UPPSALA
UNIVERSITET

HANNA ARONSSON

# Biophysical Analysis of Novel Viral-Host Protein-Protein Interactions Identified by Proteomic Peptide Phage Display

*Degree project E in Biochemistry*

June 2020

Department of Medical Biochemistry and Microbiology (IMBIM)
Supervisors: PhD student Filip Mihalič and Prof. Per Jemth

## Abstract

Viral hijacking of cellular processes can be performed in numerous ways. One example of such interference is viral proteins mimicking regions of native proteins which interrupts the naturally occurring protein-protein interactions (PPIs) in host cells. To do this, viral proteins often contain short linear motifs (SLiMs) that are present in non-structural regions, and are prone to evolve quickly. Here, several peptides derived from viral proteins have been investigated in their interaction with two human target proteins, the C-terminal domain of Poly(A)-binding protein (PABPC) and Microtubule-associated proteins 1A/1B light chain 3C (LC3C). PABPC binds to the poly(A)-tail of mRNA and is involved in translation initiation. LC3C is involved in autophagy, where it is a part of the formation of the autophagosome, into which cargo can be brought for degradation. As a starting point, proteomic peptide phage display (ProP-PD) was performed prior to this work, to identify interacting peptides by screening a large viral library. Peptides showing binding capacity were then chosen for further validation utilizing fluorescence polarization (FP) and isothermal titration calorimetry (ITC). The affinities could be determined for all peptides with FP, while most of the ITC data is yet inconclusive. These measurements can be utilized as initial results in an investigation of viral hijacking of translation and autophagy, which will lead to a larger understanding of how viruses take over these pathways in cells for their own benefit.

## Svensk populärvetenskaplig sammanfattning

Mitt i den rådande COVID-19 pandemin, känns det mer aktuellt än på länge att få en djupare förståelse för hur virus påverkar våra celler. Virus kan kapa olika processer i värdceller genom att interagera med cellens egna proteiner och på så sätt störa de naturliga interaktionerna som är viktiga för cellens processer. Proteiner består av länkade aminosyror och de antar vissa strukturer som i sin helhet utgör proteinet. Det finns dock vissa delar av proteiner som inte antar någon specifik struktur. Dessa delar kan tolerera fler punktmutationer jämfört med de strukturella delarna av proteiner. Detta i kombination med att virus också tenderar att ha en relativ hög mutationsfrekvens, gör att deras sådana delar av virala proteiner lättare kan ändras under kort tid och att de därmed kan utvecklas för att påverka vissa delar av en cells processer. I detta projekt har flera peptider som motsvarar dessa ostrukturerade delar av virala proteiner identifierats och validerats i deras interaktioner med två humana målproteiner, Poly(A)-bindande protein (PABP) och Autofagosom proteinet LC3C.

PABP är ett protein som binder till ena änden av mRNA som består av upprepningar av basen adenin. Detta protein är viktigt när andra proteiner ska byggas upp i cellen, där det påverkar strukturen av mRNA, som fungerar som en mall för byggandet. LC3C är ett protein som spelar en stor roll i autofagi processer i cellen. Dessa processer bryter ner olika cellkomponenter så att de till exempel kan återanvändas till att bygga andra eller nya nödvändiga komponenter för cellen.

Genom fagdisplay, har man tidigare kunnat identifiera peptider som verkar interagera med dessa två humana målproteiner och detta har legat som grund för detta projekt. Här har peptiderna och proteinernas interaktioner validerats med metoderna fluorescenspolarisation och kalorimetri, för att undersöka hur starkt de binder till varandra. På så sätt kan man undersöka vilka virala proteiner som är av intresse för viral kapning av PABPs och LC3Cs processer. Alla peptider som har undersökts i detta projekt visade sig binda till målproteinerna, där vissa binder starkare än andra. En fortsättning på detta projekt behövs för att säkerställa att interaktionerna sker i celler och att de sker när proteinerna är i deras fulla längd och inte bara i den korta versionen som är i en peptid.

# List of abbreviations

| | |
|---|---|
| ΔH | Change in enthalpy |
| $A_{215}$ | Absorbance at 215 nm |
| $A_{280}$ | Absorbance at 280 nm |
| DNA | Deoxyribonucleic acid |
| DTT | Dithiothreitol |
| *E. coli* | *Escherichia coli* |
| FITC | Fluorescein isothiocyanate |
| FP | Fluorescence polarization |
| FT | Flow-through |
| GSH | Glutathione |
| GST | Glutathione S-transferase |
| His-tag | Histidine-tag (6x His residues) |
| IDR | Intrinsically disordered region |
| IMAC | Immobilized-metal ion affinity chromatography |
| ITC | Isothermal titration calorimetry |
| LB | Lysogeny broth |
| LC3C | Microtubule-associated proteins 1A/1B light chain 3C |
| LIR | LC3 interacting region |
| mRNA | messenger ribonucleic acid |
| MS | Mass spectrometry |
| Mw | Molecular weight |
| MWCO | Molecular weight cut-off |
| nm | nanometer |
| $OD_{600}$ | Optical density at 600 nm |
| PABPC | C-terminal domain of Poly(A)-binding protein |
| PAGE | Polyacrylamide gel electrophoresis |
| PAIP | PABP interacting protein |
| PBS | Phosphate buffered saline |
| Poly(A) | Polyadenylate |
| PPI | Protein-Protein interaction |
| PPL | Plane polarized light |
| ProP-PD | Proteomic Peptide Phage Display |
| RRM | RNA recognition motif |
| SDS | Sodium dodecyl sulfate |
| SLiMs | Short Linear Motifs |
| ssDNA | Single-stranded DNA |
| w/v | Weight by volume |

# Table of contents

# 1 Introduction

## 1.1 Short linear motifs

Short linear motifs (SLiMs) are short stretches of up to 10 amino acids often found in intrinsically disordered regions (IDRs) of proteins. These motifs have been shown to be involved in, and important for, many different protein-protein interactions (PPIs).[1] IDRs are, as the name suggest, peptide segments in the protein that does not possess a defined secondary or tertiary structure. Since protein function is usually linked to the structure, these regions were for a long time overlooked and the same can be said for the interactions of SLiMs. Since SLiMs have a small interface with their interaction partners, the interactions are often transient and have affinities, $K_d$, in the low to medium micromolar range. This makes them more difficult to identify and study compared to other PPIs.[1] Upon interaction of a SLiM, the IDR might fold into secondary structures such as α-helices or β-sheets and this could influence the affinity and specificity of the binding.[1,2] SLiMs often contain binding motifs, which have specific amino acid sequences that can be identified and described.[2] The affinity and specificity of SLiM-mediated PPIs are achieved through several aspects, such as core motif binding determinants, the context of the motif in the sequence and properties that changes over time, for example concentration of interacting species.[1]

## 1.2 Viral hijacking

Because SLiMs are short and disordered, they make a relatively easy target for viral mimicry.[3] Viral mimicry means that viruses can create a sequence or a structure that resembles a structure of the host on the molecular level and that can be used in a beneficial way.[3] Viruses can evolve SLiMs that compete with PPIs of the host cell and thus hijack different regulatory and signaling pathways.[3,4] SLiMs are subjects of rapid evolution and converges to similar motifs. Because of this, viral SLiMs which are part of already rapidly evolving proteins, can adapt quickly to disturb host cell functions by competing out the common interactors.[3]

## 1.3 Phage display screening

There are several types of high-throughput methods for identification of SLiMs-based interactions, varying from arrays and protein fragmentation assays, to display methods.[5] Phage display was introduced in the 1980's, and can be used to identify and isolate peptides or proteins of interest that bind to a specific target.[6] Filamentous phages are widely used for this technique. On their surface, different coat proteins are present which are used to display a peptide of interest. To do this, the DNA sequences encoding various peptides are fused to the coat protein gene, allowing large combinatorial libraries to be produced with ease. Since the genes of the peptides of interest are fused into the phages such that they are displayed on the surface, this creates the possibility to link pheno- and genotype.[7] Phage libraries that are a final product of such procedures contain a wide variety of peptides presented on their surface.[7]

One application of phage display is proteomic peptide phage display (ProP-PD), which can be used for identification of interactions between proteins and SLiMs. The libraries are produced to display defined regions (peptides) of the target proteome. For phage selection (*Figure 1*), the protein of interest can be immobilized onto a surface, to which the phage library is added. The phages that present peptides which interacts with the target protein will remain during a washing step, while unbound phages will be discarded. The bound phages can be eluted followed by

amplification through infection of *E. coli*. Upon amplification. the selected phages can be used for further rounds of selection by repeating the procedure. When phages have gone through a few rounds of selection, sequencing is performed for identification of interacting peptides.[8]

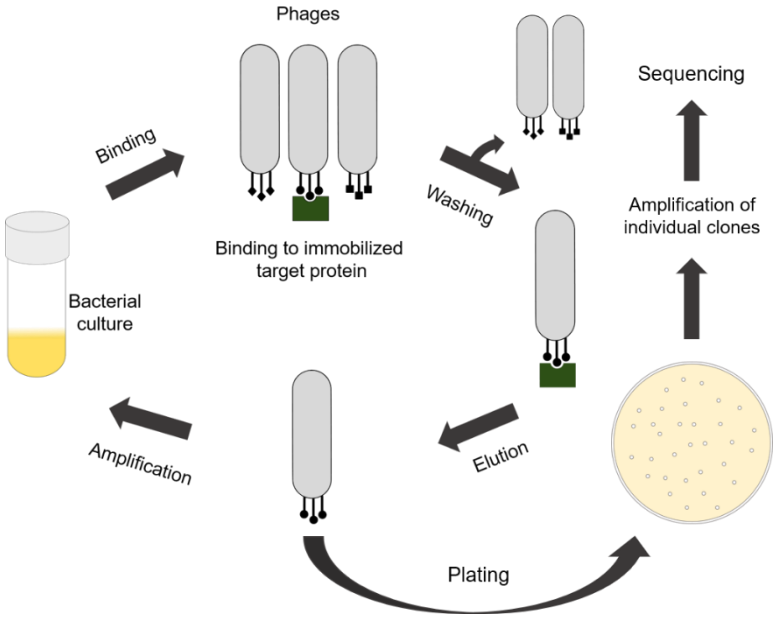


***Figure 1.*** **Phage display selection process**. Adapted from Huang *et. al*. [9] Phages displaying different peptides are added to an immobilized protein followed by a washing step, where non-binders will be discarded and binders will remain. Phages that display binders are then eluted and can be amplified in a bacterial culture to repeat the procedure. After a few rounds, individual clones are amplified and sequenced to obtain information about which peptides interact with the target protein.

---

## 1.4 Fluorescence polarization

To further validate and quantify interactions of SLiMs, several biophysical methods can be used and fluorescence polarization (FP) is a cheap and fast technique for this purpose. It is based on linking a fluorophore to the protein or peptide of interest, which will emit light of a certain degree of polarization when it is excited by polarized light. The degree of polarized light emission is inversely proportional to the rate of rotation of the molecule, which will change upon interaction with another molecule (*Figure 2*). When the fluorescently labeled small molecule (e.g. peptide) is free in solution, it will emit largely depolarized light after excitation by polarized light because it is rotating rapidly. If the labeled peptide interacts with a larger molecule (e.g. a protein), the emitted light will still be polarized after excitation by polarized light since the peptide rotates more slowly in this case.[10] Fluoresence polarization can be used for both direct measurement and displacement assays, where in both cases the concentration of fluorescent ligand is kept constant. For the direct measurement, protein concentration is varied which will result in a saturation of the binding when the concentration is increased incrementally. With the displacement assay, both a labeled and an unlabeled ligand are used, where the fluorescent one is allowed to interact with the protein and will be displaced by the unlabeled ligand. The unlabeled ligand concenctration is varied, while the fluorescent ligand and protein concentrations are kept constant. The signal will then decrease with increased

unlabeled peptide concentration and yield a sigmoidal curve when a logarithmic scale is used, as the unlabeled peptide will displace the labeled peptide from the protein molecule.[11]

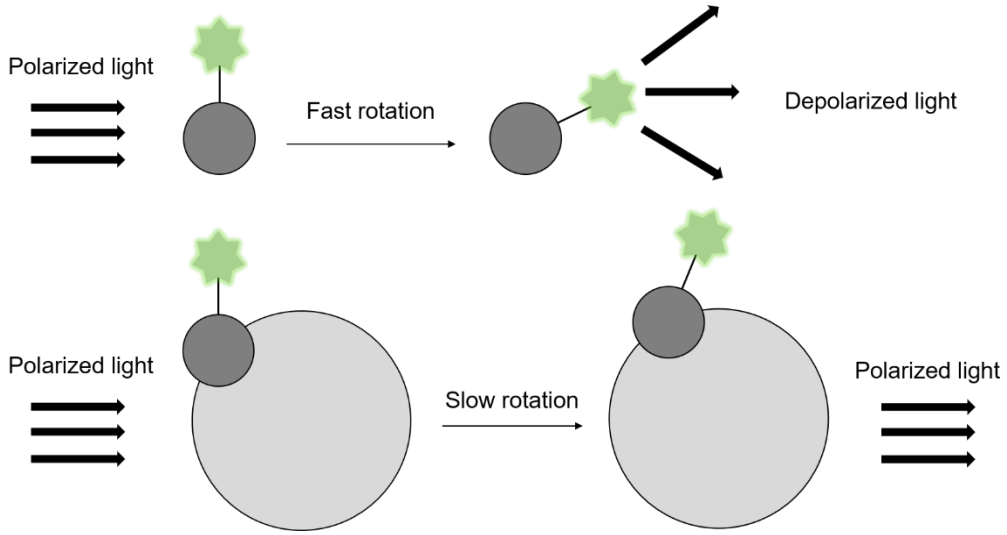


***Figure 2.*** **General principal for Fluorescence Polarization (FP).** Adapted from Moerke [10] and Rossi et. al[11]. When a fluorescently labeled molecule (e.g. a peptide) is free in solution, it will rotate fast and convert incoming polarized light to depolarized light. If it is interacting with a larger molecule (e.g. a protein) the rotation will be slower and the incoming polarized light will remain polarized. The ratio between depolarized light and polarized light can be measured and used to determine affinity of the interaction.

