yellow fever. Science 361, 847–848.
Baumlova, A., Chalupska, D., Rozycki, B., Jovic, M., Wisniewski, E., Klima, M., Dubankova, A., Kloer, D.P., Nencka, R., Balla, T., Boura, E., 2014. The crystal structure of the phosphatidylinositol 4-kinase IIalpha. EMBO Rep. 15, 1085–1092.
Blake, L.E., Garcia-Blanco, M.A., 2014. Human genetic variation and yellow fever mortality during 19th century. U.S. epidemics. mBio 5 e01253-01214.
Butcher, S.J., Grimes, J.M., Makeyev, E.V., Bamford, D.H., Stuart, D.I., 2001. A mechanism for initiating RNA-dependent RNA polymerization. Nature 410, 235–240.
Byszewska, M., Smietanski, M., Purta, E., Bujnicki, J.M., 2014. RNA methyltransferases involved in 5' cap biosynthesis. RNA Biol. 11, 1597–1607.
Coutard, B., Barral, K., Lichiere, J., Selisko, B., Martin, B., Aouadi, W., Lombardia, M.O., Debart, F., Vasseur, J.J., Guillemot, J.C., Canard, B., Decroly, E., 2017. Zika virus methyltransferase: structure and functions for drug design perspectives. J. Virol. 91.
Couto-Lima, D., Madec, Y., Bersot, M.I., Campos, S.S., Motta, M.A., Santos, F.B.D., Vazeille, M., Vasconcelos, P., Lourenco-de-Oliveira, R., Failloux, A.B., 2017. Potential risk of re-emergence of urban transmission of Yellow Fever virus in Brazil facilitated by competent Aedes populations. Sci. Rep. 7, 4848.
Dong, H., Chang, D.C., Xie, X., Toh, Y.X., Chung, K.Y., Zou, G., Lescar, J., Lim, S.P., Shi, P.Y., 2010. Biochemical and genetic characterization of dengue virus methyltransferase. Virology 405, 568–578.
Douam, F., Ploss, A., 2018. Yellow fever virus: knowledge gaps impeding the fight against an old foe. Trends Microbiol. 26, 913–928.
Duan, W., Song, H., Wang, H., Chai, Y., Su, C., Qi, J., Shi, Y., Gao, G.F., 2017. The crystal structure of Zika virus NS5 reveals conserved drug targets. EMBO J. 36, 919–933.
Dubankova, A., Humpolickova, J., Klima, M., Boura, E., 2017. Negative charge and membrane-tethered viral 3B cooperate to recruit viral RNA dependent RNA polymerase 3D (pol). Sci. Rep. 7, 17309.
Emsley, P., Lohkamp, B., Scott, W.G., Cowtan, K., 2010. Features and development of Coot. Acta crystallographica. Acta Crystallogr. Sect. D Biol. Crystallogr. 66, 486–501.
Epelboin, Y., Talaga, S., Epelboin, L., Dusfour, I., 2017. Zika virus: an updated review of competent or naturally infected mosquitoes. PLoS Neglected Trop. Dis. 11, e0005933.
Eyer, L., Nencka, R., Huvarova, I., Palus, M., Joao Alves, M., Gould, E.A., De Clercq, E., Ruzek, D., 2016. Nucleoside inhibitors of Zika virus. J. Infect. Dis. 214, 707–711.
Ferrero, D.S., Ruiz-Arroyo, V.M., Soler, N., Uson, I., Guarne, A., Verdaguer, N., 2019. Supramolecular arrangement of the full-length Zika virus NS5. PLoS Pathog. 15, e1007656.
Geiss, B.J., Thompson, A.A., Andrews, A.J., Sons, R.L., Gari, H.H., Keenan, S.M., Peersen, O.B., 2009. Analysis of flavivirus NS5 methyltransferase cap binding. J. Mol. Biol. 385, 1643–1654.
Godoy, A.S., Lima, G.M., Oliveira, K.I., Torres, N.U., Maluf, F.V., Guido, R.V., Oliva, G., 2017. Crystal structure of Zika virus NS5 RNA-dependent RNA polymerase. Nat. Commun. 8, 14764.
Gong, P., Kortus, M.G., Nix, J.C., Davis, R.E., Peersen, O.B., 2013. Structures of coxsackievirus, rhinovirus, and poliovirus polymerase elongation complexes solved by engineering RNA mediated crystal contacts. PLoS One 8, e60272.
Gong, P., Peersen, O.B., 2010. Structural basis for active site closure by the poliovirus RNA-dependent RNA polymerase. In: Proceedings of the National Academy of Sciences of the United States of America. vol. 107. pp. 22505–22510.
Hercik, K., Brynda, J., Nencka, R., Boura, E., 2017a. Structural basis of Zika virus methyltransferase inhibition by sinefungin. Arch. Virol. 162, 2091–2096.
Hercik, K., Kozak, J., Sala, M., Dejmek, M., Hrebabecky, H., Zbornikova, E., Smola, M., Ruzek, D., Nencka, R., Boura, E., 2017b. Adenosine triphosphate analogs can efficiently inhibit the Zika virus RNA-dependent RNA polymerase. Antivir. Res. 137, 131–133.
Hsu, N.Y., Ilnytska, O., Belov, G., Santiana, M., Chen, Y.H., Takvorian, P.M., Pau, C., van der Schaar, H., Kaushik-Basu, N., Balla, T., Cameron, C.E., Ehrenfeld, E., van Kuppeveld, F.J., Altan-Bonnet, N., 2010. Viral reorganization of the secretory pathway generates distinct organelles for RNA replication. Cell 141, 799–811.
Hyde, J.L., Diamond, M.S., 2015. Innate immune restriction and antagonism of viral RNA lacking 2-O methylation. Virology 479–480, 66–74.
Kabsch, W., 2010. Xds. Acta crystallographica. Acta Crystallogr. Sect. D Biol. Crystallogr. 66, 125–132.
Krauer, F., Riesen, M., Reveiz, L., Oladapo, O.T., Martinez-Vega, R., Porgo, T.V., Haefliger, A., Broutet, N.J., Low, N., Grp, W.Z.C.W., 2017. Zika virus infection as a cause of congenital brain abnormalities and guillain-barre syndrome: systematic review. PLoS Med. 14.
Krug, M., Weiss, M.S., Heinemann, U., Mueller, U., 2012. XDSAPP: a graphical user interface for the convenient processing of diffraction data using XDS. J. Appl. Crystallogr. 45, 568–572.
Lim, S.P., Noble, C.G., Seh, C.C., Soh, T.S., El Sahili, A., Chan, G.K., Lescar, J., Arora, R., Benson, T., Nilar, S., Manjunatha, U., Wan, K.F., Dong, H., Xie, X., Shi, P.Y., Yokokawa, F., 2016. Potent allosteric dengue virus NS5 polymerase inhibitors: mechanism of action and resistance profiling. PLoS Pathog. 12, e1005737.
Lim, S.P., Noble, C.G., Shi, P.Y., 2015. The dengue virus NS5 protein as a target for drug discovery. Antivir. Res. 119, 57–67.
Lu, G., Gong, P., 2013. Crystal Structure of the full-length Japanese encephalitis virus NS5 reveals a conserved methyltransferase-polymerase interface. PLoS Pathog. 9, e1003549.
Lu, G., Gong, P., 2017. A structural view of the RNA-dependent RNA polymerases from the Flavivirus genus. Virus Res. 234, 34–43.
Monaghan, A.J., Sampson, K.M., Steinhoff, D.F., Ernst, K.C., Ebi, K.L., Jones, B., Hayden, M.H., 2018. The potential impacts of 21st century climatic and population changes on human exposure to the virus vector mosquito Aedes aegypti. Clim. Change 146, 487–500.
Mueller, U., Darowski, N., Fuchs, M.R., Forster, R., Hellmig, M., Paithankar, K.S., Puhringer, S., Steffien, M., Zocher, G., Weiss, M.S., 2012. Facilities for macromolecular crystallography at the helmholtz-zentrum berlin. J. Synchrotron Radiat. 19, 442–449.
Ng, K.K., Arnold, J.J., Cameron, C.E., 2008. Structure-function relationships among RNA-dependent RNA polymerases. Curr. Top. Microbiol. Immunol. 320, 137–156.
Reed, W., Carroll, J., Agramonte, A., Lazear, J.W., 1900. The etiology of yellow fever-A preliminary note. Public Health Pap Rep 26, 37–53.
Robert, X., Gouet, P., 2014. Deciphering key features in protein structures with the new ENDscript server. Nucleic Acids Res. 42, W320–W324.
Rusanov, T., Kent, T., Saeed, M., Hoang, T.M., Thomas, C., Rice, C.M., Pomerantz, R.T., 2018. Identification of a small interface between the methyltransferase and RNA polymerase of NS5 that is essential for Zika virus replication. Sci. Rep. 8, 17384.
Sebera, J., Dubankova, A., Sychrovsky, V., Ruzek, D., Boura, E., Nencka, R., 2018. The structural model of Zika virus RNA-dependent RNA polymerase in complex with RNA

for rational design of novel nucleotide inhibitors. Sci. Rep. 8, 11132.
Selisko, B., Papageorgiou, N., Ferron, F., Canard, B., 2018. Structural and functional basis of the fidelity of nucleotide selection by flavivirus RNA-dependent RNA polymerases. Viruses, vol. 10.
Stephen, P., Baz, M., Boivin, G., Lin, S.X., 2016. Structural insight into NS5 of Zika virus leading to the discovery of MTase inhibitors. J. Am. Chem. Soc. 138, 16212–16215.
Upadhyay, A.K., Cyr, M., Longenecker, K., Tripathi, R., Sun, C., Kempf, D.J., 2017. Crystal structure of full-length Zika virus NS5 protein reveals a conformation similar to Japanese encephalitis virus NS5. Acta Crystallogr F Struct Biol Commun 73, 116–122.
Wu, J., Bera, A.K., Kuhn, R.J., Smith, J.L., 2005. Structure of the Flavivirus helicase: implications for catalytic activity, protein interactions, and proteolytic processing. J. Virol. 79, 10268–10277.
Yang, X., Smidansky, E.D., Maksimchuk, K.R., Lum, D., Welch, J.L., Arnold, J.J., Cameron, C.E., Boehr, D.D., 2012. Motif D of viral RNA-dependent RNA polymerases determines efficiency and fidelity of nucleotide addition. Structure 20, 1519–1527.
Zhang, Y., Corver, J., Chipman, P.R., Zhang, W., Pletnev, S.V., Sedlak, D., Baker, T.S., Strauss, J.H., Kuhn, R.J., Rossmann, M.G., 2003. Structures of immature flavivirus particles. EMBO J. 22, 2604–2613.
Zhao, B., Yi, G., Du, F., Chuang, Y.C., Vaughan, R.C., Sankaran, B., Kao, C.C., Li, P., 2017. Structure and function of the Zika virus full-length NS5 protein. Nat. Commun. 8, 14762.
Zhao, Y., Soh, T.S., Zheng, J., Chan, K.W., Phoo, W.W., Lee, C.C., Tay, M.Y., Swaminathan, K., Cornvik, T.C., Lim, S.P., Shi, P.Y., Lescar, J., Vasudevan, S.G., Luo, D., 2015. A crystal structure of the Dengue virus NS5 protein reveals a novel inter-domain interface essential for protein flexibility and virus replication. PLoS Pathog. 11, e1004682.