---

## 1.5 Isothermal titration calorimetry

Another method that can be used for validation and quantification of interactions of SLiMs is isothermal titration calorimetry (ITC). The principle of ITC is based on the fact that when two species, for example a protein and a ligand, interact with each other, heat will either be released or absorbed. This heat is proportional to the amount of ligand binding to the protein. For detection, the amount of power required to maintain a certain constant temperature between the reaction and reference cells is measured (*Figure 3*).[12] The raw data from ITC is a thermogram with energy vs time, which can be integrated into a sigmoidal curve (*Figure 3*), from which $K_d$, $\Delta H$ and $n$ (N) can be determined. To do the analysis, the binding model is assumed (e.g. one-site ($n = 1$) or several identical sites etc.), and the fitting process includes a nonlinear regression. The fitting is performed iteratively, where the parameters are initially estimated followed by a step-by-step adjustment of the fit to come closer to the actual data points.[13]

## 1.6 Circular dichroism

Circular dichroism (CD) is a technique that can be used to determine the secondary structure of biomolecules such as proteins. The obtained spectra from these measurements are of a certain characteristic, dependent on the geometry of the molecule. By comparing CD-spectra with known protein structures which have been obtained with X-ray crystallography, several empirical algorithms have been developed. The pattern of the spectra and at what wavelengths signals are obtained allows determination of overall secondary structure of a protein.[14] The technique can also be used for stability measurements, where temperature scans are performed to denature the protein which will lead to a loss of secondary structure and a concomitant loss

of preferential absorption of right or left polarized light. The temperature can then be lowered again to investigate whether the protein re-folds after being denatured, which would give the same or a similar signature spectra as before the temperature scan.[15] The principle is based on that plane polarized light (PPL) interacts with asymmetric molecules. Having two PPL waves perpendicular to each other, while having the same wavelength and amplitude, but ¼ wavelengths out of phase, will lead to a wave that oscillates in a circular manner. This wave can be right or left circularly polarized. When PPL interacts with proteins, which consists of asymmetric units, the right and left components will absorb differently. The signal will be dependent on the difference in absorbance between these components.[16]

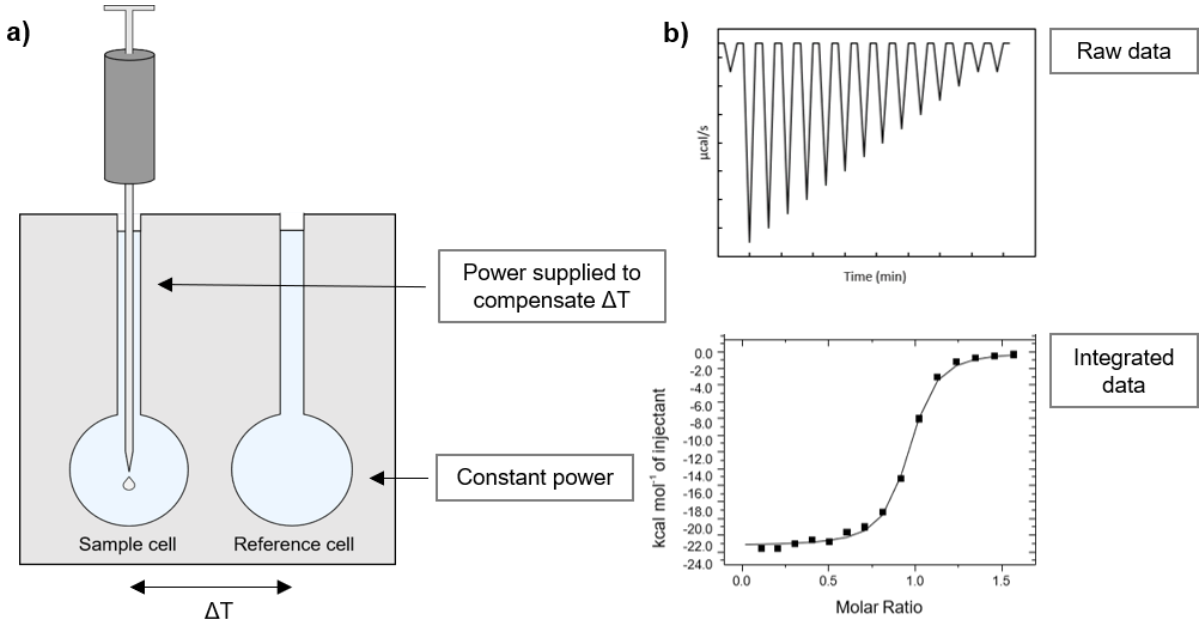


***Figure 3*. General setup of an ITC experiment.** Adapted from Leavitt *et. al* [12] and Freyer *et al.* [13]. a) In the sample cell, one of the species is present, while the other is titrated in. Upon interaction, heat will either be taken up or released and the power required to maintain the same temperature in the sample cell as for the reference cell is measured. b) Example of raw data and integrated data from which $K_d$ can be determined.

## 2 Project description

A novel ProP-PD library of disordered regions from viral proteomes known to infect humans was recently developed. The library has been screened towards several human protein domains and the aim of this project was to validate some of the results obtained from ProP-PD screening. The long-term aim for the project is to confirm the interactions in a biological context with cell-based *in vitro* experiments to further confirm potential viral hijacking, although this was not included in this project. In the big picture, this project is part of deepening the understanding of how viral proteins interact with human proteins to hijack host cell functions. It would be of interest to know if it is achieved through the proposed motifs, since the ProP-PD screening could then be established as a valid screening method for such interactions. For validation, FP and ITC were used to obtain affinity values ($K_d$/$K_i$) for interactions between proteins and peptides. The peptides were chosen on several criteria; previous knowledge, recognition motif,

overlapping peptides in screens, location in host, high counts in screens and promiscuity. Attempts were also made to crystallize one of the proteins with the chosen peptides.

## 2.1 Poly(A)-binding protein 1

The first protein to be studied was the C-terminal domain of Poly(A)-binding protein 1 (PABPC). PABPs are highly conserved proteins that interact with the polyadenylate (poly(A)) tail at the 3'-end of the majority of eukaryotic mRNAs. When PABPs were discovered, they were only thought to protect mRNAs from decay, while over the years they have been shown to have several different functions.[17] PABPs do in fact interact with poly(A) tails throughout the mRNAs lifetime, from production to degradation, for example determining the length of the mRNA by binding to it during poly(A) tail synthesis.[18,19] Another function is to facilitate the transport of mRNA to the cytoplasm, where it is sorted for decay or translation. There are several different PABP variants and the number of different ones vary between two to eight in different species.[17,20]

PABP1 is one of the most studied in this class of proteins. It has been shown to shuttle between the cytoplasm and nucleus.[21] PABP1 has four N-terminal RNA-recognition motifs (RRMs) that can interact with RNA. Its C-terminal domain consists of five α-helices which form a domain called PABPC or MLLE (*Figure 5*).[22] The name MLLE is used because this region contains the conserved motif Met-Leu-Leu-Glu. Between the N-terminal RRM motifs and the MLLE domain, there is a proline-rich unstructured linker (*Figure 4*).[17]

PABP1 is believed to be involved in translation initiation when it is bound to the poly(A) tail of the mRNA, which is one of two important structures for stimulation of protein synthesis.[23] The other one is the 5'-cap structure, to which the eukaryotic initiation factor 4G (eIF4G) can bind. eIF4G functions as a scaffold to bring several other factors into the cap-binding initiation complex (eIF4F) (*Figure 4*). PABP1 can interact with eIF4G through two of its four RRMs and to the release factor RF3 with its C-terminal domain, while being associated with the poly(A) tail. This will create a loop structure of the mRNA, which is believed to be important for translation initiation and ribosome recycling.[17]

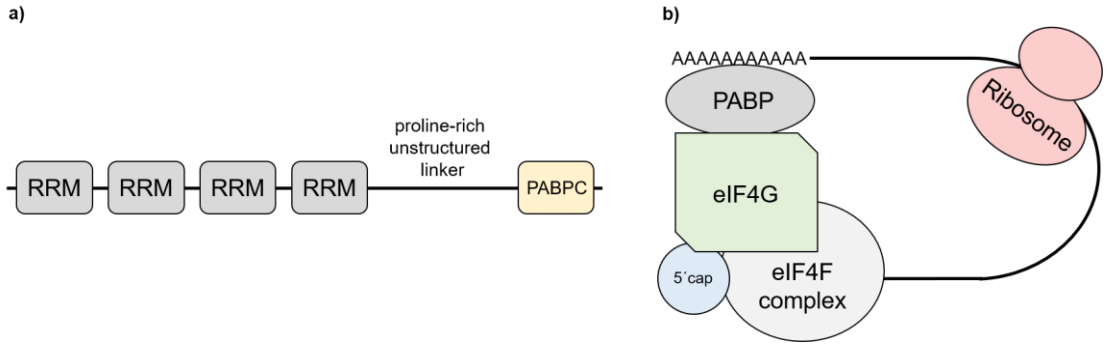


***Figure 4.*** **Domain organization of PABP and role of PABPC in translation initiation.** a) Domain organization of PABP, with four RRMs and the PABPC domain linked by a proline-rich unstructured linker. b) PABP is important for translation initiation and interacts with the poly(A) tail of mRNA and eIF4G, which acts like a scaffold to bring in factors that forms the eIF4F complex. Adapted from Eliseeva *et al* [17].

The PABPC domain is known to interact with PAM2 motifs of other proteins.[17] The PAM2 motif can for example be found in PABP interacting proteins PAIP1 and PAIP2 as well as in RF3.[24] The consensus motif of PAM2 is [LFP][NS][PICTAFL]xAxx[FY]x[PYLF], where the brackets denote that either of these amino acids are possible and x stands for any amino acid.[25] Several structures of PABPC alone and in complex with peptides from interacting proteins are available. As for the interaction between PABPC and PAIP2, two different structures suggest different binding conformations (*Figure 5*). The crystal structure (PDB: 3KUT) propose that there is no significant conformational change, while the NMR structure (PDB: 1JGN) shows a large conformational change in the PABPC domain upon interaction with the peptide. The peptides differ somewhat in length, with PAIP2 in the NMR structure being extended with three residues in the N-terminal and two residues in the C-terminal.[26,27] Other structures of PABPC in interaction with peptides shows more agreement with the crystal structure 3KUT, where no substantial conformational change occur. Since NMR allows more flexibility in the structure determination compared to the static x-ray crystallization, the NMR structure might be more consistent with what actually occurs in cells.

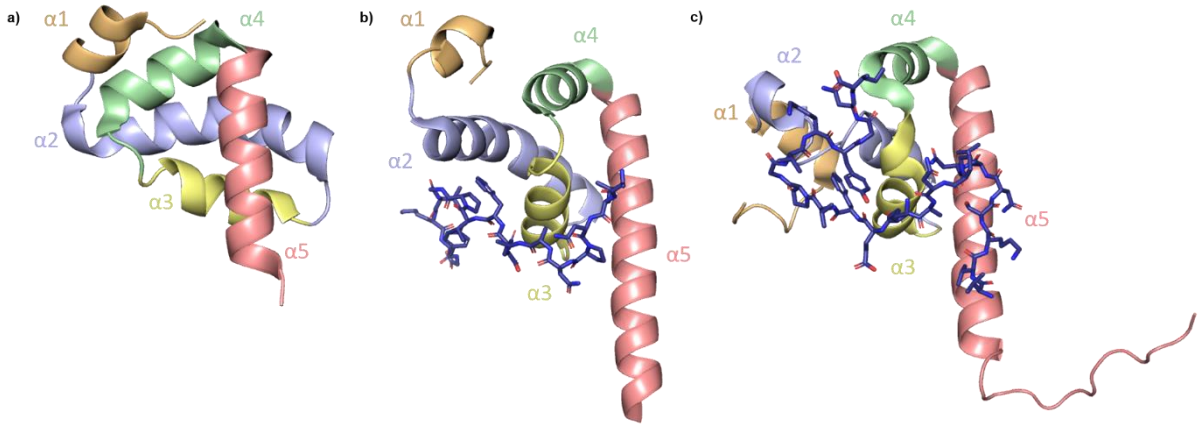


***Figure 5.* PABPC structures.** a) PABPC consists of five alpha-helices, represented by different colors and numbered from N- to C-terminal. b) Crystal structure of PABPC interacting with PAIP2, suggesting a smaller conformational change upon binding. c) NMR structure of PABPC interacting with PAIP2, suggesting a conformational change in PABPC involving helices α1 and α4. The α5 helix is longer in the two structures in (b) and (c) because longer constructs were used in the experiments. The figures were generated in PyMoL. PDB accession numbers: 3KUR (a), 3KUT (b) and 1JGN (c).

---

PABP1 has previously been shown to be a viral target, being cleaved by a variety of virally encoded proteases.[28] The most studied virus attacking PABP1 is poliovirus, an *Enterovirus* belonging to the *Picornaviridae* family. This virus is known to shut down the translational machinery of the host cell and was first thought to do so by proteolytic cleavage of eIF4G.[29] Further experiments showed that complete cleavage of eIF4G only resulted in partial host cell shut off, suggesting that some other process must also be involved in the translation inhibition. Later, it has been shown that PABP1 is also a target for proteases from several viruses from the *Picornaviridae* family, such as poliovirus.[28,30] The viral proteases cleave PABP1 at several sites in the linker region between the N- and C-terminal, which results in a weakened or abolished interaction between the PABP and mRNA. PABP1 has also been shown to be cleaved

by proteases from viruses within the *Caliciviridae* family and the *Retroviridae* family, to which HIV-1 and -2 belong.[28]

For PABPC, six peptides of interest (*Table 1*) from the viral proteome were chosen according to the criteria already mentioned, for further validation. The capsid protein from rubella virus does not contain the canonical PAM2 motif, but was included in the analysis because it has been previously reported to bind to PABPC and was found as a low count hit from the Pro-PD screen.[31] The affinities of these peptides for PABPC together with two peptides of known binders, PAIP1 and PAIP2 were determined with FP and ITC (*Table 1*).[27] For FP, a competition assay was performed. A FITC-labeled PAIP2 peptide was used to obtain a saturation curve, from which $K_d$ could be determined. The FITC-labeled PAIP2 was also used for displacement by the other peptides. The genome polyprotein from Tick-borne encephalitis virus (TBE) is called NS3 from now on, because the protein is cleaved into the protease NS3.[32]

*Table 1.* Peptides selected from ProP-PD for further validation of their interaction with PABP1. The bold residues in the respective sequence make up the recognition motif. If the peptides did not contain any residue that absorbs at 280 nm, a tyrosine was added at the end to ensure accurate concentration measurements.

| Protein | Virus | Peptide |
|---|---|---|
| Genome polyprotein (NS3) | Tick-borne encephalitis virus (TBE) | FEVKDGVYRI**FSP**GLL |
| Non-structural protein V (NSV) | Hendra | SAG**LNPAAVPFVP**KNQY |
| Non-structural protein V (NSV) | Nipah | TTG**LNPTAVPFTL**RNLY |
| Nucleoprotein (NP) | Human coronavirus 229E (HCoV) | SSQ**LNPAAPSW**IPPHA |
| Viral IRF4-like protein (ViRF4) | Human Herpesvirus 8 (HHV-8) | HP**LNPSALEFNP**SQTY |
| Capsid protein | Rubella virus | SWLWSEGEGAVFYRVD |
| Poly(A) binding protein 1 (PAIP1) | N/A | MSK**LSVNAPEFYP**SGY |
| Poly(A) binding protein 2 (PAIP2) | N/A | KSN**LNPNAKEFVP**GVKY |

## 2.2 Microtubule-associated proteins 1A/1B light chain 3C

The second protein of interest during this project was Microtubule-associated proteins 1A/1B light chain 3C (LC3C). LC3C is a ubiquitin-like protein that belongs to the autophagy-related protein 8 family which play a crucial role in autophagy processes. LC3C is one of three LC3 proteins in this family together with LC3A and LC3B.[33] Autophagy is a process through which the cells recycle damaged components, e.g. proteins, organelles and macromolecular complexes, by promoting their degradation. During the process, the autophagosome is formed, which has a double-membrane structure and contains cytoplasmic components. This structure will eventually fuse with the lysosome to form an autolysosome where its content will be degraded.[34] LC3C is synthesized in a soluble form which later associates with the membrane lipid phosphatidylethanolamine (PE) to become a part of the isolation membrane that will form the autophagosome. It can also attach to cargo receptors, which are brought into the autophagosome for degradation (*Figure 6*).[33]

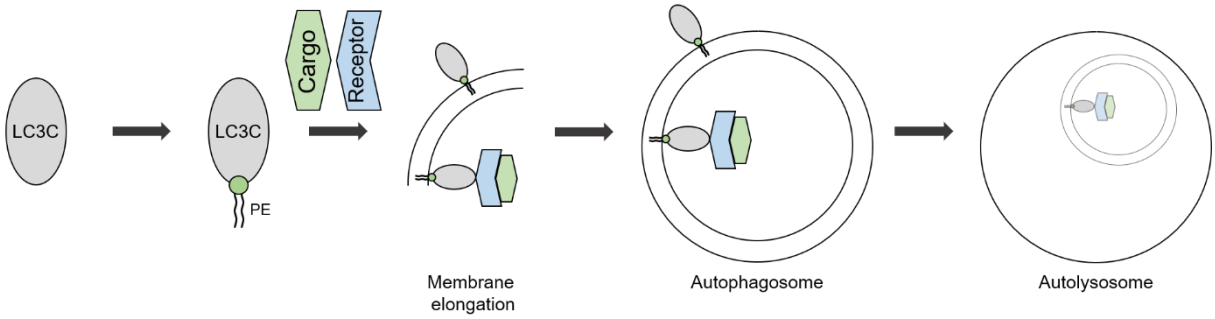


***Figure 6.*** **Autophagosome formation involving LC3C.** Adapted from Martens.[35] LC3C is synthesized in a soluble form that will associate with the membrane lipid PE and be able to become a part of the membrane that forms the autophagosome. The membrane will be elongated, finally being integrated into the autolysosome where the LC3C-PE complex can bring cargo for degradation.

---

Many known interactors of LC3C have the mutual recognition motif called LC3-interacting region (LIR). This motif starts with a negatively charged residue, either acidic or a phosphorylated serine or threonine. This is followed by a [WFY]xx[LIV] motif, in which the aromatic residue has been shown to be most important.[36] As for LC3 (not defined for LC3A, LC3B or LC3C) interactions with viral proteins, two proteins from Influenza A virus (IAV), Matrix-2 protein (M2) and Nucleoprotein (NP) have been proposed to be binders.[37] LC3 can relocate to form spots in cells which can be detected and indicates that the autophagy process is ongoing. M2 and NP has been shown to interact with LC3 within those spots during infection which seems to be related to autophagy, which in turn was shown to be related to viral replication regulation.[37]

For LC3C, six peptides (*Table 2*) from viral proteomes were chosen for further validation. Replication protein E1 was found as a hit for various strains of human papillomaviruses (HPV) in the Pro-PD screen. For FP, several FITC-labeled peptides were used to obtain saturation curves and evaluate which one to use for displacement experiments (*Table 3*).

*Table 2.* Peptides chosen for further validation of their interaction with LC3C. The bold residues in the sequence make the recognition motif. Similarly, to PABPC-binding peptides, tyrosine was added to the end of the peptide if it lacked residues absorbing at 280 nm.

| Protein | Virus | Peptide |
|---|---|---|
| Replication protein E1 (E1) | Human papillomavirus type 8 (HPV8) | KEGLS**EWCIL**EAECSD |
| Replication protein E1 (E1) | Human papillomavirus type 39 (HPV39) | GSGCNG**WFLV**QAIVDK |
| Matrix-2 protein (M2) | Influenza A virus (IAV) | YSAVDVD**DGHFVNI**VLE |
| Thymidylase kinase (TMK) | Variola virus | EGE**DIHWQII**SSEFEE |
| Glycoprotein | Rabies virus (RABV) | GDEAE**DFVEV**HLPDVHY |
| Protein UL56 | Human herpesvirus 2 (HHV-2) | GASAGQ**FVVI**DIDTPTY |

*Table 3.* FITC-labeled peptides used to obtain saturation curves and to evaluate which to use for displacement measurements for LC3C-peptide interactions. The p62 protein is a known human interactor, while the other three were found in Pro-PD screens for LC3A, LC3B and LC3C as potential novel targets (indicated in first column).

| Protein | Organism | Peptide |
|---|---|---|
| Ubiquitin binding protein p62 (p62) | Human | FITC-DDD**WTHL**SSK |
| Replication protein E1 (E1) (LC3A) | Human papillomavirus type 12 (HPV12) | FITC-KEGLSD**WCIL**EAECSD |
| Replication protein E1 (E1) (LC3B) | Human papillomavirus type 70 (HPV70) | FITC-GSGCNG**WFLV**QAIVDK |
| Replication protein E1 (E1) (LC3C) | Human papillomavirus type 29 (HPV29) | FITC-AERAGG**WFMV**EAIVDR |

# 3 Methods

## 3.1 Plasmid preparation and sequencing

Transformations for plasmid preparation were made with *E. coli* Dh5α competent cells. 30 µL cells were mixed with 1-2 µL of plasmid followed by incubation on ice for 20 minutes. The cells were heat shocked at 42 °C for 90 seconds followed by another incubation on ice for 2 min. 500 µL of LB-media (0.5 % (w/v) NaCl, 1 % (w/v) yeast extract, 1.6 % (w/v) peptone) was added to the cells followed by incubation at 37 °C for 30 minutes. 50-100 µL of culture was spread on a LB-plate (0.5 % (w/v) NaCl, 0.5 % (w/v) yeast extract, 1 % (w/v) tryptone, 1.5 % (w/v) agar) containing 100 µg/mL ampicillin or 50 µg/mL kanamycin. Plates were incubated at 37 °C overnight. One colony was inoculated into 3 mL LB-media containing 100 µg/mL ampicillin or 50 µg/mL kanamycin followed by incubation at 37 °C, 220 revolutions per minute overnight. The plasmid was purified with PureYield$^{TM}$ plasmid Miniprep kit (Promega, USA) and the concentration was measured with Scientific$^{TM}$ NanoDrop$^{TM}$ One$^{C}$ (Thermo Fisher Scientific, USA). Purified plasmids were sent for sequencing (Eurofins genomic, Germany) to ensure the construct contained the correct sequence.

## 3.2 Protein expression and purification

The plasmids were transformed into *E. coli* BL21 (DE3) competent cells, in the same way as for plasmid preparation. One colony was inoculated into 40-50 mL of LB-media containing 100 µg/mL ampicillin or 50 µg/mL kanamycin followed by incubation at 37 °C, 220 revolutions per minute, overnight. The following day, 4.5-5 mL of overnight culture was inoculated into each growing flask (up to 8 of them) with 500 mL of LB media containing 100 µg/mL ampicillin or 30 µg/mL kanamycin followed by incubation at 37 °C, 220 revolutions per minute until $OD_{600}$ reached 0.5-0.7.

### 3.2.1 PABPC expression and purification

For PABPC expression, the cultures were cooled down at 4 °C for 15 minutes before expression was induced with 0.4-0.5 mM IPTG followed by incubation at 18 °C, 220 revolutions per minute overnight. The cells were harvested by centrifugation at 4 °C, 3300 *g*, for 30-40 minutes. The pellets were washed by resuspending them in 30 mL cold 1x PBS (137 mM NaCl, 2.7 mM KCl, 10 mM $Na_2PO_4$, 1.8 mM $KH_2PO_4$) with shaking at 180 revolutions per minute and another centrifugation at 4 °C, 4400 *g*, 25 minutes was performed. The pellets were either stored in -20 °C or directly used for purification.

The pellet was resuspended in lysis buffer (50 mM Tris, 300 mM NaCl, 1 tablet of cOmplete™ protease inhibitor cocktail/50 mL buffer (Roche, Switzerland), 20 µg/mL RNase (Roche, Switzerland), 20 µg/mL DNase (Roche, Switzerland), 5 mM DTT, 2 mM $CaCl_2$, 4 mM $MgCl_2$, pH 7.7) followed by lysing of cells using a Cell Disruptor CF2 (Constant Systems Ltd, United Kingdom), by passing the sample twice through the cell disruptor. The lysate was centrifuged at 4 °C, 43000 *g* for 40 minutes and the supernatant was filtered with 20 µm cutoff filter and added to GSH-beads, pre-equilibrated with equilibration buffer (50 mM Tris, 300 mM NaCl, pH 7.7). The gel-slurries were incubated in conical tubes at 4 °C with over the top rotation for 2 h and then poured into empty plastic columns. The conical tubes were rinsed with wash buffer (50 mM Tris, 300 mM NaCl, 2 mM DTT, pH 7.7) to ensure all beads were transferred into the columns. The beads were then washed with wash buffer until absorbance at 280 nm dropped below 0.01, which was measured with NanoDrop.

When $A_{280}$ dropped below 0.01, 1 mL of elution buffer (50 mM Tris, 300 mM NaCl, 2 mM DTT, 10 mM glutathione (GSH), pH 7.7) was added to the beads, which were mixed by gentle pipetting up and down. The columns were left for 3 minutes to let the gel sediment followed by collection of the eluate. After the first collection, another 2 mL of elution buffer was added to the columns and the same procedure as for the first elution round was repeated, where all eluates were collected in the same tube. The absorbance at 280 nm was measured for some of the last drops from this elution to evaluate if another elution would be necessary ($A_{280} > 0.1$). If this was the case, 1 mL elution buffer was added to the beads, following the same collection procedure as previously. The protein concentration of the total elution pool was measured with $A_{280}$ to estimate the amount of thrombin (GE healthcare, USA) necessary for cleavage of the GST tag. Thrombin was added to the eluate followed by incubation at 4 °C overnight in the elution buffer or simultaneously with dialysis in equilibration butter at 4 °C overnight. For the dialysis, a 3.5 kDa MWCO standard regenerated cellulose dialysis tubing (Repligen, USA) was used.