SCIENTIFIC REPORTS

OPEN

The structural model of Zika virus RNA-dependent RNA polymerase in complex with RNA for rational design of novel nucleotide inhibitors

Received: 4 April 2018
Accepted: 9 July 2018
Published online: 24 July 2018

Jakub Šebera[1], Anna Dubankova[1], Vladimír Sychrovský[2], Daniel Ruzek[3,4], Evzen Boura[1] & Radim Nencka[1]

Zika virus is a global health threat due to significantly elevated risk of fetus malformations in infected pregnant women. Currently, neither an effective therapy nor a prophylactic vaccination is available for clinical use, desperately necessitating novel therapeutics and approaches to obtain them. Here, we present a structural model of the Zika virus RNA-dependent RNA polymerase (ZIKV RdRp) in complex with template and nascent RNAs, Mg^{2+} ions and accessing nucleoside triphosphate. The model allowed for docking studies aimed at effective pre-screening of potential inhibitors of ZIKV RdRp. Applicability of the structural model for docking studies was illustrated with the NITD008 artificial nucleotide that is known to effectively inhibit the function of the ZIKV RdRp. The ZIKV RdRp – RNA structural model is provided for all possible variations of the nascent RNA bases pairs to enhance its general utility in docking and modelling experiments. The developed model makes the rational design of novel nucleosides and nucleotide analogues feasible and thus provides a solid platform for the development of advanced antiviral therapy.

Zika virus (ZIKV) is an emerging mosquito-borne member of the family *Flaviviridae*, genus *Flavivirus*. This genus includes several major human pathogens like the dengue virus, Japanese encephalitis virus, West Nile virus, and tick-borne encephalitis virus. Virtually the entire human population lives where at least one flavivirus species is endemic[1]. The ZIKV was first identified in 1947 in Uganda[2]. Until its sudden emergence in Brazil in 2015 and explosive spread through South America, Central America, and the Caribbean, the virus had been thought to produce a rare and mild, self-limiting disease in humans, with common symptoms that include fever, rash, joint pain, and conjunctivitis[3]. However, during the recent outbreak it was discovered that ZIKV infection is linked to adverse pregnancy and birth defects, most notably microcephaly and other serious brain anomalies[4]. In adults, ZIKV can cause rare but severe neurological complications, including Guillain-Barré syndrome (a poorly understood immunopathological disease affecting the nervous system that may lead to paralysis or death), acute disseminated encephalomyelitis (an immune-mediated inflammatory demyelinating condition that predominately affects the white matter of the brain and spinal cord), acute myelitis, meningitis or encephalitis[5–9] with fatal consequences in isolated cases[7,10]. During the most recent outbreak in Brazil, an estimated 440 000–1 300 000 cases of ZIKV infection have been reported[11].

Flaviviruses are small, enveloped viruses carrying genomes which consist of non-segmented single-stranded positive-sense RNA (+ RNA). The genome contains one large open reading frame (ORF), which is flanked with

[1]Gilead Sciences Research Centre at IOCB Prague, Institute of Organic Chemistry and Biochemistry of the Czech Academy of Sciences, Praha, Czech Republic. [2]Institute of Organic Chemistry and Biochemistry of the Czech Academy of Sciences, Praha, Czech Republic. [3]Veterinary Research Institute, Hudcova 70, CZ-62100, Brno, Czech Republic. [4]Institute of Parasitology, Biology Centre of the Czech Academy of Sciences, Branisovska 31, CZ-37005, Ceske Budejovice, Czech Republic. Jakub Šebera and Anna Dubankova contributed equally to this work. Correspondence and requests for materials should be addressed to E.B. (email: boura@uochb.cas.cz) or R.N. (email: nencka@uochb.cas.cz)

untranslated regions[12]. Viral RNA bears a type-1 cap structure (m7GpppAm) at the 5′-end and has no poly(A) tail at the 3′-end. The ORF encodes a single polyprotein that is subsequently processed into three structural (C, prM/M, and E) and seven non-structural (NS1, NS2A, NS2B, NS3, NS4A, NS4B, and NS5) proteins[12]. Central enzymatic activities related to virus replication are encoded by NS3 and NS5 proteins. NS3 is the second largest viral protein and possesses both the protease and helicase activities, with NS2B serving as a cofactor for the protease[13]. The NS5 protein is the largest virus-encoded protein, comprised of about 900 amino acids (AAs), and is also the most highly conserved protein among flaviviruses[14]. Approximately 270 AAs at the N-terminus form the methyltransferase (MTase) domain, and approximately 620 AAs at the C-terminus form the RNA-dependent RNA polymerase (RdRp)[15]. The MTase is involved in type 1 cap synthesis at the 5′-end of the viral RNA. At least two activities of the MTase were recognized; i.e., guanine N7 methyltransferase and nucleoside 2′-*O* ribose methyltransferase[16]. No MTase domain was found in the other genera of the *Flaviviridae* family, i.e., hepaciviruses and pestiviruses, which contain an internal ribosome entry site at its 5′-end of viral genome instead of the type 1 cap[17].

The polymerase domain of NS5′ catalyses *de novo* RNA synthesis to first generate −RNA (negative-sense RNA) using the viral +RNA as a template. In the next round, the −RNA serves as a template for the synthesis of more +RNA strands which are used as mRNA for viral polyprotein synthesis or are packaged into arising viral particles[12]. NS5 is considered to be one of the most promising targets for antiviral drugs because both domains (MTase and RdRp) are essential for virus replication and are not present in non-infected cells[18–22].

Recently, the structural biology of ZIKV RdRp significantly advanced as several crystal structures of the RdRp were analysed[23,24] in addition to the previously solved structure by AbbVie Inc[15]. Additionally, several crystal structures of other flavivirus polymerases were solved in complex with small molecules[25,26], however, neither of these structures contained an RNA molecule nor any nucleotide or nucleotide analog, which hinders their use in structure based design of novel inhibitors. For these reasons we endeavoured to build a sophisticated model of RNA and nucleotide bound ZIKV RdRp based on the published crystal structures of ZIKV RdRp and its most similar RdRp bound to RNA that can be found in the PDB database which is the Hepatitis C virus (HCV) RdRp. Such a functional model of ZIKV RdRp containing all the necessary components is indispensable for successful docking experiments, which are an essential part of the rational design of inhibitors that might serve as drugs against ZIKV and other flaviviruses. Correctness of the ZIKV RdRp:RNA structural model developed in this work is demonstrated biochemically using point mutants predicted from the structural model and the usefulness of our model is documented using docking experiments.

Material and Methods

Structural Model. The QM/MM structural model was derived from the ZIKV RdRp (PDB ID 5TFR)[15] and HCV (PDB ID 4WTD)[27] crystal structures. The protein without substrate RNA was captured in 5TFR whereas the template RNA, nascent RNA and accessing ADP nucleotide within catalytic site were captured in 4WTD. The *S*-adenosylhomocysteine cofactor, bound to the MTase domain, was removed from the ZIKV structure. The structural alignment of ZIKV and HCV (chain B) performed using the Measures of Structural Similarity algorithm[28] implemented in the Chimera[29] yielded our starting structural model. The alignment of HCV and ZIKV RdRp involved 174 amino acid pairs and the RMSD between 174 pruned atom pairs was 1.123 Å (Figures S1 and S2). The two Mn^{2+} ions within the HCV catalytic core including their x-ray-resolved water ligands (two water molecules) were included in the structural model but were replaced by Mg^{2+} atoms because magnesium is the physiological ligand. The starting structural model thus corresponded to ZIKV 5TFR structure where the RNA molecules plus the x-ray resolved solvent within the catalytic core were adopted from HCV 4WTD structure. The ADP, which is the accessing RNA nucleotide in HCV, was in our structural model modified to ATP by adding a protonated γ-phosphate. The geometry of added γ-phosphate was adopted from the geometry of ATP in the x-ray structure of human norovirus polymerase (PDB ID 3H5Y)[30]. One of the two x-ray-resolved water molecules was removed because it occupied the position of the added γ-phosphate. The template RNA and nascent RNA strands were shortened to include only the last A^3-U^8 and consecutive nascent $U^2 \ldots A^{602}$ base pairs, where the A^{602} was accessing ATP (Fig. 1). The numbering of RNA nucleotides was adopted from HCV 4WTD structure and used throughout the text. The ZIKV RdRp complex was protonated assuming pH 7.0 using the Propka method[31] and the Maestro program[32].

The initial structural model was solvated to describe the effect on QM/MM-optimized geometries consistently because only one water ligand of the metal cation was crystallographically resolved within the catalytic site (Figs 1A and S3). Moreover, two alternative structural models were prepared assuming a different protonation state of the O3′ oxygen atom in nucleotide U^8 that terminated the nascent RNA strand. One model included O3′-protonated ribose while the other included O3′-deprotonated ribose of U^8. Geometries of the two structural models were preliminarily QM/MM-optimized and the complexes were solvated with the SPC explicit water[33] using the Maestro program. A water layer with a thickness ~10 Å surrounding the ZIKV-RNA complex included 32 354 water molecules. A neutral charge of the solvated complexes was achieved by adding Cl^- anions using the Maestro program. The geometries of the protein, RNAs, two Mg^{2+} ions and the x-ray-resolved water within catalytic core were constrained during MD simulation by imposing the force constant 100.0 kcal mol^{-1} $Å^{-1}$. Hence, only dynamics of the added water molecules and counter-ions was MD-calculated freely, without restraints to encompass the ZIKV-RNA complex with the explicit solvent coherently. The MD simulations were carried out by employing the Desmod program[34]. The OPLS-AA(2005) force field was used in all MD simulation[35]. First, the system was equilibrated employing the constraints and periodic boundary conditions, T = 300 K and P = 101325 Pa. The equilibration was followed by 120 ps of NVT simulation. The solvation of ZIKV-RNA complex was obtained upon 120 ps NPT MD simulation. The solvated structural model used in QM/MM calculations included only the water molecules within the 5 Å layer surrounding the ZIKV RdRp:RNA complex (Fig. 1A). The structural model of the protein-RNA complex was assembled by means of the two x-ray structures. When the interfacial parts of merged moieties are only loosely defined, e.g. in weak and transient complexes, their mutual