SDS-PAGE analysis was performed on samples from the first purification procedure, where 20-40 µL samples were saved after each step, to evaluate the purification progress. To all samples, dye (1.4 M β-mercaptoethanol, 25 mM Tris-HCl pH 6.8, 10 % glycerol, 0.8 % SDS, 0.04 % (w/v) orange G) was added and the dilutions were made with Milli-Q water. The samples were boiled for 5-10 minutes on a heating block and then loaded onto an SDS-PAGE gel (either pre-casted, mini-PRTOEAN® TGX Stain-free™, 4-15 % or in-house made 15 %) together with a Mw-marker (PageRuler™ Prestained Protein Ladder 10 to 180 kDa (Thermo Fisher Scientific, USA) for in-house made gels and Precision Plus Protein Unstained Standards (BioRad, USA) for pre-casted gels). The gels were run at 180V until the samples reached the bottom of the gel for pre-cast gels, or at 65V until the samples reached the running gel followed by 150V until the samples reached the bottom for self-made gels. The gels were stained and then imaged with ChemiDoc MP (BioRad, USA). The gels were stained with staining solution (0.5 % Coomassie brilliant blue R-250, 10 % acetic acid, 25 % isopropanol), microwaved for 45 seconds, followed by incubation at room temperature with shaking for 30-60 minutes. After staining, destaining solution (10 % acetic acid, 10 % isopropanol) was added to the gel which was microwaved for 45 seconds followed by incubation at room temperature with shaking for 15 minutes with paper on the sides to increase the destaining rate. The destaining procedure was repeated one time.

The columns described below were connected to an ÄKTA explorer system and all buffers were degassed prior to use. If no dialysis was performed, the cleavage reaction sample was desalted on a desalting column (HiPrep$^{TM}$ 26/10 Desalting, GE Healthcare, USA). This was to remove DTT and other salts. The desalting column was equilibrated with equilibration buffer followed by addition of the cleavage reaction sample. The sample was then eluted with the same equilibration buffer, where sample collection was started when $A_{280}$ increased and stopped when it decreased. To remove GST and thrombin from the sample, reverse IMAC followed by a benzamidine column (HiTrap® Benzamidine Fast Flow, GE Healthcare, USA) were used. The GST protein contained a His-tag, which interacted with the IMAC $Ni^{2+}$-column and thrombin interacted with the benzamidine column. This means that the pure protein was in the flow-through from the columns. For PABPC, no tryptophans or tyrosines are present, which means it does not absorb at 280 nm and absorbance at 215 nm ($A_{215}$) was instead monitored for collection. The columns were equilibrated with equilibration buffer, the sample was loaded and PABPC (flow-through) was collected when $A_{215}$ increased. GST was eluted with imidazole buffer (30 mM Tris, 180 mM NaCl, 400 mM imidazole, pH 7.7) and collected with the same procedure as for the other collections. For buffer exchange, the desalting column was equilibrated with potassium phosphate buffer (50 mM $KH_2PO_4$, pH 7.4) followed by addition of the flow-through sample from reverse IMAC. The sample was eluted with the potassium phosphate buffer, for the protein to be stored in. The final protein sample was concentrated by centrifugation using Amicon® Ultra-15, 3000 MWCO tubes (Merck, Germany), divided into 55 µL aliquots, flash frozen in liquid nitrogen and stored in -80 °C.

An SDS-PAGE analysis was performed with samples collected throughout the purification procedure with the ÄKTA explorer system. The same procedure as for the first SDS-PAGE analysis was used. The purity of the final sample was also analyzed with mass spectrometry (MS) which was performed by another person in the group.

The PABPC concentration was measured using absorbance at 280 nm with NanoDrop in triplicates using a cuvette with 1 cm width. The concentration was measured at 205 nm and the protein was diluted 400-1000 times in MilliQ-water, using an extinction coefficient of 292230 $M^{-1}$ $cm^{-1}$.

### 3.2.2 LC3C expression and purification

Expression and purification of LC3C followed the same procedure as for the PABPC with a few exceptions. Lysing of the cells was performed by either Cell Disruptor CF2 or sonication (BANDELIN SONOPULS UW2200, Bandelin, Germany). For sonication, the samples were treated twice with 70 % cycles, 4 min and 30 % of the max power. The GST-tag had a different cleavage site, and PreScission Protease (in-house produced) was used in parallel with dialysis in equilibration buffer (50 mM Tris, 300 mM NaCl, pH 7.6) at 4 °C overnight. Reverse IMAC was used to remove both GST and the PreScission Protease and collection of the flow-through was performed. The concentration was measured in triplicates using absorbance with NanoDrop and an extinction coefficient of 8940 $M^{-1}$ $cm^{-1}$.

## 3.3 Circular dichroism

To confirm that the proteins were properly folded and to investigate the stability after purification, J-1500 circular dichroism spectrometer (Jasco, Japan) was used. The protein was diluted to 20 µM in potassium phosphate buffer for the measurement, which was performed at 25 °C, scanning wavelengths between 260-190 nm, with four accumulations. For stability

measurements, temperature scans were performed between 10-95 °C, increasing the temperature with a rate of 1 °C/min to denature the protein. For these measurements, a set wavelength dependent on the previous wavelength scan was used. For PABPC, 223 nm was used and for LC3C, 217 nm. They were chosen because of the strong signal at these wavelengths. After the temperature scan, the sample was cooled to 25 °C and another standard measurement was performed to investigate if the protein would re-fold after the denaturation.

### 3.4 Fluorescence polarization

Peptide solutions were prepared by dissolving lyophilized powder in potassium phosphate buffer with Tween (50 mM $KH_2PO_4$, 0.05 % Tween) followed by incubation with rotation, 4 °C, 2 h. Peptide concentrations for unlabeled peptides were measured with Nanodrop in triplicates at 280 nm. For FITC-labeled peptides, the concentrations were measured in triplicates using a cuvette with 1 cm width at 490 nm and 200 times dilution in Milli-Q water, using an extinction coefficient of 730000 $M^{-1}$ $cm^{-1}$.

To measure the affinity between the protein and different peptides, competitive binding/displacement method was used. The measurements were performed in black opaque 96-wells plates on SpectraMax® iD5 (Molecular Decvices, USA). For all plates, serial dilutions of peptides were performed with 2x dilution in each consecutive well, spanning from 2 to 4096 times dilution. A pilot plate was prepared with the protein and FITC-labeled peptide in the first row and peptides in the following rows. In the last row, there were 3 wells with buffer (50 mM $KH_2PO_4$, 0.05 % Tween), 3 wells with buffer and FITC-labeled peptide and 3 wells with protein and FITC-labeled peptide for references. For displacement measurements, several plates were prepared and the protein/peptides were added in triplicates. A solution of protein and FITC-labeled peptide was added to each peptide for the displacement.

For all measurements, a concentration of 5 nM for the probe was used. For PABPC, the highest protein concentration was 32.5 µM in the saturation experiment and was kept at 2 µM for the displacement experiments. Two saturation curves were obtained for PABPC, because the displacement for capsid rubella peptide was performed at a different timepoint and therefore a second saturation was made together with this experiment. For LC3C, the highest protein concentration was 36.9 µM in the saturation experiment and was kept at 2 µM for the displacement experiments. All LC3C-peptides were measured during the same round, where one saturation curve for the probe was obtained and used for determination of the affinity for the displacing peptides. The choice of concentration of the probe was based on the pilot measurement, where the raw signal needs to be sufficient for the experiments. For sufficient signal in displacement experiment, the protein and probe need to be in complex with each other. Therefore, a protein concentration approximately 2-3 times higher than $K_D$ was used. To calculate $K_i$ for the peptides, (*Equation 1*) was used.

$$K_i = \frac{[I]_{50}}{\frac{[L]_{50}}{K_d} + \frac{[P]_0}{K_d + 1}} \qquad \textit{(Equation 1)}$$

$[I]_{50}$ is the concentration of free competing peptide at 50 % of inhibition, $[L]_{50}$ is the concentration of free labeled peptide (denoted "probe" in this report) at 50 % inhibition, $[P]_0$ is the concentration of free protein at 0 % inhibition and $K_d$ is the dissociation constant for the protein and labeled ligand. These parameters can be computed through *Equation 2 – Equation 4*.

$$[P]_0^2 + (K_d + [L]_T) \times [P]_0 - [P]_T \quad \textit{(Equation 2)}$$

$[L]_T$ is the total concentration of free ligand and $[P]_T$ is the total concentration of free protein.

$$[L]_{50} = [L]_T - \frac{[PL]_0}{2} \quad \textit{(Equation 3)}$$

$[PL]_0$ is the concentration of protein-labeled peptide complex at 0 % inhibition, which is the case for the controls with only protein and labeled peptide.

$$[I]_{50} = IC_{50} - [P]_T + \frac{K_d \times [PL]_{50}}{[L]_{50}} + [PL]_{50} \quad \textit{(Equation 4)}$$

$IC_{50}$ is the concentration of competing peptide that is required to competitively dissociate 50 % of the labeled peptide and $[PL]_{50}$ is the concentration of protein-labeled peptide concentration at 50 % inhibition. Equations were in part adopted from Nikolovska-Coleska, Z. *et al.*[39]

## 3.5 Isothermal titration calorimetry

To further validate some of the affinities, ITC was used (MicroCal iTC200, GE Healthcare, USA). To ensure that the buffer would not influence the measurements, a joint dialysis was performed for the peptides and the protein in potassium phosphate buffer. Microdialysis bags were used, standard RC membrane, 100-200 µL or 100-500 µL with 100-500 MWCO (Repligen, USA). The dialysis was performed at 4 °C, with stirring overnight. Concentrations of peptides were measured in triplicates. Any dilutions of peptides and protein were made with the dialysis buffer. Since previous affinities were obtained with FP, these were used to estimate what concentrations would be necessary for ITC and samples were made accordingly (*Table 4*). The measurements were performed at 25 °C.

*Table 4.* Protein and peptide concentrations used in ITC measurements. The protein concentrations represent the concentration in the sample cell and the peptide concentrations represents the concentration in the syringe.

| Peptide | [PABPC] (µM) | [Peptide] (µM) | Peptide | [LC3C] (µM) | [Peptide] (µM) |
|---|---|---|---|---|---|
| PAIP2 HSa | 29.1 | 223 | E1 HPV-8 | 46.8 | 492 |
| NSV Hendra | 99.2 | 983 | M2 IAV | 67.9 | 724 |
| NP HCoV | 129 | 1214 | G RABV | 60.8 | 602 |
| Capsid Rubella | 20.5 | 182 | p62 HSa (FITC) | 9.88 | 91.3 |
| NSV Nipah | 97.5 | 1025 | E1 HPV-29 (FITC) | 13.8 | 123 |

## 3.6 Crystallization

Attempts were made to crystallize PABPC with selected peptides. Different conditions were screened with the MORPHEUS[38] kit and a customized Ammonium sulfate kit. For the ammonium sulfate kit, PABPC was crystallized without any peptide present, followed by seeding to increase the chances of obtaining crystals with PABPC and peptides and to increase crystal quality. One crystal from the PABPC setup was chosen and used for seeding. 2 µL of reservoir solution was added to the drop followed by crushing of the crystal with a crystal crusher followed by transferring the drop into an Eppendorf tube. Another 5 µL of reservoir solution was added to the drop which was also transferred into the Eppendorf tube. This

procedure was repeated until the volume of seed stock solution was 50 µL. The seeds were then vortexed to ensure they were crushed, with 30 second intervals for 2 min in total and incubation on ice for 1 min between every interval. For the setup with seeding, 1:1 mixes of protein and peptide were dialyzed in 2 L of 10 mM MES, 100 mM NaCl, pH 6.3 for 4 h before crystallization. Another seeding attempt was made with crystals obtained after the first seeding with crystals from the PABPC-ViRF4 HHV8 plate.

# 4 Results

## 4.1 Expression and purification of PABPC

The Phh1030-PABPC plasmid was transformed into *E. coli* Dh5α competent cells for plasmid preparation. The coding region of PABPC was verified by sequencing. The prepared plasmid was then transformed into *E. coli* BL21 (DE3) competent cells for protein expression. The cells grew as expected and expressed protein that was purified with GSH affinity resin followed by a reverse IMAC step. The concentration of the final PABPC sample was 1300 µM. For purity analysis, SDS-PAGE and mass spectrometry (MS) were used. The expected mass of PABPC is 9627.1 Da (*Figure 7*). The expression of PABPC was high and the purity analysis shows rather pure final samples, which was sufficient for the affinity experiments that were performed.

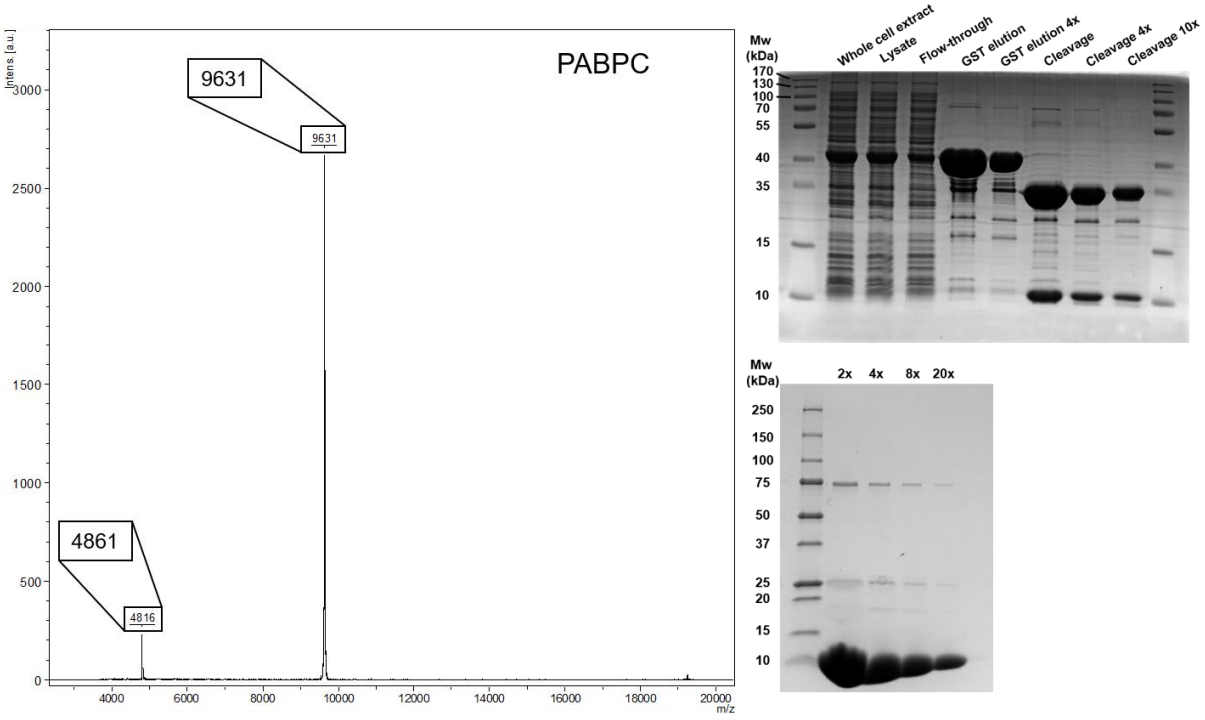


***Figure 7.*** **Purity analysis of PABPC.** MS-spectra which confirms the size and suggests that the protein has been succesfully purified with minimal contamintation. SDS-PAGE analysis, top gel visualizes the samples throughout the purification and the lower gel shows the final sample. For the bottom gel, the final protein sample was diluted 2, 4, 8 and 20 times. The gels confirm high expression of PABPC with the GST tag (whole cell extract, lysate) and rather pure final sample.

## 4.2 Expression and purification of LC3C

Two plasmids, pGEX-4T1-LC3C and pETM33-LC3C, containing the LC3C gene were transformed into *E. coli* Dh5α competent cells for plasmid preparation. The coding region of LC3C was verified by sequencing. The prepared plasmids were then transformed into *E. coli* BL21 (DE3) competent cells for protein expression. For pGEX-4T1-LC3C, the cells were growing with an expected rate until an $OD_{600}$ of 0.5-0.6, while after induction with IPTG, the growth rate was reduced. The cells were collected with centrifugation, where rather small pellets were obtained. After two failed purification attempts, the pETM33-LC3C plasmid was used instead. With this plasmid, the cells grew better and the protein was expressed and several pellets for purification were obtained. Several of the pellets were purified with GSH affinity resin followed by a reverse IMAC. The concentration of the final sample varied between 485-591 µM. For purity analysis, SDS-PAGE and (MS) were used. The expected mass of LC3C is 14850.36 Da (*Figure 8*).

Several attempts to express and purify LC3C pGEX-4T1 were made. The cells did not grow well; Very small pellets were obtained and the purifications from these pellets were unsuccessful. It should be mentioned that the purification attempts for this construct were performed by cleavage directly on the GSH beads which could have contributed to the low yield and failed attempts. During these attempts, LC3C also showed a tendency to aggregate and where much sample was lost during filtrations. However, since such small pellets were obtained and the purification attempts failed, the choice was made to try another plasmid, and it worked well. From the pETM33 plasmid, the expression and purification were more successful, although several rounds had to be performed to obtain enough protein for all measurements because of low expression level. LC3C from the pETM33 plasmid also had some tendency to precipitate, but less so compared to LC3C from pGEX-4T1. The purity analysis shows some contamination in all samples throughout the purification as well as in the final sample. Perhaps another purification step should have been added to ensure that no contamination would affect the affinity measurements. However, for the initial FP measurements, a protein from another purification than shown in *Figure 8* was used and this protein sample showed a somewhat less contaminated final sample and was deemed sufficiently pure for the measurements.

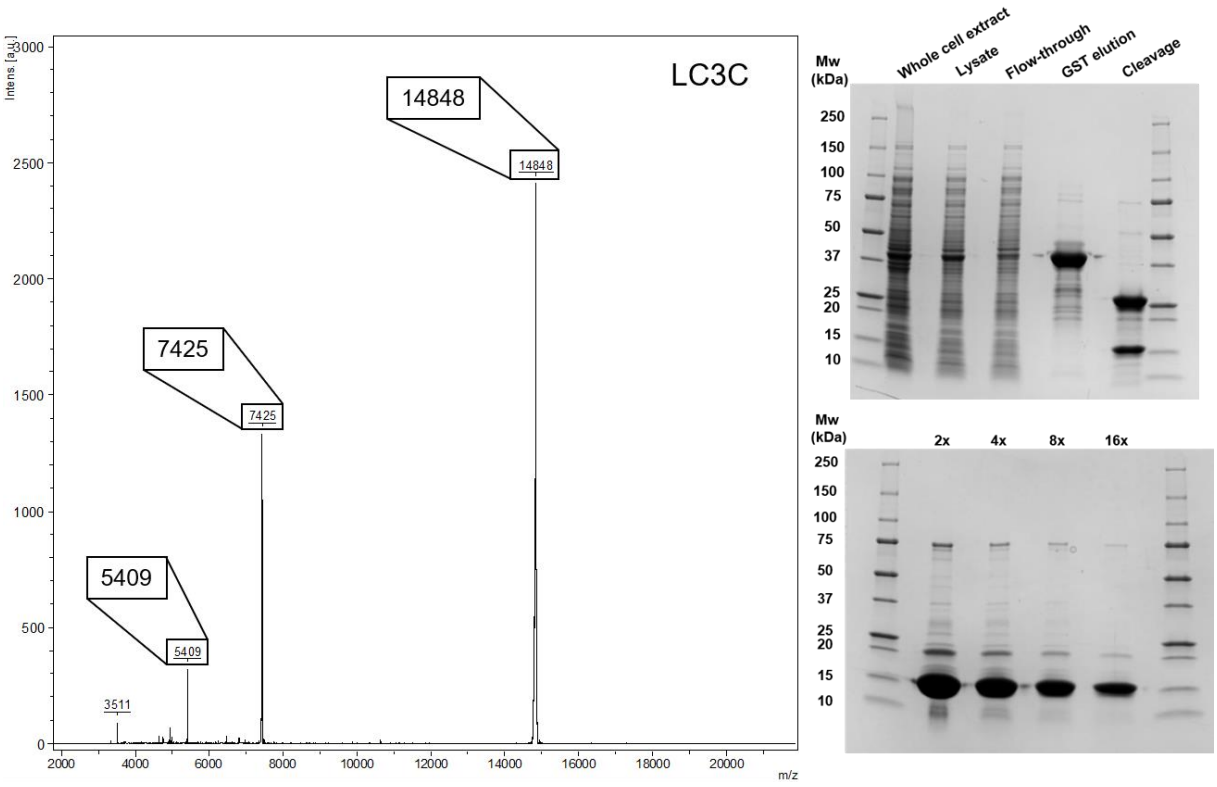


***Figure 8.*** **Purity analysis of LC3C**. MS-spectra which confirms the size and suggests that the protein has been succesfully purified. SDS-PAGE analysis, top gel visualizes the samples throughout the purification and the lower gel shows the final sample after concentration for two separate purifications. . For the bottom gel, the final protein sample was diluted 2, 4, 8 and 16 times The gels confirm expression of LC3C with the GST tag (Whole cell extract, Lysate) and final sample with some contamination.

## 4.3 Circular dichroism

CD measurements were performed to confirm the folding of the proteins in the conditions that were used for FP and ITC. All experiments with a set temperature were performed between 190-250 nm or 190-260 nm. A first measurement at 25 °C was performed to obtain a CD spectrum, followed by a temperature scan between 10-95 °C to study the stability of the protein. Another measurement at 95 °C was performed to obtain a spectrum for the unfolded protein followed by a final measurement at 25 °C after the measurement at 95 °C to study refolding of the protein. The melting temperatures for PABPC and LC3C have not been estimated in this study since it was not the focus or part of the aim

PABPC only consists of alpha helices, which is easily traced with CD since alpha helices give the specific shape of the CD spectrum with two distinct minima around 208 and 222 nm. This is clearly the case for the PABPC sample, which shows this characteristic pattern and confirms that the protein is folded. The temperature scan suggested that the protein is very stable, as a clear plateau is not achieved even going up to 95 °C, although a quite sharp signal change is obtained. It does also refold upon lowering the temperature down to 25 °C after heat denaturation (*Figure 9*).

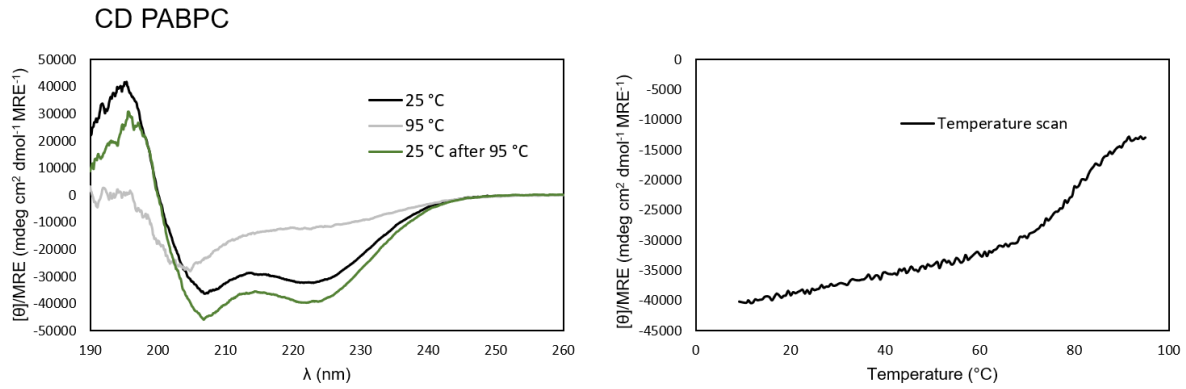


***Figure 9*. CD experiment for PABPC under different conditions**. a) Comparison of spectra for experiments at different temperatures, 25 °C, 95 °C after temperature scan and 25 °C after returning from 95 °C. b) Thermal denaturation of PABPC (10-95 °C) monitored at 223 nm.

---

LC3C consists of both alpha helices and beta sheets, which does not yield a curve that is as obvious as for PABPC. However, it is clear that something happens to the structure during the temperature denaturation, where a sharp change in signal and a plateau is observable. Also, the measurement performed at 25 °C after the heat denaturation is similar to the one at 95 °C, which means that the protein does not refold. These results are also consistent with the fact that LC3C seems to have a tendency to precipitate, which indicates that it is prone to aggregation (*Figure 10*).

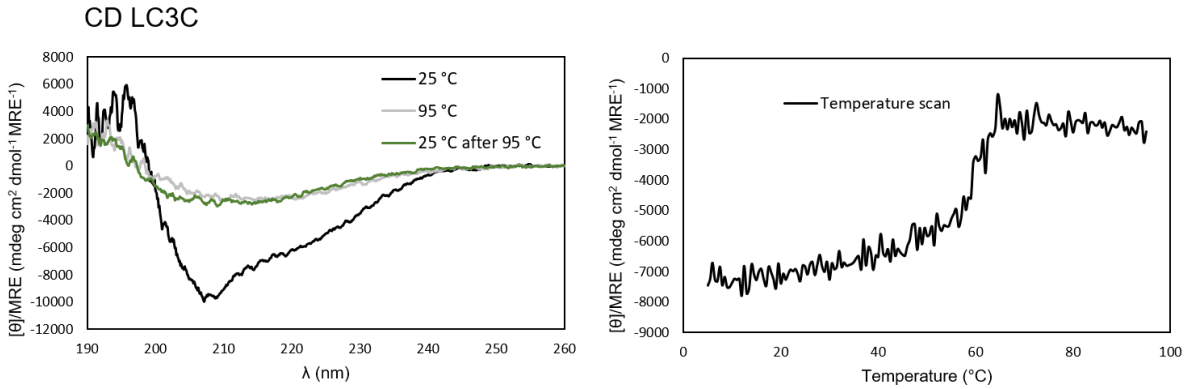


***Figure 10*. CD experiment for LC3C under different conditions**. a) Comparison of spectra for experiments at different temperatures, 25 °C, 95 °C after temperature scan and 25 °C after returning from 95 °C. b) Thermal denaturation of LC3C (10-95 °C) monitored at 217 nm.