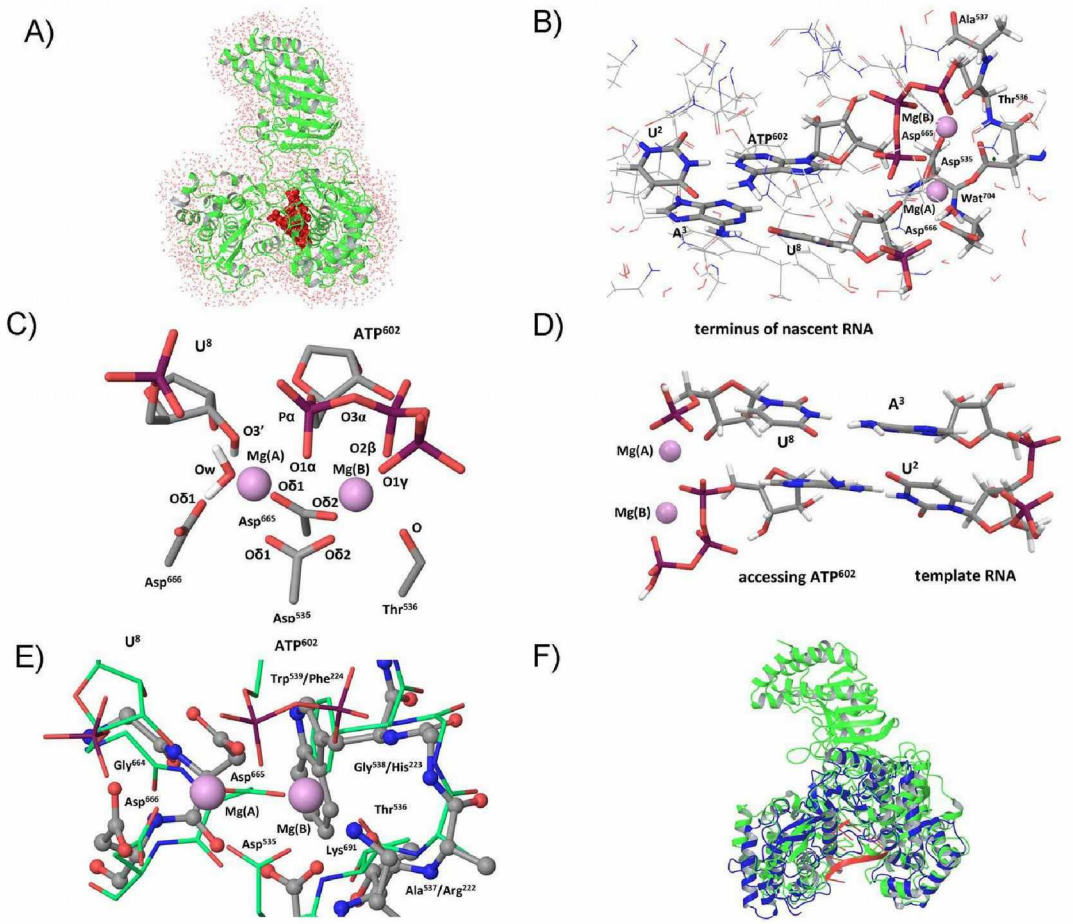


Figure 1. The ZIKV RdRp – RNA complex. (**A**) Sketch of the solvated structural model derived using the x-ray structures of HCV and Zika RdRps that was employed in QM/MM calculations: ZIKV (green), RNA (red), water molecules (red dots) (**B**) Sketch of the QM part (highlighted in bold) within structural model that included ATP^{602} and U^8 nucleotides, U^2 and A^3 bases of paired nucleotides within template RNA, amino acids Asp^{665}, Asp^{666}, Asp^{535}, The^{536}, Ala^{537}, two Mg^{2+} ions and the x-ray resolved water ligand of Mg atom and Wat^{704}. The numbering of RNA was adopted from HCV 4WTD structure. The numbering of the amino acids was adopted from the ZIKV RdRp crystal structure 5TFR (**C**) The schematic sketch of the residues within catalytic core including their numbering used throughout the text. The balls Mg(A) and Mg(B) represent metal cations. (**D**) Sketch of the RNA molecules within catalytic core including two Mg atoms. (**E**) The alignment of catalytic core of HCV (thin-tube depiction including green carbon) and ZIKV RdRp (ball & stick depiction including gray carbon) structures obtained using the Chimera. The Gly^{538}/His^{223}, Trp^{539}/Phe^{224}, and Ala^{537}/Arg^{222} denote the amino acid residues in ZIKV/HCV, respectively. (**F**) The overlay of ZIKV (green) and HCV (blue) inclusive RNA molecule (red) obtained using the Chimera program.

arrangements should be extensively MD-sampled[36]. By contrast, the catalytic core of Zika-RNA can be regarded structurally well-defined and conserved within the family of polymerases[37]. The MD approach was, therefore, used only for relaxation of the initial complex while the QM/MM modelling approach ensured accurate optimizations of the catalytic core including key protein-RNA interactions in response to particular RNA sequences. Lastly, the structural models including all possible nascent RNA base pairs were prepared. For the structural model including O3′-deprotonated ribose of U^8 four structural models included A^{602}-U^2, U^{602}-A^2, G^{602}-C^2 and C^{602}-G^2 base pairs. The structural model including normal U^8 included merely A^{602}-U^2. The PDB files including respective five structural models can be found in the Supporting Information.

The QM/MM Calculation. The QM/MM method was used for optimization of geometry of the ZIKV-RNA complex. The QM and MM part of structural model was described employing the QSite 6.1 program. The QM part included nucleotide-triphosphate N^{602}, N=A, U, G, C, nucleotide U^8, bases of the paired nucleotides within template RNA, amino acids Asp^{665}, Asp^{666}, Asp^{535}, The^{536}, Ala^{537}, two Mg^{2+} ions and the x-ray resolved water ligand of Mg (Fig. 1B). The numbering of the amino acids adopted from 5TFR was used throughout the text. The charge of the QM part was -3 e^- and -4 e^- for structural model including O3′-deprotonated U^8 or A^8 and normal U^8 nucleotide, respectively. The total charge of QM/MM model that included O3′-deprotonated U^8 or A^8 was 6 e^-. The total charge of the model with normal U^8 was 8 e^-. The QM part was calculated using the DFT method employing the PBE functional[38] and the lacvp* basis[39]. The MM part was calculated using the OPLS-AA(2005)

force field[35]. Only the x-ray geometry of the sugar-backbone part of the template RNA was fixed in all QM/MM geometry optimizations, the rest of the geometric parameters was optimized. All the QM/MM-calculated energy minima were validated with the vibration analysis employing the latest Hessian as implemented within the QSite program[40]. Note, actual accuracy of the calculated structures de facto corresponds to the accuracies of original x-ray structures employed in preparation of initial Zika-RNA model. Somewhat higher precision of geometrical parameters throughout the text, however, corresponds to the accuracy assumed for the QM/MM method.

Protein expression and purification. Wild type ZIKV RdRp and all the six mutants were expressed in *E. coli* Rosetta Gami B (DE3) cells and purified as described in detail previously[20] using our using our standard protocols[41,42]. Briefly, the ZIKV RdRp (MR766 strain, residues 35–903 which corresponds to residues 2551–3419 of the polyprotein) was expressed with a His_8SUMO purification and solubilisation tag. Bacterial lysate was prepared using French pressure cell press (Thermo). Proteins were purified by combination of Ni-NTA chelation chromatography followed by the cleavage of the purification tag by the Ulp1 protease and finally purified using size exclusion chromatography (SEC). The purified protein was concentrated to 10 μM concentration and stored at −80 °C in SEC buffer (500 mM NaCl, 20 mM Tris pH 8, 3 mM βME, 5% glycerol, 2 mM $MgCl_2$) until needed.

Enzymatic activity assays. The activity of ZIKV RdRp was measured as described previously[20]. Briefly, it was measured as a frequency of radioactivity labeled ATP* (alpha – P32) incorporation into prolonging chain. We used 0.2 ng/μl poly-U (Polyuridylic acid homopolymer, Sigma) as a template, 10 nM 15-mer poly-A primer and 1 μM ZIKV RdRp in buffer (50 mM Tris-HCl pH 7.5, 40 mM NaCl, 1 mM DTT, 2.5 mM MgCl2, 0.1 mg/ml BSA and 0.2 U/μl RNase inhibitor, Promega). The reaction was started by addition of 2.5 μM ATP and 3000 Ci/mmol ATP* (alpha -P32) and was running for 40 minutes at 30 °C. The RNA product was caught on anion exchange paper (Whatman, UK). Unbound radioactivity was washed out 10x with 125 mM Na_2HPO_4 (pH 9), twice with water and once with an ethanol flush. The radiation was measured by a Phosphorimager (Typhoon 9410) and the signal was quantified using ImageJ software[43]. Reactions without RdRp were used as a negative control and its signal was subtracted from all testing signals. The normalized activity was calculated as a ratio of wild type RdRp (100% activity).

Docking. *Protein preparation.* The QM/MM-optimized structural model including ATP (au_o3.pdb) was used for all the docking experiments. Initially, the ATP as well as water molecules were removed. The pdbqt file was prepared with the AutoDock Tools 1.5.6[44] using the default methodology (the polar hydrogen atoms were maintained in the model). The charges of Mg^{2+} ions were calculated for the QM part of QM/MM-optimized protein-RNA complex that included ATP^{602} and U^8 nascent RNA base pair excluding the A^3 and U^2 nucleobases. The four hydrogen caps bridging the QM part with the MM part were replaced by hydrogen atoms that were geometry optimized with the PBE method and lapcv* basis using the Jaguar 8.2 program[45]. The RESP[46] charges of the magnesium atoms were calculated using the B3LYP[47,48] method and 6–31 G(d,p)[49,50] basis set with the Gaussian09.D01[51] quantum chemistry package and Amber tool antechamber[52]. The charge +1.1 e^- was calculated for the two magnesium atoms by using both alternatives of the structural models including normal and O3′-deprotonated ribose of U^8.

The QM-calculated charges of the magnesium ions were manually inserted into the pdbqt file. We also extracted several water molecules (number 10, 15, 5056, 5594 and 5382 in our supplementary model files), which play an essential role in the binding site, from the original pdb file into a separate pdb file, prepared pdbqt file containing these water molecules in AutoDock Tool 1.5.6 and manually inserted them back into the pdbqt file with protein and ions.

Ligand preparation. ATP was extracted from the original model into a separate pdb file and converted to pdbqt by AutoDock Tools 1.5.6. We used ACD/ChemSketch 12.01 for *de novo* preparation of the ligand NIT008 and converted the 2D structure to a 3D structure using standard function with the force field based on CHARMM parametrization. The 2′-endo (South) and 3′-endo (North) conformations of this compound were prepared and optimized using B3LYP-D3[53] method and 6–31(d,p) basis set with inclusion of the implicit water environment modelled by Polarizable Continuum Model (PCM)[54]. The necessary conversion of the format was performed using AutoDock Tools 1.5.6.

Docking validation and docking experiments. We have performed several validation experiments, which show the importance of all the components (RNA, nucleotide, RdRp) for successful docking experiments. All these experiments were performed using Autodock Vina 1.1[55] with the default docking method at larger exhaustiveness = 200 and greater binding space (binding pocket of 30 × 30 × 30 Å centered at −2.55, −7.53 and 1.67 Å) in comparison with the subsequent docking experiments. The docking experiments were also performed with the Autodock Vina 1.1 using the default docking method with exhaustiveness = 100 under similar conditions as in our previous work[56,57]. The docking was performed into the smaller binding pocket of 20 × 20 × 20 Å centred at −2.55, −7.53 and 1.67 Å.