---

## 4.4 Fluorescence polarization

Fluorescence polarization measurements were performed for PABPC and LC3C with several different peptides to validate their interaction and to estimate affinities. Saturation curves for the FITC-labeled peptides were produced to determine $K_d$ for the probe which was used to calculate $K_i$ of the displacing peptides. Every peptide was measured in triplicates, where each curve was used to calculate one $K_i$ value, from which the mean was taken and the standard error was calculated from the mean of these values. To calculate $K_i$ for each peptide, *Equation 1* was used.[39] All peptides that were measured with FP displaced the FITC-labeled peptide, since all curves show a decrease in signal with higher concentration of the displacing peptide

### 4.4.1 PABPC

Affinities between PABPC and peptides were determined with FP-monitored displacement experiments, using a FITC-labeled PAIP2 peptide as probe (*Figure 11, Table 5, Supplementary figure S1*). For PABPC, the affinities vary slightly between the different peptides. The strongest interaction is for NSV Hendra and Capsid Rubella which makes these interesting for further experiments. The weakest binders are NSV Nipah and NS3 TBE. NSV Nipah has a similar motif as NSV Hendra, which is why it would be interesting to perhaps measure the interaction with another technique, e.g. ITC to validate that the binding is weak compared to the other peptides. However, two FP measurements have been performed for this peptide and in both cases, the $K_i$ is around 20 µM, so it is possible that this interaction is indeed weaker than the interaction between PABPC and NSV Hendra. For NS3 TBE, the peptide is found in the C-terminal part of the protein, which is why it does not contain the complete motif and this could be an explanation to the weak affinity. For this protein, it would therefore be interesting to perform experiments with the full-length protein to investigate if the interaction would occur in the context of full-length proteins.

Two saturation experiments were performed for PABPC at two different timepoints, with the same probe. The affinities obtained from these measurements differ by 0.16 µM, which could be considered to be large since it is in the same magnitude as the actual $K_d$ values. This might give some uncertainty to the $K_i$ values that are calculated from the $K_d$ values. However, this is also the reason why two separate saturations were performed, both for their respective displacements. To validate that the $K_i$ values are reliable, the affinities should be determined with another technique, such as ITC, which has also been performed for some of the peptides.

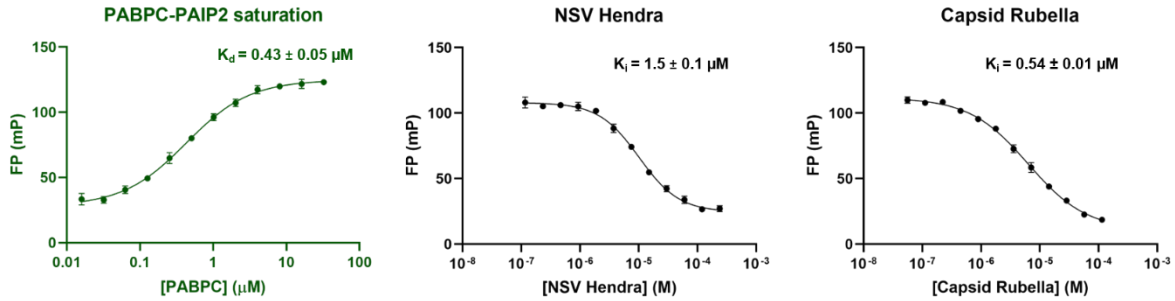


***Figure 11.*** **FP-monitored displacement measurements for PABPC**. Peptides NSV Hendra and Capsid Rubella, with respective $K_i$ values. The $K_d$ of the FITC-PAIP2/PABPC complex was determined in a saturation experiment (left panel). The saturation curve has a log-scale on the x-axis. For the Capsid Rubella peptide, this was measured at a different timepoint than the others and therefore another PABPC saturation was performed in parallel with this. Therefore, the $K_i$ for Capsid Rubella is calculated with the $K_d$ of FITC-PAIP2 from the second saturation (*Supplementary figure S1*). The $K_d$ differs between the saturations by 0.16 µM.

*Table 5.* Affinities between PABPC and the different peptides obtained from FP-monitored displacement measurement.

| Protein | Peptide | $K_i$ (µM) |
|---|---|---|
| NS3 TBE | FEVKDGVYRI**FSP**GLL | 28 ± 1.6 |
| NSV Hendra | SAG**LNPAAVPFVP**KNQY | 1.5 ± 0.1 |
| NSV Nipah | TTG**LNPTAVPFTL**RNLY | 22 ± 3.5 |
| NP HCoV 229E | SSQ**LNPAAPSW**IPPHA | 5.9 ± 0.5 |
| ViRF4 HHV-8 | HP**LNPSALEFNP**SQTY | 9.7 ± 2.1 |
| Capsid protein Rubella | SWLWSEGEGAVFYRVD | 0.54 ± 0.01 |
| PAIP1 | MSK**LSVNAPEFYP**SGY | 1.5 ± 0.06 |
| PAIP2 | KSN**LNPNAKEFVP**GVKY | 0.20 ± 0.06 |

### 4.4.2 LC3C

Pilot experiments of four different FITC-labeled peptides were performed, where saturation curves for each were obtained. Affinities between LC3C and the chosen peptides were determined with FP displacement measurements, using a FITC-labeled p62 peptide as probe (*Figure 12*, *Table 6*, *Supplementary figure S2*). The peptide derived from E1 HPV-39 could not be dissolved in the correct buffer due to its hydrophobic nature and was therefore excluded from the measurements.

The LC3C peptides all bind with quite strong affinities comparable to those of PABPC and its peptides. The strongest binders seem to be E1 proteins from different HPV strains. All FITC-labeled E1 peptides resulted in saturation curves and could have been used as probes for the displacement measurements. It would be interesting to perform further experiments on each peptide since they interact tightly with LC3C. However, the three peptides with strongest affinities were chosen for ITC measurements together with the FITC-labeled peptides E1 HPV-29 and p62. All peptides contain the known LIR motif, except for UL56, which does not contain the initial charged residue, which could be the explanation why this interaction with LC3C is weaker compared to the other peptides.

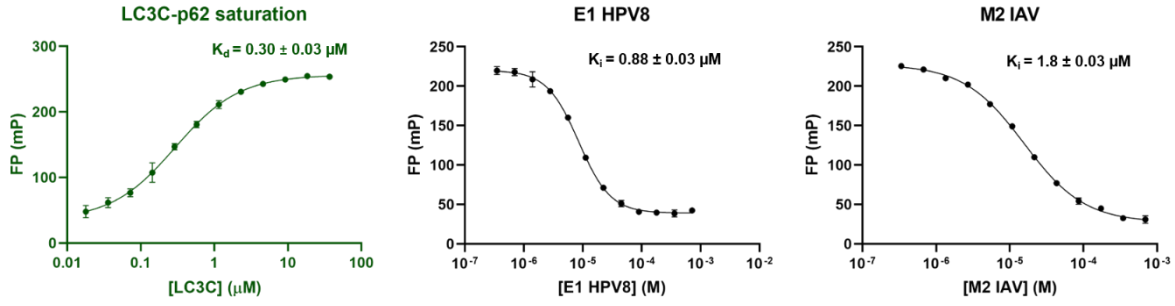


***Figure 12.*** **FP-monitored displacement measurements for LC3C.** Peptides E1 HPV8 and M2 IAV, with respective $K_i$ values. The $K_d$ of the FITC-p62/LC3C complex was determined in a saturation experiment (left panel). The saturation curve has a log-scale on the x-axis.

*Table 6*. Affinities between LC3C and the different peptides obtained from FP displacement measurement.

| Protein | Peptide | $K_i$ (μM) |
|---|---|---|
| E1 HPV8 | KEGLS**EWCIL**EAECSD | 0.88 ± 0.03 |
| M2 IAV | YSAVDVD**DGHFVNI**VLE | 1.8 ± 0.03 |
| TMK Variola | EGE**DIHWQII**SSEFEE | 3.7 ± 0.17 |
| Glycoprotein RABV | GDEAE**DFVEV**HLPDVHY | 1.0 ± 0.11 |
| UL56 HHV-2 | GASAGQ**FVVI**DIDTPTY | 6.3 ± 0.50 |

## 4.5 Isothermal titration calorimetry

ITC was performed for selected peptides for both PABPC and LC3C. The selection was based on the FP results and some of the strongest binders were used. For the experiments with PABPC, the raw data was integrated while for LC3C, the data could not be analyzed. The raw data was integrated and fitted iteratively, assuming a 1:1 binding, from which the stoichiometry (N), equilibrium association constant ($K_a$) and change in enthalpy (ΔH) were obtained. The affinities ($K_d$) could be calculated from the $K_a$ values.

### 4.5.1 PABPC

The ITC data (*Figure 13, Table 7*) obtained for PAIP2 and NSV Hendra yielded sigmoidal curves and affinities could be extracted. For NP HCoV, the fitted curve did not show a proper top and bottom making curve fitting and the resulting parameters more uncertain. PAIP2 showed the largest integrated heats of binding, while NP HCoV showed quite low values in comparison. For PAIP2, the affinity coincides between the two techniques, NSV Hendra differs somewhat, but the order of magnitude is the same. The peptide derived from NP HCoV differs quite a bit. The affinity between PABPC and PAIP2 has been previously reported to be 0.43 μM from ITC measurement on a peptide with three more residues on each side compared to the peptide used here.[26] This value is comparable to what is obtained for this ITC measurement since they have the same magnitude and because the peptides differ in length, the affinity could also differ. The ITC measurement for NP HCoV yielded small peaks as compared to the other peptides. The integrated data does not show a distinct top and bottom of the curve, which results in an uncertain $K_d$ value that does not correspond well to the $K_i$ from the FP measurement. ITC might not be an optimal technique for determination of weaker affinities. PAIP2, which binds with a quite strong affinity compared to the other peptides show similar results with both FP and ITC. For the peptides with weaker affinity, the difference between the techniques is larger and for the NP HCoV experiment, already high concentration of peptide was used and still the data was not useful. This suggests that it is difficult to measure weak affinities with ITC, since such high concentrations are necessary. This measurement should perhaps be performed with higher concentrations of protein and peptide if possible. Another peptide of interest to verify with ITC was Capsid Rubella, since it has been previously reported and does seem to bind to PABPC tighter than the other viral peptides. This peptide does not contain the obvious recognition motif, which makes it perhaps more interesting than the other ones.

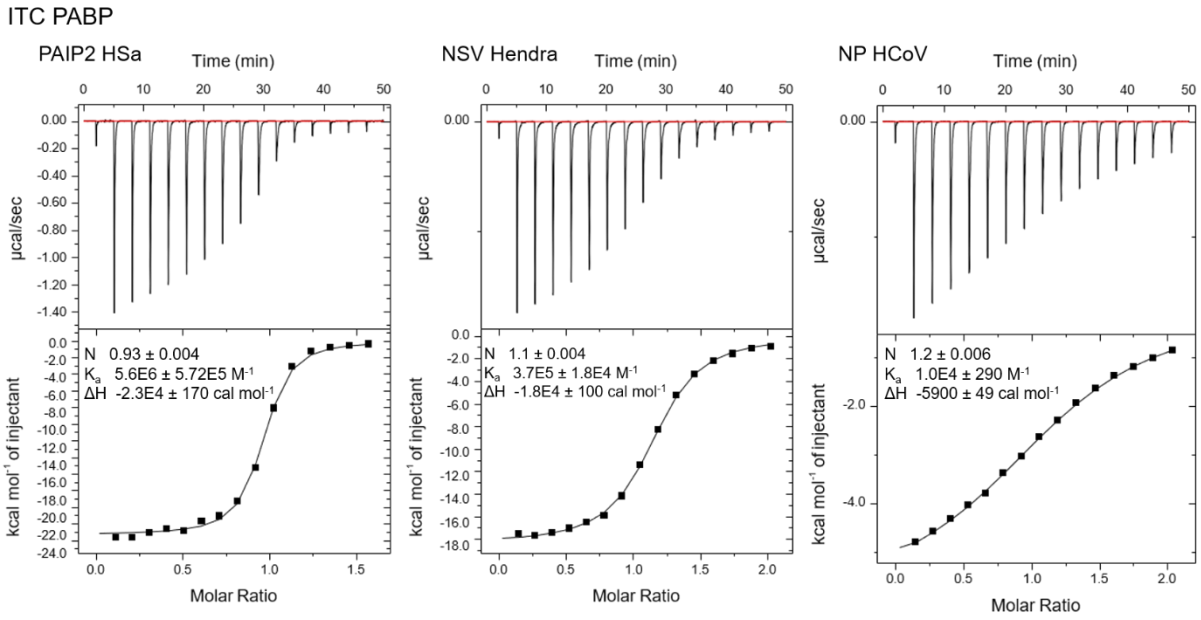


***Figure 13.*** **ITC experiments for PABPC.** Three of the peptides, PAIP2 HSa, NSV Hendra and NP HCoV, with respective obtained values. On the top of each panel is the raw data obtained from the measurement and below is the integrated data which has been fitted to yield values for stoichiometry (N), equilibrium association constant ($K_a$) and change in enthalpy (ΔH) were obtained.