Results

The QM/MM calculations of ZIKV RdRp:RNA complex. The ZIKV RdRp:RNA model employed in hybrid quantum mechanics/molecular mechanics (QM/MM) calculations was derived from the ZIKV RdRp (PDB ID 5TFR) and HCV (PDB ID 4WTD) crystal structures because no flaviviral RdRp was so far crystallized in complex with RNA. The two structures were superimposed using the Measures of Structural Similarity algorithm[28] as implemented in the Chimera[58]. The alignment of HCV and ZIKV RdRp involved 174 amino acid pairs and the RMSD between 174 pruned atom pairs was 1.123 Å (Figures S1 and S2). Next, the RNA and

x-ray-resolved solvent in the HCV structure were transposed into the ZIKV structure and the two Mn^{2+} ions atoms were replaced by the physiological Mg^{2+} ions. Finally, the ADP from HCV was modified to ATP as detailed in the M&M section. The RNA strands were shortened to include only the A^3-U^8 and nascent U^2…A^{602} base pairs (the numbering of residues was adopted from the HCV RdRp structure 4WTD). The protein-RNA complex was protonated assuming pH = 7.0 and solvated by explicit water molecules keeping all geometry parameters derived from the two crystal structures fixed as is summarized in Fig. 1. Two alternative structural models, one including normal ribose of U8 (Fig. 1C, supplementary PDB file au_o3h) and one included O3′-deprotonated ribose of U^8 (Fig. 1D, supplementary PDB files au_o3′, cg_o3′, gc_o3′, ua_o3′) describe the two possible states of the principal catalytic nucleotide. The solvated structural models furthermore included four possible nascent RNA base pairs to model the catalytic core including all the four NTP^{602}, N=A, C, G, U.

We discovered that the ordering of the RNA molecules within catalytic site indicated the A-form of the nascent RNA albeit local geometries of nucleotides and their mutual arrangements did not correspond precisely to the canonical A-RNA. At first, the class of backbone of the template RNA was not the A-form although the majority of the backbone torsions fit the A-RNA values (Table S1). Secondly, the bases within base pairs did not lie in one plane. Interestingly, the starting crystal structure was more of a buckle-type of mutual declination (Figure S4) while the deviations from their ideally planar arrangements within QM/MM-optimized base pairs was more likely a propeller-twist (Figures S5–S9). In any case, the mutual declination of bases in base pairs was noticeable. The propeller-twist in QM/MM-optimized base pairs ranged from ca 20° to 30° and in the U^{602}-A^2 pair it was even 41° (Table S4). The buckle angle of the base-planes in the case of the HCV polymerase bound to RNA (pdb code 4WTD) was far smaller (Table S4). This incoherence is most likely due to no geometry restraints in the QM/MM method compared to the crystallographic refinement employing planarity geometric restrains for the nucleobases in structures with a resolution worse than 1 Å (The original crystal structure 4WTD was solved at 2.7 Å resolution.). Therefore, the out-of-plane deviations of the inner-ring atoms of the nucleobases within DNA and RNA molecules can be observed only in crystal structures obtained at ultra-high resolution (1.0 Å or better) that are refined without geometry based restrains[59]. Just minutely disordered geometries of the RNA nucleotides in the RdRp:RNA complex as compared to canonical A-RNA inevitably resulted in slightly imperfect geometries of the nascent base pairs (Figure S4). However, the base pairing pattern adopted from the catalytic core of the HCV polymerase was sustained during all QM/MM calculations (Figures S5–S9), which illustrated considerable versatility of the catalytic site with respect to all possible sequences of captured RNAs. The hydrogen bonds within the canonical Watson-Crick base pairs can be derived from the respective interatomic distances between the H-bonded atoms in the crystal structure (Figure S4) and in all QM/MM-calculated geometries (Figures S5–S9). The QM/MM geometry optimization resulted in rather directional (linear) character of the hydrogen bonds within the base pairs, however, at the expense of their notable propeller twist. In any case, both the X-ray and QM/MM geometries of base pairs illustrated rather notable deviation of the RNA strands within ZIKV RdRp catalytic core from the ideal A-form RNA duplex.

Interestingly, the extent of mutual declination of the bases in the base pairs depended on the RNA sequence (Table S4). Probably the most striking was the difference of propeller-twist angles 21° and 8° calculated in U^8-A^3 and A^8-U^3 base pair, respectively. A closer look into the two base pairs unveiled that notable out-of-plane geometries of the paired bases pairs allowed RNA sequence-specific interactions that would not be possible in canonical A-RNA (Figure S5). The inter-base pair interactions involving NH_2 and CO groups of bases were possible due to the imperfect geometries of nascent RNA duplexes and particularly due to their sequence that involved combinations of suitable stacked purine-pyrimidine bases (e.g. the upper structure in Figure S5). These interactions would be impossible in the case of the stacked purine-purine and/or pyrimidine-pyrimidine bases (e.g. the bottom structure in Figure S5). Lack of the NH_2…O=C inter-base interactions between homologous stacked bases likely explains the smaller propeller-twist of A^8-U^3 relative to that of U^8-A^3 (Figure S5). Though locally different, the overall QM/MM-calculated geometries of RNA nucleotides corresponded to the x-ray geometry of the HCV catalytic core, which represents a reliable initial state of the catalytic reaction where the accessing nucleotide is going to be connected to the growing nascent RNA strand according to the template RNA (Fig. 1D).

The QM/MM-calculated geometries of RNA molecules within the catalytic site illustrated some aspects that could be important with respect to the function of the Zika polymerase. Notably, the geometries of the two Mg^{2+} cations coordinated to both RNA and protein revealed their putative function. Close proximity of the O3′ oxygen of terminal nucleotide 8 to Mg^{2+}(A) cation and to the Pα atom of the NTP (Fig. 1C) indicates activation of the terminal RNA nucleotide within the nascent RNA strand that facilitates formation of the O3′-Pα phosphodiester linkage. The QM/MM-calculated O3′-Mg^{2+}(A) distance depended critically on whether the terminal nucleotide 8 of nascent RNA was or was not O3′-deprotonated (Table S3) while the dependence on RNA sequence was negligible (Table S2). Noteworthy, the metal-ligand distances calculated for Mg^{2+} and Mn^{2+} differed (Table S2), which may explain the metal-specific effect on catalytic function that was reported for the HCV RdRp[60]. For O3′H-form of terminal RNA nucleotide 8 the calculated O3′-Mg(A) and O3′-Mn(A) distance was 2.194 Å and 2.280 Å, respectively. The O3′-Mn(A) distance in the HCV RdRp crystal structure was 2.129 Å. For O3′-deprotonated nucleotide 8, the O3′-Mg(A) distance ranged from 1.970 Å to 1.977 Å (Table S2) and the O3′-Mn(A) distance was 1.989 Å (Table S3). The QM/MM calculations thus indicated that the O3′H-form of U^8 was captured in the HCV RdRp crystal structure. The activation of the RNA molecules with regard to the catalytic reaction due to O3′-deprotonation of terminal nucleotide of nascent RNA was further demonstrated by shortening the distance between the O3′ of nucleotide 8 and Pα of NTP^{602}. For ATP, the O3′ nucleotide 8-Pα shortened upon O3′-deprotonation from 3.142 Å to 2. 926 Å. For the remaining three NTPs and O3′-deprotonated ribose of nucleotide 8, the O3′-Pα ranged from 2.892 Å to 2.929 Å (Table S5). The O3′-Pα in HCV crystal structure was 3.366 Å, which again indicates that the prevalent, normal form of the terminal RNA nucleotide 8 was captured in the HCV RdRp crystal. The O3′-Pα calculated with Mn^{2+} for normal and O3′-deprotonated ribose of nucleotide 8 was 3.268 Å and 3.076 Å, respectively. The activated state of the RNA substrate, however, assumes

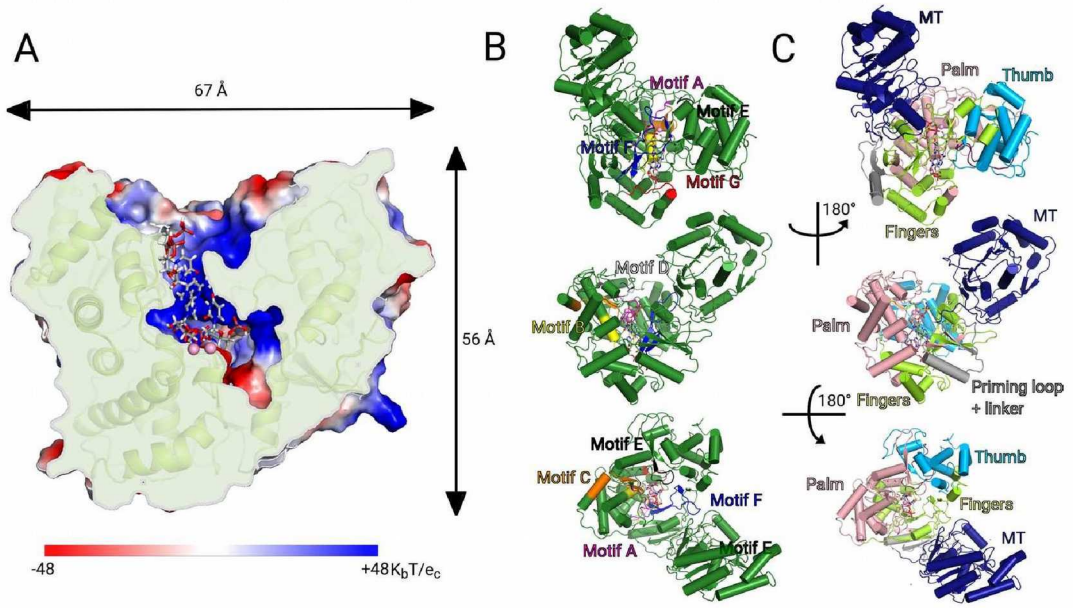


Figure 2. ZIKV RdRp model – biological significance. (**A**) The cross-section of ZIKV RdRp, the electrostatic surface is shown in red – blue conversion. Positively charged areas are blue and negatively charged areas are red. The RNA binding cavity in the protein catalytic core has positive charge that is complementary to the negatively charged RNA. (**B**) Structural motifs representations. Motif A (pink), Motif B (lime), Motif C (orange), Motif D (white), Motif E (black), Motif F (blue), Motif G (red) are highlighted on the ZIKV RdRp model. (**C**) Domains representations. On ZIKV RdRp can be distinguish 4 domains: Methyltransferase – MT (blue), Palm (pink), Fingers (green) and Thumb (cyan).

O3′-deprotonation of the terminal nucleoside within the nascent RNA that takes place in the catalytic core[61]. Consequently, the majority of the QM/MM calculations employed O3′-deprotonated nucleotide 8 to illustrate the reaction state shortly before formation of the O3′-Pα phosphodiester bond.

Biochemical validation of the ZIKV RdRp:RNA complex model. We chose six AAs playing key roles in RNA replication based on our model of the ZIKV RdRp:RNA complex (Fig. 2). Using PyMol (The PyMOL Molecular Graphics System, Version 1.8 Schrödinger, LLC) software, we identified four AAs (Lys^{369}, Arg^{371}, Asn^{373} and Arg^{449}, orange color code in Fig. 3) that our model predicts to be important because they create H – bonds with RNA at a maximum distance of 3.2 Å. We also chose two AAs interacting with Mg^{2+} ions (Asp^{501} and Asp^{631}) that are thus predicted essential (red color code in Fig. 3). As a negative control we used Trp^{505} and Val^{572} (grey color code in Fig. 3) that are located in the proximity of RNA, however, their interactions with RNA or Mg^{2+} are not essential on our model (Fig. 3). Next, we prepared individual recombinant ZIKV RdRps each bearing a single mutation in aforementioned AAs (mutation to Ala).

The activity of all these recombinant RdRp was measured as the ability to incorporate ATP* (alpha – P32) into elongated polyA RNA and compared with the wild type ZIKV RdRp (Fig. 3B). The ZIKV RdRp mutants involving AAs that create H – bond(s) with RNA (Lys^{369}Ala, Arg^{371}Ala, Asn^{373}Ala, and Arg^{449}Ala) had significantly lower activity compared to the wild type protein (decrease of 65%). Total loss of enzymatic activity was measured with the mutants involving AAs that interact with Mg^{2+} (Asp^{501} or Asp^{631}mutated to alanine). The enzymatic activity of RdRp with mutation Trp^{501}Ala was also reduced (40% reduction). Tryptophan is a large aromatic amino acid and although Trp^{505} does not directly participate in RNA polymerization (based on the model) it might be important for the stability of ZIKV RdRp. As was assumed, the mutation Val^{572}Ala in ZIKV RdRp had negligible effect on the polymerase activity.

Docking validation and docking experiments. The docking experiments and validation were performed by the open-source software AutoDock Vina 1.1 using the default methodology[55]. Water molecules were omitted for the docking experiments except for the water molecules within the active site.

Initially, we have performed several experiments to validate the importance of the distinct components of the presented model. Firstly, we have successfully tested this approach by the redocking of ATP into the model using two different grid-box sizes. Both these experimental setups resulted in similar results (Tables S6 and S7) that are summarized in the Fig. 4. Congruently, as also shown in Fig. 5A and Figure S11, the pose of re-docked ATP nicely correlates with its original position within the binding site of the RdRp that was QM/MM-optimized. In the second part of the validation process, we have gradually removed the parts of the model in the active site and evaluated the effect on docking. Docking of ATP into the model missing all the implemented components (nascent and template RNA, Mg^{2+} and waters) resulted in completely incorrect pose of ATP in the binding site (Fig. 4C). Docking into a model with removed Mg^{2+} ions and water molecules gave ATP with significant distortion of the

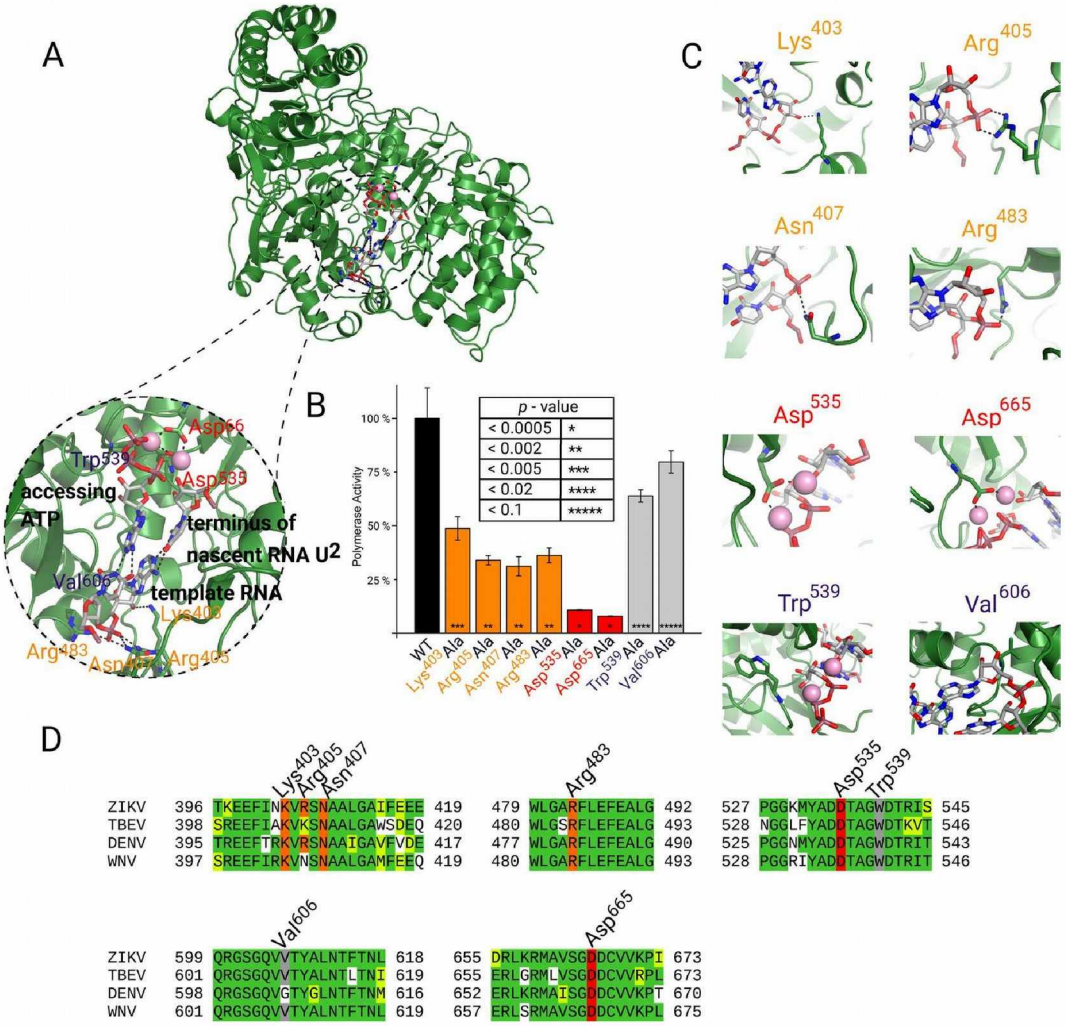


Figure 3. Enzymatic analysis. (**A**) ZIKV RdRp:RNA complex model with an enlarged RNA binding cavity. The RNA template is a dinucleotide composed of an A^3 – U^8 base pair. Nascent U^2 and ATP are present. In enlarged picture are highlighted amino acids related to activity assay. (**B**) Activity of ZIKV RdRp: Mutations Lys^{403}Ala, Arg^{405}Ala, Asn^{407}Ala, Arg^{483}Ala significantly decrease activity (orange), whereas mutation Asp^{535}Ala or Asp^{665}Ala (red) totally disrupts RdRp enzymatic activity. In mutations Trp^{539}Ala and Val^{606}Ala (gray) were not achieved any significant difference. Values in box are Student t-test *p* values. (**C**) Detail of mentioned AAs – ligand wt ZIKV RdRp interaction: Lys^{403}, Arg^{405}, Asn^{407} and Arg^{483} provide interactions with template RNA, Asp^{535} and Asp^{665} provide interaction with Mg2+, Trp^{539} and Val^{606} are in the proximity of active site RdRp. (**D**) Amino acids sequence alignment RdRp through Flaviviruses: Zika virus 1 – ZIKV, Tick-borne encephalitis virus Hypr - TBEV, Dengue Virus 1– DENV, West Nile virus – WNV.

triphosphate moiety but maintained interaction with the nascent and template RNA (Fig. 4D). In contrast, when the model lacks either nascent RNA strand or both RNA components, the docking fails due to distortion of the nucleoside moiety (Fig. 4E,F). In the validation step we have also mutated Arg^{473} and Lys^{458} to glycine. This experiments show that while the removal of Arg^{473} has a significant effect on the position of the triphosphate moiety of the docked ATP (Fig. 4G), the mutated Lys^{458} has none or only marginal effect on the docking results (Fig. 4H).

Subsequently, we used this validated model for docking of the well-known flavivirus RdRp inhibitor NITD008 (in its triphosphate form, Figure S10), which exerts significant activity against ZIKV and other flaviviruses[62,63]. We selected this model compound over numerous other 2′-substituted nucleoside derivatives, since compound NITD008 possess significantly sterically demanding ethynyl substituent at 2′-position. This ethynyl substituent requires significant space in the active site of the enzyme in order to be efficiently recognized. For the purpose of docking, we prepared two starting structures of NITD008 in two distinct conformations of the ribose ring – 2′-endo (South) and 3′-endo (North) because the docking algorithm employing AutoDock Vina does not alter the conformation of the ribose ring. The two conformers of NITD008 employed in the docking studies were geometry optimized using the B3LYP-D3[53] method, 6–31(d,p) basis set and by including the effect of implicit water modelled by the Polarizable Continuum Model (PCM)[54].

The results of NITD008 docking clearly show that only the 3′-endo (North) conformation (Fig. 5) fits nicely into the active site without any significant NITD008 distortion. As shown in Fig. 5B,C, the position of the