*Table 7.* Affinities between PABPC and three of the peptides obtained from ITC measurements

| Protein | Peptide | $K_d$ (µM) |
|---|---|---|
| PAIP2 | KSN**LNPNAKEFVP**GVKY | 0.18 ± 0.02 |
| NSV Hendra | SAG**LNPAAVPFVP**KNQY | 2.7 ± 0.13 |
| NP HCoV | SSQ**LNPAAPSW**IPPHA | 96 ± 2.8 |

### 4.5.2 LC3C

Some of the LC3C peptides were also chosen for ITC experiments, including two of the FITC-labeled peptides to examine if the probe seemed to affect the measurement. The peptides definitely interacted with LC3C and it would have been of good interest to verify the affinities with another technique than FP. All protein and peptides mentioned in the method section, were dialyzed and diluted to suitable concentrations for ITC measurements, but the instrument stopped working properly. Attempts were anyway made to obtain data for the LC3C peptides E1 HPV-8 and M2 IAV, where the signal was too low and the curve could not be fitted and therefore no affinities were obtained. The other peptides were never measured due to struggling with consistency of the instrument. These measurements should however be performed in the future.

## 4.6 Crystallization of PABPC

The crystallization of PABPC was performed in several rounds, where the first round of screening mixes of PABPC and peptides with the MORPEHUS kit, led to no useful results. For the second round, PABP was screened alone with the customized ammonium sulfate kit, without any peptide present to produce apo-crystals that could potentially be used for seeding.

From this, several crystals were obtained, either as microcrystals or proper crystals. One crystal was selected for seeding to set up a new round of screen for protein-peptide mixes using the customized ammonium sulfate screen. From this round, several of the mixes showed microcrystals and the PABPC-ViRF4 HHV8 had grown several crystals, where one was chosen to be sent to the beam (*Figure 14*). The crystal was however lost during mounting and could not be sent to the beam. Therefore, another round of seeding was performed in the same manner with seeds produced from microcrystals and one proper crystal present in the PABPC-ViRF-HHV8 plate. From this round, another three crystals were obtained for the PABPC-ViRF-HHV8 plate and these will be mounted and sent to the beam. Microcrystals were also present in this plate, which could be used for further seeding experiments (*Figure 14*). The crystallization of PABPC with the peptides required optimization and some trial and error. The first round that was performed in the MORPHEUS screen, showed some crystals in the first row of conditions ($MgCl_2$ and $CaCl_2$)[38] which is prone to form salt crystals and therefore these were neither mounted or used for seeding. One problem with the crystallization of PABPC is that it does not contain any aromatic residues. If it would have, the crystals could be imaged with UV-light and this is one way of distinguishing salt crystals from protein crystals. Instead, to identify which crystals might be protein, the crystals can be touched with a loop and if it seems to be very fragile, this indicates that it is protein and not salt. However, if the crystal would break during mounting, it is not possible to further process it and obtain a crystal structure which makes this method less useful. Because of this reason, the crystals obtained for the MORPHEUS screen were regarded as salt even though this has not been proven. Instead, the approach became more directed towards screening conditions that has previously been used to crystallize PABPC.[26] One crystal that showed potential from the first Ammonium sulfate screen was crushed and used for seeding crystals with protein-peptide mixes. This crystal seemed to be fragile, which indicates that it was indeed a protein crystal, although this cannot be verified.

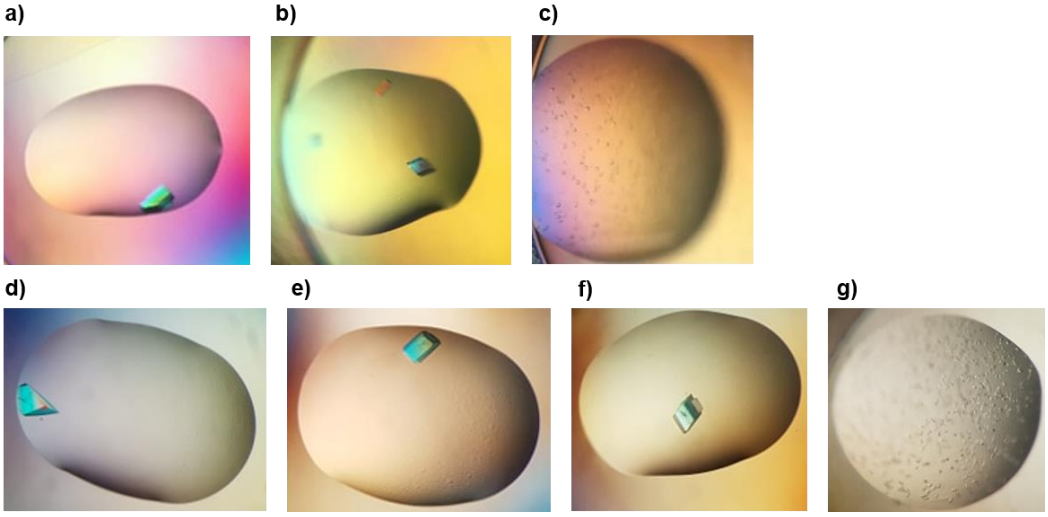


***Figure 14.*** **Crystals that were obtained after the seeding attempts in the PABPC-ViRF4 HHV8 plate.** The crystals shown in panel a-c were obtained during the first round of seeding and those in d-g were obtained in the second round. a) One crystal that was attempted to be mounted for beamline, but was lost during this step. b) Several crystals were obtained in this drop, where the one to the right was used to produce seeds. c) An example of microcrystals that were used for seeding. g) Microcrystals obtained in the second seeding round that could be used for another round of seeding.

# 5 Discussion

## 5.1 Fluorescence polarization

Comparing the affinities between the protein and the peptides, the LIR motif is quite short compared to the PAM2 motif. It has been shown that the interaction of LIR motifs is mostly dependent on the aromatic residue in the beginning of the motif. The PAM2 motif consists of many residues, where many of them seem to be important for binding. Because the motif contains so many residues, there are more factors playing in to the interaction, making the LC3C-peptide contacts clearer. One good example of this is again the NS3 TBE peptide, which is in the C-terminal end of the protein. Therefore, the entire motif is not present and neither can it be centered, which is preferred for these measurements. This peptide has a much weaker affinity compared to the other peptides, which is probably a result of the shortened motif. All the other PABPC peptides do have the entire motif, except for capsid Rubella, but there are also several options for each position, where some are probably favored over others. The most interesting peptide that has been picked up in this study is the Capsid Rubella peptide. It does not contain the PAM2 motif and still it binds with the strongest affinity of the viral peptides and also binds tighter than the known PAIP1. Further experiments for this peptide or the full-length protein are therefore desired.

## 5.2 Crystallization

For the crystallization, it would be interesting to obtain structures for several of the peptides to investigate if the binding mode is similar for all of them. Since the crystal structures for PABPC suggest that no significant conformational change occurs upon peptide binding, this is probably the most probable outcome for these peptides as well. However, since the PAIP2 peptide in complex with PABPC show different structures dependent on the technique that has been used, it is not clear which structure is closer to reality. Structures that are determined with NMR do allow more flexibility as compared to crystallization. Therefore, the NMR structure might show a conformational change that occurs but does not show in crystallization structures. Sometimes, conformational changes can also be obtained in crystallization structures and it would be of interest to find out what the case is for the viral peptides. If a structure of Capsid Rubella peptide in complex with PABPC could be obtained, it would be interesting to see if it binds in the same site as the other peptides and which residues are involved in the interaction. The peptide does contain some residues common to the motif and perhaps this will show a similar binding mode, suggesting that the motif can also be somewhat different. Also, for some of the positions in the PAM2 motif, there are a lot of options for residues and they do not all have common features. Perhaps therefore, Capsid Rubella do in fact interact in a similar manner to the other peptides, or a new motif has been evolved.

## 5.3 Viral hijacking

The choice of peptides to investigate was based on several criteria, with location in host and previous knowledge about the suggested interaction. The peptide derived from NS3 TBE originates from a serine protease, which could perhaps act like known proteases on PABPC by cleavage in the proline-rich linker, leading to shut-off of the translational machinery when combined with proteolytic cleavage of eIF4G. Perhaps it would be interesting to examine if the same peptide would bind to eIF4G and this could enable us to draw better conclusions. Since this interaction in particular is weak in comparison of the others, more experiments should be performed to investigate if the interaction would and could happen in a biological context. The other peptides do not have any known function linked to host translation mechanisms, but could

very well interact with PABPC to interrupt its interaction with the poly(A)-tail and thus also affect the translation. PABPC also seems to protect the poly(A)-tail during the mRNA's lifetime, which could lead to mRNA decay if this interaction is disturbed. This will in the big picture of course also affect the translation. The viruses might also be able to recruit PABP to efficiently translate the viral mRNA which could be another route of hijacking.

Viral hijacking linked to autophagy is a subject that has been somewhat investigated, but where more research is necessary. However, there is evidence of negative-strand RNA viruses that induces autophagy in the host cell for their own benefit.[40] Rabies virus and Influenza A virus are examples of negative-strand RNA viruses, from which the peptides G RABV and M2 IAV originate. The peptide M2 IAV has also been previously reported to interact with LC3C, although without quantitative data for the affinity and no further investigation of the motif.[37] These two peptides showed a low micromolar affinity to LC3C which suggest that these viruses could utilize autophagy of the host cell for their own replication. For example, the M2-LC3 interaction has been shown to redirect LC3-coupled membranes to the cell membrane which in turn is needed for IAV budding.[40] To induce autophagy might seem like a counterproductive measure, since it will degrade cell components and could actually be used by the host cell for protection against viruses. As for the M2-LC3 example, the viral proteins might induce autophagy-like processes, but that does not mean that the process will be carried out until the end, where components are degraded in the autolysosome. It could also induce autophagy to degrade cellular components which in turn can be used as building blocks to accelerate viral production. However, it has to make sure the cell stays intact to be able to survive in the host. The remaining peptides which originates from DNA viruses, does not have the same clear evidence of action linked to autophagy. Some DNA viruses have been shown to suppress autophagy to enhance their own survival. If this is achieved through interactions with LC3C or not seems to be unclear.[41] Since these peptides do show low micromolar affinity to LC3C, this might be something that could be further examined to conclude what their mode of action is and why they have seemed to evolve proteins with the LIR motif.

Since previously known binders have been picked up in the ProP-PD screen this shows that this method is indeed valid for identification of viral peptide – human protein interactions. Hopefully this technique together with the validation methods that have been proposed in this study can provide insight into how viruses utilize or disrupt cell processes for their own benefit.

# 6 Future outlooks

The first step in a continuation of the project would be to finish the ITC measurements for both the remaining PABPC peptides and the LC3C peptides. This could give some confirmation of the affinities and validate the $K_i$ values.

Another experiment that should be performed is GST pulldowns, with the full-length proteins to confirm the interactions in cells and in the context of full-length proteins. Since only short peptides are investigated in the techniques that have been used, the interactions should be confirmed also for the full-length proteins and it is also helpful to see if they would indeed interact in cells and not only as purified proteins. This would indicate some real biological function of the interaction which can be further investigated to determine how they affect the processes involving PABPC and LC3C. To confirm the motifs, it would also be of interest to perform targeted mutagenesis to strengthen the conclusion that these are the important residues for the binding.

The crystallization should be finished and potential crystals should be sent to a beam to obtain structures for PABPC with the different peptides. It would also be of interest to try and optimize the conditions for the Capsid Rubella peptide in particular, because the lack of the motif is difficult to explain without knowing the interface of the interaction.

# 7 Acknowledgements

First of all, I would like to thank prof. Per Jemth for giving me the opportunity to work on this project. I would like to thank him for creating an environment that promotes good learning opportunities.

I would also like to express my sincere gratitude to my lab supervisor Filip Mihalič, for all his help and understanding of my uncertainties during my project. You have been of great help in every aspect of the project and I am grateful to have been your first student.

To the rest of the Jemth group: thank you for everything, for making me feel safe and for all the favors and discussions when I needed it. A special thanks to Elin, for always being quick to help me with anything I have needed, either lab work or my annoying questions about equations and what not.

I would also like to thank Ivarsson group, with whom we have collaborated on this project. Thank you for everything that you have helped with and contributed to this work. A special thanks to Johanna, for helping me when needed and for always checking in on me. Also, thank you for all the laughs you have provided even when I felt I was struggling.

Lastly, I would like to thank my fellow programme students for all the talks during lunch and fika. Thank you for letting me ventilate my struggles and for the good times we have had. A special thanks to Mia, who has been by my side since our first year of the Bachelor's, thank you for being a great friend and for all your support in life and in this project.

## References


[1] Ivarsson, Y. & Jemth, P. (2019) Affinity and specificity of motif-based protein – protein interactions. *Curr. Opin. Struct. Biol.* **54**, 26–33.

[2] Tompa, P., Davey, N. E., Gibson, T. J. & Babu, M. M. (2014) A Million Peptide Motifs for the Molecular Biologist. *Mol. Cell* **55**, 161–169.

[3] Gibson, T. J., Davey, N. E. & Trave, G. (2011) How viruses hijack cell regulation. *Trends Biochem. Sci.* **36**, 159–169.

[4] Via, A., Uyar, B., Brun, C. & Zanzoni, A. (2015) How pathogens use linear motifs to perturb host cell networks. *Trends Biochem. Sci.* **40**, 36–48.

[5] Blikstad, C. & Ivarsson, Y. (2015) High-throughput methods for identification of protein-protein interactions involving short linear motifs. *Cell Commun. Signal.* **13**, 1–9.

[6] Smith, G. P. & Petrenko, V. A. (1997) Phage Display. *Chem. Rev.* **97**, 391–410.