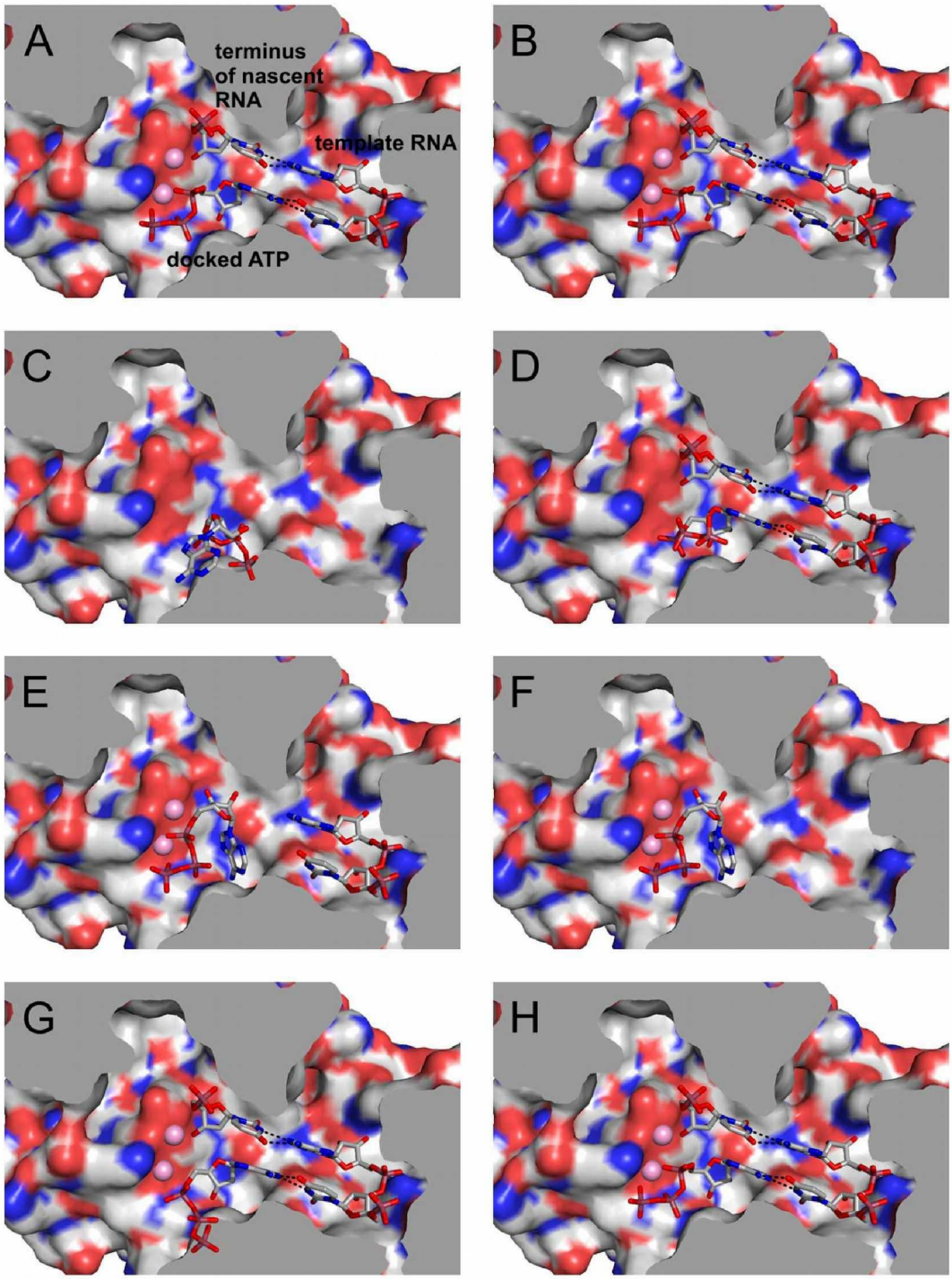


Figure 4. Results of the docking validation. (**A**) Docking of ATP using the standard procedure. (**B**) As in A but using expanded grid-box (30 × 30 × 30 Å). (**C**) Docking into the structure of Zika RdRp without any ligands (**D**) As in C but only magnesium ions and waters in the active site are omitted. (**E**) As in C but only the terminus of the nascent RNA was omitted. (**F**) As in C but only both RNA components were omitted. (**G**) As in A but the with mutated Arg^{473}. (**H**) As in A but with mutated Lys^{458}.

NITD008 is similar to that of ATP and NITD008 occupies a lipophilic cavity of the RdRp protein adjacent to the 2′-position of the ribose moiety. Although the 2′-endo (South) conformer of the NITD008 derivative also directed the ethynyl substituent into this lipophilic cavity, the score calculated for this conformer was significantly lower (9.3 kcal over 10.1 kcal – Tables S14 and S15, Figures S12 and S13) than the score of the 3′-endo (North) counterpart.

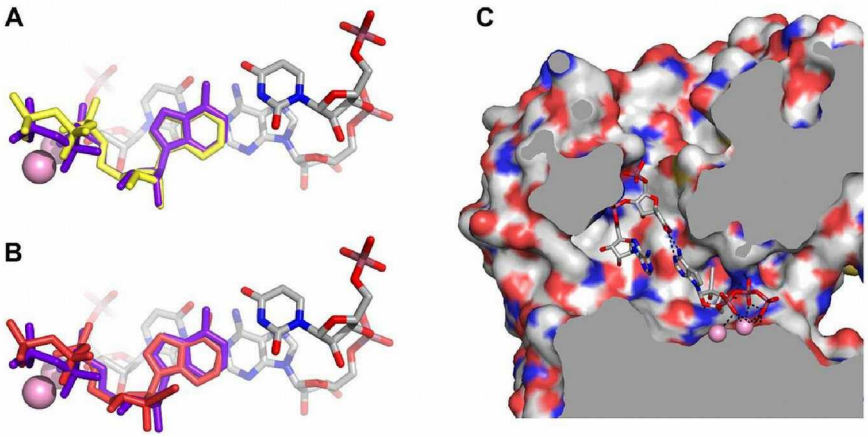


Figure 5. Results of the docking experiments. (**A**) Control re-docking of ATP into the model of ZIKV RNA-dependent RNA polymerase. The overlay of ATP from the original model (purple) and the structure docked by AutoDock Vina (yellow). The docking pose with the highest score is shown. (**B**,**C**) Docking of compound NITD008. (**B**) Overlay of the original ATP in the model (purple) and the highest score pose of NITD008 with *anti* orientation of the nucleobase (red). (**C**) The docking pose of NITD008 reveals significant lipophilic cavity occupied by the large 2′-ethynyl substituent.

Discussion

The QM/MM-optimized model of the ZIKV RdRp:RNA complex highlighted many important interactions among all the partners within the catalytic core that involved Mg^{2+} ions, accessing NTP, template, and nascent RNAs and the protein. First, we will focus on the local geometric parameters to illustrate the structural basis for rationalizing the docking studies and the mutation experiments.

The geometries of the Mg^{2+} ions (Tables S2 and S3) were affected by RNA sequence less than the RNA-protein geometries (Table S5). This illustrates well-conserved structure of the very catalytic core near the key catalytic metal (A) (Fig. 1B) as compared to the rather versatile structure of RNA-protein interface that must be naturally adjusted according to RNA sequence. The lengthening of O3′α-Pα bond in ATP upon O3′-deprotonation of the terminal nucleotide of nascent RNA by 0.028 Å illustrates activation of the accessing NTP. The O3′-deprotonation of the terminal RNA nucleotide 8, at the same time, resulted in shortening of the coordination bond between the catalytic metal and O3′ oxygen of 8 that is favourable with respect to consequential formation of the O3′-Pα phosphodiester bond. The catalytic Mg^{2+} ion (A) clearly provides the principal linkage of the nascent RNA with the accessing NTP that is activated only upon O3′-deprotonation of the terminal nucleotide within the nascent RNA. This assumption is coherent with recent benchmark calculations of the P-O bond formation operated by DNA Polymerases β and λ[61]. Our QM/MM calculations furthermore revealed that the interactions between the NTP phosphate and other RNA molecules remained preserved, only respective bond distances varied upon altered RNA sequence (Table S5). These results clearly show versatility of the forthcoming catalytic reaction irrespective of the RNA sequence. By contrast, the interactions of RNA nucleobases with protein depended far more on RNA sequence, however, some interactions, such as the $O2'^{602}$ – $N\delta_2$ Asn^{612}, $O2'^{8}$ – N Asp^{655} and $O3'^{3}$ – Nε Lys^{403}, were preserved irrespectively of the RNA sequence (Table S5, section Base(RNA)-protein). Otherwise, variety of sequence-dependent changes occurred within the network of nucleobase-protein interactions. Firstly, the replacement of the nucleobases in the RNA molecules resulted in virtually the same interaction, only the contact was mediated by a different atom. For example, the Nε nitrogen of Lys^{458} interacted with N7 nitrogen of ATP^{602}, O6 oxygen of GTP^{602} and N4 nitrogen of UTP^{602} and CTP^{602} in dependence of the RNA sequence (Table S5). Locally, the Nε of Lys^{458} established H-bond with N7 and O6 atoms of the purine and with the C4-amino group of the pyrimidine bases. These adjustments of the local interactions according to the atoms and groups of atoms within RNA nucleobase that occupy one, rather distinct position within catalytic core were also accompanied by a change of the interacting atom within the amino acid. For example, the C_β, $C_{\gamma 2}$ or $C_{\delta 1}$ carbon atoms of Ile^{475} amino acid were involved in highly sequence-dependent interactions with the nucleotides 2 and 602 (Table S5). The calculations of RNA-protein interactions illustrated versatility of the amino acid residues within the ZIKV RdRp catalytic site. Proper deposition of the RNA molecules within the catalytic core, indispensable with regard to the catalytic reaction, was shown to be maintained irrespective of the RNA sequence.

The QM/MM calculations also demonstrated a rather conservative pattern of the RNA phosphate-metal-protein interactions within ZIKV RdRp catalytic core as compared to somewhat more versatile nucleobase-protein interactions that were obviously adjusted according to the particular RNA sequence. Hence, the homologous parts of the RNA molecules involving particularly phosphate and ribose units are deposited within the ZIKV RdRp core in a uniform way regardless of RNA sequence namely due to the effect of the two catalytically important Mg^{2+} cations. Especially the Mg^{2+}(A) was critical with respect to RNA activation as it provides principal chemical linkage of the nascent RNA strand with the accessing NTP. Highly variable nucleobase-protein interaction network demonstrated versatility of the amino acid residues within catalytic core that is indispensable for proper deposition and capturing of RNA molecules irrespectively of their sequence. The contacts among RNAs, the two Mg^{2+} cations and the ZIKV

RdRp, which were geometrically characterized in Tables S2–S5 were employed for design of biochemical experiments for independent validation of our model.

ZIKV RdRp has a right - handed fold with palm, fingers, and thumb domains typical for all polymerases. Seven structural motifs can be distinguished throughout the RdRps[64]. Motifs A – E are included in the palm domain and motifs F – G in the fingers domains (Fig. 2B). Two universally conserved aspartic acids residues (Asp^{535} that is located in motif A and Asp^{665} is in motif C) are responsible for interaction with Mg^{2+} ions and mediate the two metal ion catalysis[65] that enhances substrate recognition and catalytic specificity[66]. Mutation of either Asp^{535} or Asp^{665} results in a total loss of ZIKV RdRp polymerization activity (Fig. 3B). Motif G winds around the RNA template strand and is specific for RdRps. It contains the important residues Lys^{403}, Arg^{405}, Asn^{407} and Arg^{483}. In ZIKV RdRp the Arg^{405}, Asn^{407} and Arg^{483} residues are predicted to interact with the phosphodiester backbone of the template RNA. Mutation of any of these residues to Ala results in a significant loss of polymerization activity (Fig. 3B). Lys^{403} (localized in the beginning of motif G) is highly evolutionary conserved in tick-borne encephalitis virus (TBEV), Dengue virus (DENV) and ZIKV RdRps. Our model predicts that residue Lys^{403} interacts with ribose 3′-hydroxyl group of the template RNA (Fig. 3C). Not surprisingly, the replacement of Lys^{403} by Ala had a significant effect on polymerase activity in accordance with our prediction (Fig. 3B).

Motif B is located in the palm domain and contains Ser - Gly residues responsible for RNA trans-location. However, polymerases with mutated Ser residues retained their ability to incorporate multiple nucleotides[67]. Val^{606} is in close proximity to catalytically important Ser – Gly pair. However, our model predicted that Val^{606} is not important for the enzymatic activity of RdRp. In accordance, we observed only a slight decrease of polymerase activity in the Val^{606}Ala mutant. Motif A participates in the coordination of two catalytic magnesium ions[68] and contains Trp^{559} that is located close to the catalytic centre of ZIKV RdRp. Although our model predicted that Trp^{559} does not form any hydrogen bonds with the substrate we observed somewhat reduced enzymatic activity (~40%) in the Ala mutant. We therefore suggest that the large aromatic Trp^{559} residue is required for proper folding and/or stability of ZIKV RdRp. Alternatively, Trp^{559} could contribute to membrane binding as often aromatic residues do[69,70] and viral RdRps are known to be membrane associated[71,72], however, such an aromatic residue needs to be surface exposed and Trp^{559} is rather buried within close proximity of the active site.