[7] Azzazy, H. M. E. & Highsmith, W. E. (2002) Phage display technology: clinical applications and recent innovations. *Clin. Biochem.* **35**, 425–445.

[8] Davey, N. E. *et al.* (2017) Discovery of short linear motif-mediated interactions through phage display of intrinsically disordered regions of the human proteome. *FEBS J.* **284**, 485–498.

[9] Huang, H. & Sidhu, S. S. (2011) Studying Binding Specificities of Peptide Recognition Modules by High-Throughput Phage Display Selections. in *Network Biology: Methods and Applications* (eds. Cagney, G. & Emili, A.) 87–97 (Humana Press, 2011). doi:10.1007/978-1-61779-276-2_6.

[10] Moerke, N. J. (2009) Fluorescence Polarization(FP) Assays for Monitoring Peptide-Protein or Nucleic Acid–Protein Binding. *Curr. Protoc. Chem. Biol.* **1**, 1–15.

[11] Rossi, A. M. & Taylor, C. W. (2011) Analysis of protein-ligand interactions by fluorescence polarization. *Nat. Protoc.* **6**, 365–387.

[12] Leavitt, S. & Freire, E. (2001) Direct measurement of protein binding energetics by isothermal titration calorimetry. *Curr. Opin. Struct. Biol.* **11**, 560–566.

[13] Freyer, M. W. & Lewis, E. A. (2008) Isothermal Titration Calorimetry: Experimental Design, Data Analysis, and Probing Macromolecule/Ligand Binding and Kinetic Interactions. *Methods Cell Biol.* **84**, 79–113.

[14] Whitmore, L. & Wallace, B. A. (2008) Protein secondary structure analyses from circular dichroism spectroscopy: Methods and reference databases. *Biopolymers* **89**, 392–400.

[15] Greenfield, N. J. (2006) Using circular dichroism collected as a function of temperature to determine the thermodynamics of protein unfolding and binding interactions. *Nat. Protoc.* **1**, 2527–2535.

[16] Narhi, L. (2001) Circular dichroism. *BioPharm* **14**, 16–18.

[17] Eliseeva, I. A., Lyabin, D. N. & Ovchinnikov, L. P. (2013) Poly(A)-binding proteins: Structure, domain organization, and activity regulation. *Biochem.* **78**, 1377–1391.

[18] Brown, C. E. & Sachs, A. B. (1998) Poly ( A ) Tail Length Control in Saccharomyces

cerevisiae Occurs by Message-Specific Deadenylation. **18**, 6548–6559.

[19] Mangus, D. A., Amrani, N. & Jacobson, A. (1998) Pbp1p , a Factor Interacting with Saccharomyces cerevisiae Poly ( A ) -Binding Protein , Regulates Polyadenylation. **18**, 7383–7396.

[20] Mangus, D. A., Evans, M. C. & Jacobson, A. (2003) Poly(A)-binding proteins: Multifunctional scaffolds for the post-transcriptional control of gene expression. *Genome Biol.* **4**, 1–14.

[21] Afonina, E., Stauber, R. & Pavlakis, G. N. (1998) The human poly(A)-binding protein 1 shuttles between the nucleus and the cytoplasm. *J. Biol. Chem.* **273**, 13015–13021.

[22] Pieler, T. (1996) Xenopus Poly(A) Binding Protein: Functional Domains in RNA Binding and Protein – Protein Interaction. *J. Mol. Biol.* **256**, 20–30.

[23] Wells, S. E., Hillner, P. E., Vale, R. D. & Sachs, A. B. (1998) Circularization of mRNA by Eukaryotic Translation Initiation Factors. **2**, 135–140.

[24] Khaleghpour, K., Kahvejian, A., Ekiel, I., Kozlov, G. & Gehring, K. (2001) Structure and function of the C-terminal PABC domain of human poly(A)-binding protein. *Proc. Natl. Acad. Sci.* **98**, 4409–4413.

[25] Gouw, M. *et al.* (2018) The eukaryotic linear motif resource - 2018 update. *Nucleic Acids Res.* **46**, D428–D434.

[26] Kozlov, G., Ménade, M., Rosenauer, A., Nguyen, L. & Gehring, K. (2010) Molecular Determinants of PAM2 Recognition by the MLLE Domain of Poly(A)-Binding Protein. *J. Mol. Biol.* **397**, 397–407.

[27] Kozlov, G. *et al.* (2004) Structural basis of ligand recognition by PABC, a highly specific peptide-binding domain found ligase. *EMBO J.* **23**, 272–281.

[28] Smith, R. W. P. & Gray, N. K. (2010) Poly(A)-binding protein (PABP): a common viral target. *Biochem. J.* **426**, 1–11.

[29] Barry J. Lamphear, Riqiang Yan, F. Y., Debra Waters, H.-D. L., Hannes Klump, Ernst Kuechler, T. S. & Rhoads, and R. E. (1993) Mapping the Cleavage Site in Protein Synthesis Initiation Factor eIF-4y of the 2A Proteases from Human Coxsackievirus and Rhinoviru. *J. Biol. Chem.* **268**, 19200–19203.

[30] Kuyumcu-Martinez, N. M., Van Eden, M. E., Younan, P. & Lloyd, R. E. (2004) Cleavage of Poly(A)-Binding Protein by Poliovirus 3C Protease Inhibits Host Cell Translation: a Novel Mechanism for Host Translation Shutoff. *Mol. Cell. Biol.* **24**, 1779–1790.

[31] Ilkow, C. S., Mancinelli, V., Beatch, M. D. & Hobman, T. C. (2008) Rubella Virus Capsid Protein Interacts with Poly(A)-Binding Protein and Inhibits Translation. *J. Virol.* **82**, 4284–4294.

[32] Bateman, A. (2019) UniProt: A worldwide hub of protein knowledge. *Nucleic Acids Res.* **47**, D506–D515.

[33] Nguyen, T. N. *et al.* (2016) Atg8 family LC3/GAB ARAP proteins are crucial for autophagosome-lysosome fusion but not autophagosome formation during PINK1/Parkin mitophagy and starvation. *J. Cell Biol.* **215**, 857–874.

[34] Yu, L., Chen, Y. & Tooze, S. A. (2018) Autophagy pathway: Cellular and molecular mechanisms. *Autophagy* **14**, 207–215.

[35] Martens, S. (2016) No ATG8s, no problem? How LC3/GABARAP proteins contribute to autophagy. *J. Cell Biol.* **215**, 761–763.

[36] Wild, P., McEwan, D. G. & Dikic, I. (2014) The LC3 interactome at a glance. *J. Cell Sci.* **127**, 3–9.

[37] Wang, R. *et al.* (2018) Autophagy Promotes Replication of Influenza A Virus In Vitro . *J. Virol.* **93**, 1–17.

[38] Gorrec, F. (2015) The MORPHEUS II protein crystallization screen. *Struct. Biol. Commun.* **F71**, 831–837.

[39] Nikolovska-Coleska, Z. *et al.* (2004) Development and optimization of a binding assay for the XIAP BIR3 domain using fluorescence polarization. *Anal. Biochem.* **332**, 261–273.

[40] Wang, Y., Jiang, K., Zhang, Q., Meng, S. & Ding, C. (2018) Autophagy in Negative-Strand RNA Virus Infection. *Front. Microbiol.* **9**, 1–13.

[41] Kudchodkar, S. B. & Levine, B. (2009) Viruses and autophagy. *Rev. Med. Virol.* **19**, 359–378.

# 8 Supplementary material

## Tables

*Table S1*. Conditions from which crystals were obtained in the ammonium sulfate screen.

| Well | Salt | Ammonium sulfate (M) |
|---|---|---|
| A8 | 0.2 M ammonium iodide | 2.2 |
| C3 | 0.2 M potassium fluoride | 2.2 |
| D6 | 0.2 M sodium phosphate dibasic | 2.2 |
| F6 | 0.1 M BICINE pH 9 | 2.4 |
| H5 | - | 2.2 |
| H2 | Sodium Acetate trihydrate pH 4.6 | 2.2 |

## Figures

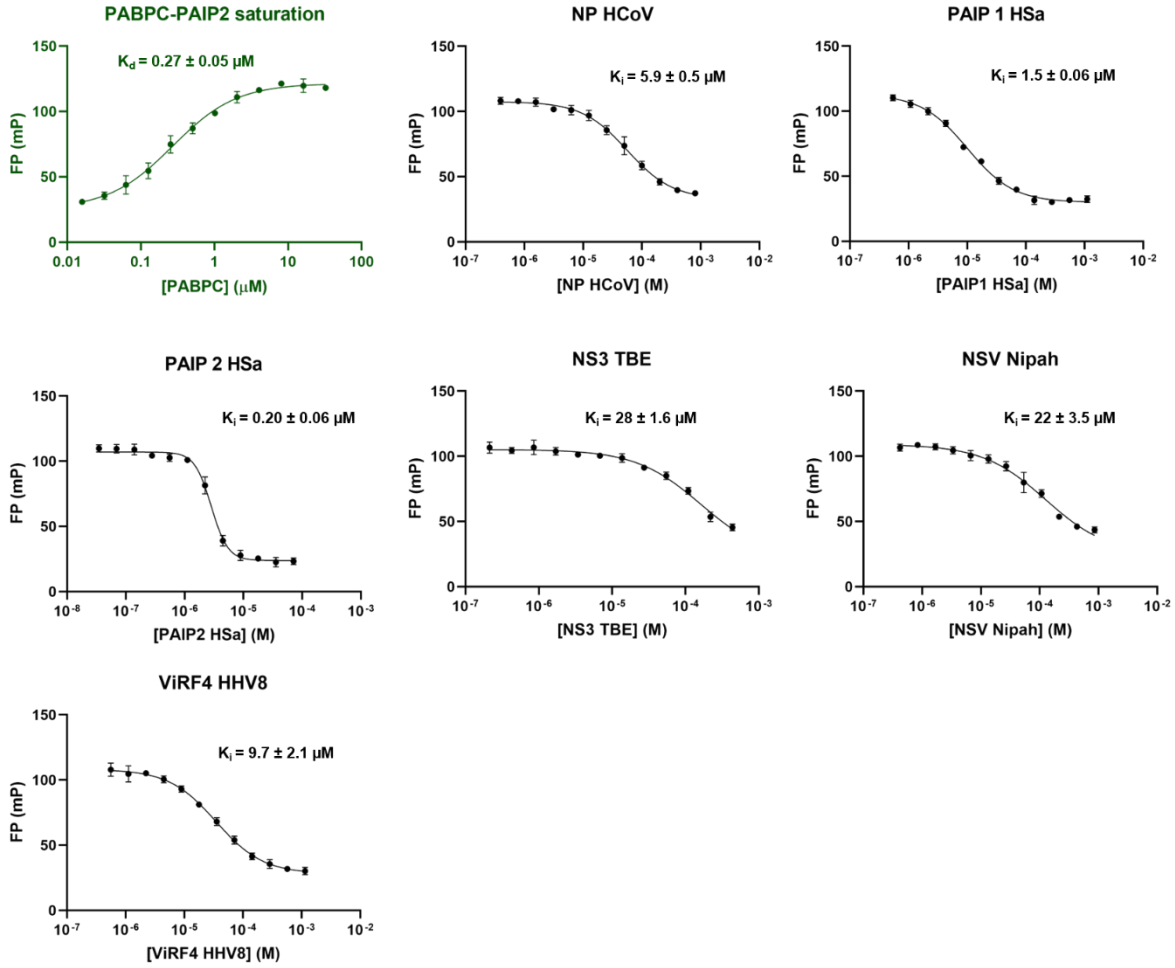


***Figure S1*. Remaining displacements curves for the PABPC-peptides.** All peptides except for NSV Hendra and Capsid Rubella, with respective $K_d$ and $K_i$ values. The $K_d$ of the FITC-PAIP2/PABPC complex was determined in a saturation experiment. The saturation curve in the left panel was obtained together with the experiment for the Capsid Rubella peptide. The saturation curve has a log-scale on the x-axis.

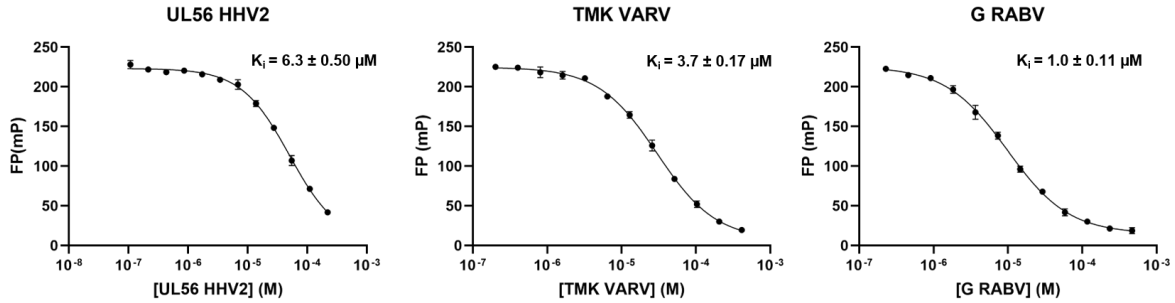


***Figure S2.*** **Remaining displacement curves for the LC3C-peptides.** Peptides UL56 HHV2, TMK VARV and G RABV, with respective $K_d$ and $K_i$ values. The $K_d$ of the FITC-p62/LC3C complex was determined in a saturation experiment (*Figure 12*).

---