The docking experiments, which we performed using our model with ATP as the accessing NTP, clearly showed that using simple docking software such as AutoDock Vina can lead to satisfactory results. Our experiments with the redocking of ATP resulted in a very nice overlay of the original ATP pose optimized by QM/MM calculations (Fig. 5A). The validation of the model proved that the inserted components (magnesium ions, RNA components and water molecules) play a vital role in docking and their removal leads to significantly incorrect docking results (Fig. 4C–H, Tables S8–S13) in comparison with the docking results obtained for QM/MM optimized model containing these essential parts of the polymerase reaction (Fig. 4A,B, Tables S6 and S7). Subsequently we demonstrated that the compound NITD008, the well-known inhibitor of flavivirus RdRp[73], can be effectively docked into the binding site despite the bulky alkyne substituent at 2′-position.

Nonetheless, several parameters appear to be essential for successful docking. Firstly, the appropriate charge of the magnesium ions seems to be critical for the optimum positioning of the docked ligand, which is in agreement with the docking studies performed by Chen and co-workers[74]. The most favorable results were obtained when the formal charges of magnesium atoms were set to $+1.1\ e^{-}$. (This charge was QM calculated for both magnesium cations). Secondly, the MD calculated and QM/MM geometry optimized solvation of the active site of the RdRp seems to be crucial for formation of appropriate interaction(s) of the triphosphate with the magnesium ions and subsequent position of the nucleobase. However, omission of the water molecules adjacent to the nucleobase might be necessary for appropriate docking of the base modified nucleoside/tide derivatives. Finally, the experiment performed with NITD008 clearly shows that conformation of the ribose ring plays a crucial role in successful docking and it can significantly affect the results. Our data showed that NITD008 in 3′-endo (North) conformation can accommodate the binding site of ZIKV RdRp in an efficient way with minimum distortion of both the nucleobase and the sugar moieties in comparison to ATP in the original model (Fig. 5B). NITD008 in the 2′-endo (South) conformation also fits into the binding site with the alkyne moiety placed in the lipophilic cavity, however, the score was approximately 1 kcal lower as compared to the score of the North conformation. In addition, both the nucleobase and the triphosphate part were significantly disordered. These docking experiments indicated that NITD008 3′-endo (North) conformation is the active conformation against the ZIKV RdRd. It is, therefore, tempting to suggest that the preferred conformation depends on the individual properties of distinct nucleotides rather than being governed by some general rule.

Conclusions

In conclusion, the comprehensive structural model of ZIKV RdRp:RNA including the protein, template and nascent RNAs, approaching nucleoside triphosphate and magnesium ions was derived from available crystal structures. The QM/MM-optimized ZIKV RdRp:RNA models including all possible nascent RNA base pairs are designated for future docking studies aimed at the design of novel nucleotide analogues against the ZIKV RdRp. These models will be also very useful for the rationalization of the biochemical experiments focused on the catalytic function of ZIKV RdRp because no structural model that includes both the protein and the RNAs is available at the moment. Exploitation of the ZIKV RdRp:RNA model was successfully demonstrated with the docking of NITD008, a nucleoside triphosphate analog. The protein-mutation study assuming the picture of catalytic core was rationalized by using our RdRp:RNA model. In particular, the classification of contacts among amino acids within the core, RNA and Mg^{2+} ions was employed in rationalization of measured catalytic function of the six mutants as compared to the activity of the wild type enzyme. The structural model highlighted the importance of particular amino acid residues in direct contact either with RNA molecules and/or magnesium ions. By contrast, mutation of the adjacent residues, which are not involved in the direct contact with these essential components of the reaction, exerted significantly lower effect on polymerase activity. The ZIKV RdRp:RNA models provide

solid basis for molecular modelling of RdRp behaviour in relation with its biochemical function. Therefore, it can be utilized by medicinal chemists for the design of novel nucleoside and nucleotide derivatives as direct-acting antiviral agents against the ZIKV and other flaviviral pathogens.

References

1. Gould, E. A. & Solomon, T. Pathogenic flaviviruses. *Lancet* **371**, 500–509, https://doi.org/10.1016/S0140-6736(08)60238-X (2008).
2. Dick, G. W., Kitchen, S. F. & Haddow, A. J. Zika virus. I. Isolations and serological specificity. *Transactions of the Royal Society of Tropical Medicine and Hygiene* **46**, 509–520 (1952).
3. Wikan, N. & Smith, D. R. Zika virus: history of a newly emerging arbovirus. *The Lancet. Infectious diseases* **16**, e119–126, https://doi.org/10.1016/S1473-3099(16)30010-X (2016).
4. Beckham, J. D., Pastula, D. M., Massey, A. & Tyler, K. L. Zika Virus as an Emerging Global Pathogen: Neurological Complications of Zika Virus. *JAMA Neurol* **73**, 875–879, https://doi.org/10.1001/jamaneurol.2016.0800 (2016).
5. Smith, D. W. & Mackenzie, J. Zika virus and Guillain-Barre syndrome: another viral cause to add to the list. *Lancet* **387**, 1486–1488, https://doi.org/10.1016/S0140-6736(16)00564-X (2016).
6. Niemeyer, B., Niemeyer, R., Borges, R. & Marchiori, E. Acute Disseminated Encephalomyelitis Following Zika Virus Infection. *Eur Neurol* **77**, 45–46, https://doi.org/10.1159/000453396 (2017).
7. Soares, C. N. *et al.* Fatal encephalitis associated with Zika virus infection in an adult. *Journal of clinical virology: the official publication of the Pan American Society for Clinical Virology* **83**, 63–65, https://doi.org/10.1016/j.jcv.2016.08.297 (2016).
8. Mecharles, S. *et al.* Acute myelitis due to Zika virus infection. *Lancet* **387**, 1481–1481, https://doi.org/10.1016/S0140-6736(16)00644-9 (2016).
9. Carteaux, G. *et al.* Zika Virus Associated with Meningoencephalitis. *The New England journal of medicine* **374**, 1595–1596, https://doi.org/10.1056/NEJMc1602964 (2016).
10. Azevedo, R. S. *et al.* Zika virus epidemic in Brazil. I. Fatal disease in adults: Clinical and laboratorial aspects. *Journal of clinical virology: the official publication of the Pan American Society for Clinical Virology* **85**, 56–64, https://doi.org/10.1016/j.jcv.2016.10.024 (2016).
11. Bogoch, I. I. *et al.* Anticipating the international spread of Zika virus from Brazil. *Lancet* **387**, 335–336, https://doi.org/10.1016/S0140-6736(16)00080-5 (2016).
12. Lindenbach, B. D. & Rice, C. M. Molecular biology of flaviviruses. *Advances in virus research* **59**, 23–61 (2003).
13. Zhang, Z. *et al.* Crystal structure of unlinked NS2B-NS3 protease from Zika virus. *Science* **354**, 1597–1600, https://doi.org/10.1126/science.aai9309 (2016).
14. Lescar, J., Lim, S. P. & Shi, P.-Y. Structure and Function of the Flavivirus NS5 Protein. *Molecular Virology and Control of Flaviviruses* 101–117 (2012).
15. Upadhyay, A. K. *et al.* Crystal structure of full-length Zika virus NS5 protein reveals a conformation similar to Japanese encephalitis virus NS5. *Acta Crystallogr F Struct Biol Commun* **73**, 116–122, https://doi.org/10.1107/S2053230X17001601 (2017).
16. Dong, H. P., Zhang, B. & Shi, P. Y. Flavivirus methyltransferase: A novel antiviral target. *Antiviral Research* **80**, 1–10, https://doi.org/10.1016/j.antiviral.2008.05.003 (2008).
17. Caillet-Saguy, C., Lim, S. P., Shi, P. Y., Lescar, J. & Bressanelli, S. Polymerases of hepatitis C viruses and flaviviruses: structural and mechanistic insights and drug development. *Antiviral research* **105**, 8–16, https://doi.org/10.1016/j.antiviral.2014.02.006 (2014).
18. Malet, H. *et al.* The flavivirus polymerase as a target for drug discovery. *Antiviral research* **80**, 23–35, https://doi.org/10.1016/j.antiviral.2008.06.007 (2008).
19. Eyer, L. *et al.* Nucleoside Inhibitors of Zika Virus. *The Journal of infectious diseases* **214**, 707–711, https://doi.org/10.1093/infdis/jiw226 (2016).
20. Hercik, K. *et al.* Adenosine triphosphate analogs can efficiently inhibit the Zika virus RNA-dependent RNA polymerase. *Antiviral research* **137**, 131–133, https://doi.org/10.1016/j.antiviral.2016.11.020 (2017).
21. Hercik, K., Brynda, J., Nencka, R. & Boura, E. Structural basis of Zika virus methyltransferase inhibition by sinefungin. *Archives of virology* **162**, 2091–2096, https://doi.org/10.1007/s00705-017-3345-x (2017).
22. Zhang, C. *et al.* Structure of the NS5 methyltransferase from Zika virus and implications in inhibitor design. *Biochemical and biophysical research communications*. https://doi.org/10.1016/j.bbrc.2016.11.098 (2016).
23. Godoy, A. S. *et al.* Crystal structure of Zika virus NS5 RNA-dependent RNA polymerase. *Nature communications* **8**, 14764, https://doi.org/10.1038/ncomms14764 (2017).
24. Zhao, B. *et al.* Structure and function of the Zika virus full-length NS5 protein. *Nature communications* **8**, 14762, https://doi.org/10.1038/ncomms14762 (2017).
25. Yin, Z. *et al.* N-Sulfonylanthranilic Acid Derivatives as Allosteric Inhibitors of Dengue Viral RNA-Dependent RNA Polymerase. *Journal of Medicinal Chemistry* **52**, 7934–7937, https://doi.org/10.1021/jm901044z (2009).
26. Yap, T. L. *et al.* Crystal structure of the dengue virus RNA-dependent RNA polymerase catalytic domain at 1.85-angstrom resolution. *J Virol* **81**, 4753–4765, https://doi.org/10.1128/JVI.02283-06 (2007).
27. Appleby, T. C. *et al.* Viral replication. Structural basis for RNA replication by the hepatitis C virus polymerase. *Science* **347**, 771–775, https://doi.org/10.1126/science.1259210 (2015).
28. Meng, E. C., Pettersen, E. F., Couch, G. S., Huang, C. C. & Ferrin, T. E. Tools for integrated sequence-structure analysis with UCSF Chimera. *BMC Bioinformatics* **7**, 339, https://doi.org/10.1186/1471-2105-7-339 (2006).
29. Pettersen, E. F. *et al.* UCSF Chimera–a visualization system for exploratory research and analysis. *Journal of computational chemistry* **25**, 1605–1612, https://doi.org/10.1002/jcc.20084 (2004).
30. Zamyatkin, D. F., Parra, F., Machin, A., Grochulski, P. & Ng, K. K. Binding of 2′-amino-2′-deoxycytidine-5′-triphosphate to norovirus polymerase induces rearrangement of the active site. *Journal of molecular biology* **390**, 10–16, https://doi.org/10.1016/j.jmb.2009.04.069 (2009).
31. Li, H., Robertson, A. D. & Jensen, J. H. Very fast empirical prediction and rationalization of protein pK(a) values. *Proteins-Structure Function and Bioinformatics* **61**, 704–721, https://doi.org/10.1002/prot.20660 (2005).
32. Maestro, v. 9.3, (Schrödinger, LLC: New York, 2013).
33. Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W. & Klein, M. L. Comparison of Simple Potential Functions For Simulating Liquid Water. *Journal of Chemical Physics* **79**, 926–935, https://doi.org/10.1063/1.445869 (1983).
34. Desmond Molecular Dynamics System, v. 3.1, (D. E. Shaw Research: New York, 2012).
35. Jorgensen, W. L., Maxwell, D. S. & TiradoRives, J. Development and testing of the OPLS all-atom force field on conformational energetics and properties of organic liquids. *Journal of the American Chemical Society* **118**, 11225–11236, https://doi.org/10.1021/ja9621760 (1996).
36. Jerabek, P., Florian, J., Stiborova, M. & Martinek, V. Flexible docking-based molecular dynamics/steered molecular dynamics calculations of protein-protein contacts in a complex of cytochrome P450 1A2 with cytochrome b5. *Biochemistry* **53**, 6695–6705, https://doi.org/10.1021/bi500814t (2014).
37. Genna, V., Vidossich, P., Ippoliti, E., Carloni, P. & De Vivo, M. A Self-Activated Mechanism for Nucleic Acid Polymerization Catalyzed by DNA/RNA Polymerases. *Journal of the American Chemical Society* **138**, 14592–14598, https://doi.org/10.1021/jacs.6b05475 (2016).

38. Perdew, J. P., Burke, K. & Ernzerhof, M. Generalized gradient approximation made simple. *Physical review letters* **77**, 3865–3868, https://doi.org/10.1103/PhysRevLett.77.3865 (1996).
39. Wadt, W. R. & Hay, P. J. Abinitio Effective Core Potentials for Molecular Calculations - Potentials for Main Group Elements Na to Bi. *Journal of Chemical Physics* **82**, 284–298, https://doi.org/10.1063/1.448800 (1985).
40. Schlegel, H. B. Estimating the Hessian for Gradient-Type Geometry Optimizations. *Theor Chim Acta* **66**, 333–340 (1984).
41. Baumlova, A. *et al.* The crystal structure of the phosphatidylinositol 4-kinase IIα. *EMBO reports* **15**, 1085–1092, https://doi.org/10.15252/embr.201438841 (2014).
42. Klima, M. *et al.* Structural insights and *in vitro* reconstitution of membrane targeting and activation of human PI4KB by the ACBD3 protein. *Scientific Reports* **6** 23641 https://doi.org/10.1038/srep23641 http://www.nature.com/articles/srep23641#supplementary-information (2016).
43. Schindelin, J., Rueden, C. T., Hiner, M. C. & Eliceiri, K. W. The ImageJ ecosystem: An open platform for biomedical image analysis. *Molecular reproduction and development* **82**, 518–529, https://doi.org/10.1002/mrd.22489 (2015).
44. Morris, G. M. *et al.* AutoDock4 and AutoDockTools4: Automated Docking with Selective Receptor Flexibility. *Journal of Computational Chemistry* **30**, 2785–2791, https://doi.org/10.1002/jcc.21256 (2009).
45. Bochevarov, A. D. *et al.* Jaguar: A high-performance quantum chemistry software program with strengths in life and materials sciences. *Int J Quantum Chem* **113**, 2110–2142, https://doi.org/10.1002/qua.24481 (2013).
46. Bayly, C. I., Cieplak, P., Cornell, W. D. & Kollman, P. A. A Well-Behaved Electrostatic Potential Based Method Using Charge Restraints For Deriving Atomic Charges - The Resp Model. *Journal of Physical Chemistry* **97**, 10269–10280, https://doi.org/10.1021/j100142a004 (1993).
47. Becke, A. D. Density-Functional Thermochemistry .3. The Role of Exact Exchange. *Journal of Chemical Physics* **98**, 5648–5652, https://doi.org/10.1063/1.464913 (1993).
48. Lee, C. T., Yang, W. T. & Parr, R. G. Development of The Colle-Salvetti Correlation-Energy Formula Into a Functional of The Electron-Density. *Physical Review B* **37**, 785–789, https://doi.org/10.1103/PhysRevB.37.785 (1988).
49. Hehre, W. J., Ditchfield, R. & Pople, J. A. Self-Consistent Molecular-Orbital Methods .12. Further Extensions of Gaussian-Type Basis Sets for Use in Molecular-Orbital Studies of Organic-Molecules. *Journal of Chemical Physics* **56**, 2257−+ https://doi.org/10.1063/1.1677527 (1972).
50. Harihara, P. & Pople, J. A. Influence of Polarization Functions on Molecular-Orbital Hydrogenation Energies. *Theoretica Chimica Acta* **28**, 213–222, https://doi.org/10.1007/bf00533485 (1973).
51. Frisch, M. J. et al., Inc.: Wallingford, CT,. Gaussian 09, revision D.01 (2013).
52. Case, D. A. *et al.* AMBER (2016).
53. Grimme, S., Antony, J., Ehrlich, S. & Krieg, H. A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu. *Journal of Chemical Physics* **132** https://doi.org/10.1063/1.3382344 (2010).
54. Tomasi, J., Mennucci, B. & Cammi, R. Quantum mechanical continuum solvation models. *Chemical Reviews* **105**, 2999–3093, https://doi.org/10.1021/cr9904009 (2005).
55. Trott, O. & Olson, A. J. Software News and Update AutoDock Vina: Improving the Speed and Accuracy of Docking with a New Scoring Function, Efficient Optimization, and Multithreading. *Journal of Computational Chemistry* **31**, 455–461, https://doi.org/10.1002/jcc.21334 (2010).
56. Mejdrova, I. *et al.* Rational Design of Novel Highly Potent and Selective Phosphatidylinositol 4-Kinase III beta (PI4KB) Inhibitors as Broad-Spectrum Antiviral Agents and Tools for Chemical Biology. *Journal of Medicinal Chemistry* **60**, 100–118, https://doi.org/10.1021/acs.jmedchem.6b01465 (2017).
57. Mejdrová, I. *et al.* Highly Selective Phosphatidylinositol 4-Kinase III beta Inhibitors and Structural Insight into Their Mode of Action. *Journal of Medicinal Chemistry* **58**, 3767–3793, https://doi.org/10.1021/acs.jmedchem.5b00499 (2015).
58. Yang, Z. *et al.* UCSF Chimera, MODELLER, and IMP: an integrated modeling system. *Journal of structural biology* **179**, 269–278, https://doi.org/10.1016/j.jsb.2011.09.006 (2012).
59. Sychrovsky, V. *et al.* Revisiting the planarity of nucleic acid bases: Pyramidilization at glycosidic nitrogen in purine bases is modulated by orientation of glycosidic torsion. *Nucleic Acids Res* **37**, 7321–7331, https://doi.org/10.1093/nar/gkp783 (2009).
60. Luo, G. *et al.* De novo initiation of RNA synthesis by the RNA-dependent RNA polymerase (NS5B) of hepatitis C virus. *Journal of virology* **74**, 851–863 (2000).
61. Klvana, M., Bren, U. & Florian, J. Uniform Free-Energy Profiles of the P-O Bond Formation and Cleavage Reactions Catalyzed by DNA Polymerases beta and lambda. *The journal of physical chemistry. B* **120**, 13017–13030, https://doi.org/10.1021/acs.jpcb.6b08581 (2016).
62. Deng, Y. Q. *et al.* Adenosine Analog NITD008 Is a Potent Inhibitor of Zika Virus. *Open Forum Infectious Diseases* **3** https://doi.org/10.1093/ofid/ofw175 (2016).
63. Lo, M. K., Shi, P.-Y., Chen, Y.-L., Flint, M. & Spiropoulou, C. F. *In vitro* antiviral activity of adenosine analog NITD008 against tick borne flaviviruses. *Antiviral Research* **130**, 46–49, https://doi.org/10.1016/j.antiviral.2016.03.013 (2016).
64. Lu, G. & Gong, P. A structural view of the RNA-dependent RNA polymerases from the Flavivirus genus. *Virus Res* **234**, 34–43, https://doi.org/10.1016/j.virusres.2017.01.020 (2017).
65. Flamand, A., Delagneau, J. F. & Bussereau, F. Rna-Polymerase Activity In Purified Rabies Virions. *Journal of General Virology* **40**, 233–238, https://doi.org/10.1099/0022-1317-40-1-233 (1978).
66. Yang, W., Lee, J. Y. & Nowotny, M. Making and breaking nucleic acids: Two-Mg2+ -ion catalysis and substrate specificity. *Molecular Cell* **22**, 5–13, https://doi.org/10.1016/j.molcel.2006.03.013 (2006).
67. Sholders, A. J. & Peersen, O. B. Distinct Conformations of a Putative Translocation Element in Poliovirus Polymerase. *Journal of Molecular Biology* **426**, 1407–1419, https://doi.org/10.1016/j.jmb.2013.12.031 (2014).
68. Wu, J., Liu, W. & Gong, P. A Structural Overview of RNA-Dependent RNA Polymerases from the Flaviviridae Family. *International Journal of Molecular Sciences* **16**, 12943–12957, https://doi.org/10.3390/ijms160612943 (2015).
69. Hurley, J. H., Boura, E., Carlson, L. A. & Rozycki, B. Membrane budding. *Cell* **143**, 875–887, https://doi.org/10.1016/j.cell.2010.11.030 (2010).
70. Boura, E. & Hurley, J. H. Structural basis for membrane targeting by the MVB12-associated beta-prism domain of the human ESCRT-I MVB12 subunit. *Proceedings of the National Academy of Sciences of the United States of America* **109**, 1901–1906, https://doi.org/10.1073/pnas.1117597109 (2012).
71. Hsu, N. Y. *et al.* Viral reorganization of the secretory pathway generates distinct organelles for RNA replication. *Cell* **141**, 799–811, https://doi.org/10.1016/j.cell.2010.03.050 (2010).
72. Dubankova, A., Humpolickova, J., Klima, M. & Boura, E. Negative charge and membrane-tethered viral 3B cooperate to recruit viral RNA dependent RNA polymerase 3D (pol). *Scientific reports* **7**, 17309, https://doi.org/10.1038/s41598-017-17621-6 (2017).
73. Yin, Z. *et al.* An adenosine nucleoside inhibitor of dengue virus. *Proceedings of the National Academy of Sciences of the United States of America* **106**, 20435–20439, https://doi.org/10.1073/pnas.0907010106 (2009).
74. Chen, D. L. *et al.* Accounting for ligand-bound metal ions in docking small molecules on adenylyl cyclase toxins. *Proteins-Structure Function and Bioinformatics* **67**, 593–605, https://doi.org/10.1002/prot.21249 (2007).