Dissertation zur Erlangung des Doktorgrades
der Fakultät für Chemie und Pharmazie
der Ludwig-Maximilians-Universität München

The alternative cap-binding complex consisting of NCBP1 and NCBP3

ANNA GEBHARDT

AUS
MÜNCHEN, DEUTSCHLAND

2018

Erklärung

Diese Dissertation wurde im Sinne von § 7 der Promotionsordnung vom 28. November 2011 von Frau Professor Dr. Elena Conti betreut.

Eidesstattliche Versicherung

Diese Dissertation wurde eigenständig und ohne unerlaubte Hilfe erarbeitet.

München, den 02.05.2018

Anna Gebhardt

Dissertation eingereicht am	01.03.2018
1. Gutachter	Prof. Dr. Elena Conti
2. Gutachter	Prof. Dr. Andreas Pichlmair
Mündliche Prüfung am	11.04.2018

- IN MEMORY OF MY BELOVED MUM -

ABSTRACT

The cap structure of polymerase II transcripts is pivotal to protect RNAs from degradation and features a platform to recruit factors involved in RNA processing and export. Among the most essential binding partners of mRNA is the cap-binding complex (CBC) consisting of the nuclear cap-binding protein 1 (NCBP1) and the direct cap-binding subunit NCBP2. The CBC was long believed to be critically required for mRNA processing and export.

In the main project of my Ph.D. thesis, I showed that the cap-binding subunit NCBP2 is dispensable for export of bulk mRNA. I investigated the cellular repertoire of cap-binding proteins to identify candidates that could compensate for the loss of NCBP2. Using mass spectrometry, I identified the uncharacterized C17orf85 as a cap-binding protein that binds the RNA cap structure depending on the N7-methyl group of the guanosine. Furthermore, C17orf85 directly binds NCBP1 to assemble an alternative cap-binding complex, which has redundant function with the canonical CBC under steady-state conditions. Considering the function of C17orf85, we renamed the protein to nuclear cap-binding protein 3 (NCBP3). Using loss of function experiments, I demonstrated that the function of NCBP3 is required to mount proper antiviral responses and to prevent viral spread highlighting the fundamental role of the alternative CBC in the adaptation to environmental stimuli.

In a minor project of my thesis, I investigated the influence of serine S225 phosphorylation of the non-structural 5A (NS5A) protein of hepatitis C virus for interactions with cellular proteins. My collaboration partners further investigated the phosphorylation-dependent NS5A interaction partners for their requirement in viral replication.

Taken together, this thesis describes the molecular and functional characterization of the alternative cap-binding complex consisting of NCBP1 and NCBP3. Furthermore, it highlights the importance of post-translational modifications like phosphorylation for the interaction potential and the function of a protein.

CONTENTS

PREFACE

This thesis is written in a cumulative style. The first part of chapter one is a general introduction to the biological background of RNA biology, particular focusing on RNA export pathways. The second and third part of chapter one describes the discrimination of self and non-self RNA , the viral manipulation of mRNA export as well as the importance of viral RNA-binding proteins for the success of viral replication. Section 1.2.1 of the introduction was published in *Journal of Cytokine and Interferon Research*:

Gebhardt, A. *, Laudenbach, B.T. *, and Pichlmair, A. (2017). Discrimination of Self and Non-Self Ribonucleic Acids. *J. Interf. Cytokine Res.* 37, 184–197. **these authors contributed equally*

The second chapter includes my results in form of three manuscripts already published or prepared for submission:

Gebhardt, A.*, Habjan, M.*, Benda, C., Meiler, A., Haas, D.A., Hein, M.Y., Mann, A., Mann, M., Habermann, B., and Pichlmair, A. (2015). mRNA export through an additional cap-binding complex consisting of NCBP1 and NCBP3. *Nat. Commun.* 6, 8192. **these authors contributed equally*

Goonawardane, N., **Gebhardt, A.**, Bartlett, C., Pichlmair, A., and Harris, M. (2017). Phosphorylation of serine 225 in hepatitis C virus NS5A regulates protein-protein interactions. *J. Virol.* JVI.00805-17.

Gebhardt, A., Schnepf, D., Moser, M., Meiler, A., Michaudel, C., Mackowiak, C., Sedda, D., Stukalov, A., Reinert, L., Paludan, S.R., Ryffel, B., Stäheli, P., and Pichlmair, A. (2018). The alternative cap-binding complex is required for antiviral response *in vivo*. (prepared for submission)

The last chapter features concluding remarks and a brief outlook.

All publications are reprinted with permission and copyright of the publications belongs to the publishers.

1 INTRODUCTION

1.1 RNA PROCESSING AND NUCLEAR EXPORT

1.1.1 Nucleocytoplasmic compartmentalization

In eukaryotic cells, evolutional pressure has led to the development of a physical separation between transcription and translation. This physical barrier, namely the nuclear envelope (NE) consisting of a double membrane, separates the nucleus that is the site of DNA storage and RNA synthesis from the cytoplasm where protein synthesis is executed. The development of this compartmentalization also triggered the establishment of transport machineries to passage macromolecules between the nucleus and the cytoplasm. One part of this machinery is the nuclear pore complex (NPC) that is embedded in the NE and controls the export and import of macromolecules together with a family of conserved transport receptors known as karyopherins or importin-ß-like transport receptors. The NPC is built up of approximately 30 different proteins called nucleoporins (Nups) that build up a 60-125 MDa structure [1]. This permits the passage of proteins, RNAs and other soluble compounds. While small molecules (less than 40 kDa) can pass the NPC by diffusion, larger complexes like messenger (m) RNA need energy to shuttle through the pore [2]. Transport receptors selectively recognize their cargos by binding short peptide signals named nuclear localization signal (NLS) and nuclear export signal (NES). Additionally, specific receptors can detect nucleotide motifs in RNAs enabling their export to the cytoplasm. Through the recognition of these signals, proteins that function in the nucleus such as histones or transcription factors can be imported into the nucleus, whereas, ribonucleic acids (RNAs) that mostly carry out their function or a processed in the cytoplasm are transported to the cytoplasm. Researchers demonstrated that one million macromolecules per minute are transported between the two compartments of a living cell [3].

Even though there are different transport pathways, they all follow the same principles. The common feature of transport is the binding of the small GTPase Ran, which can switch between a GTP- and GDP-bound state (Figure 1). Ran exists in a GTP-bound form in the nucleus and a GDP-bound form in the cytoplasm. This gradient is generated by the existence of the Ran-guanosine-nucleotide exchange factor (RanGEF), which localizes to the nucleus, and the Ran-GTPase activating protein (RanGAP), which functions in the cytoplasm. The asymmetric distribution of RanGTP and -GDP drives the nucleocytoplasmic transport direction of the transport receptors (Figure 1).

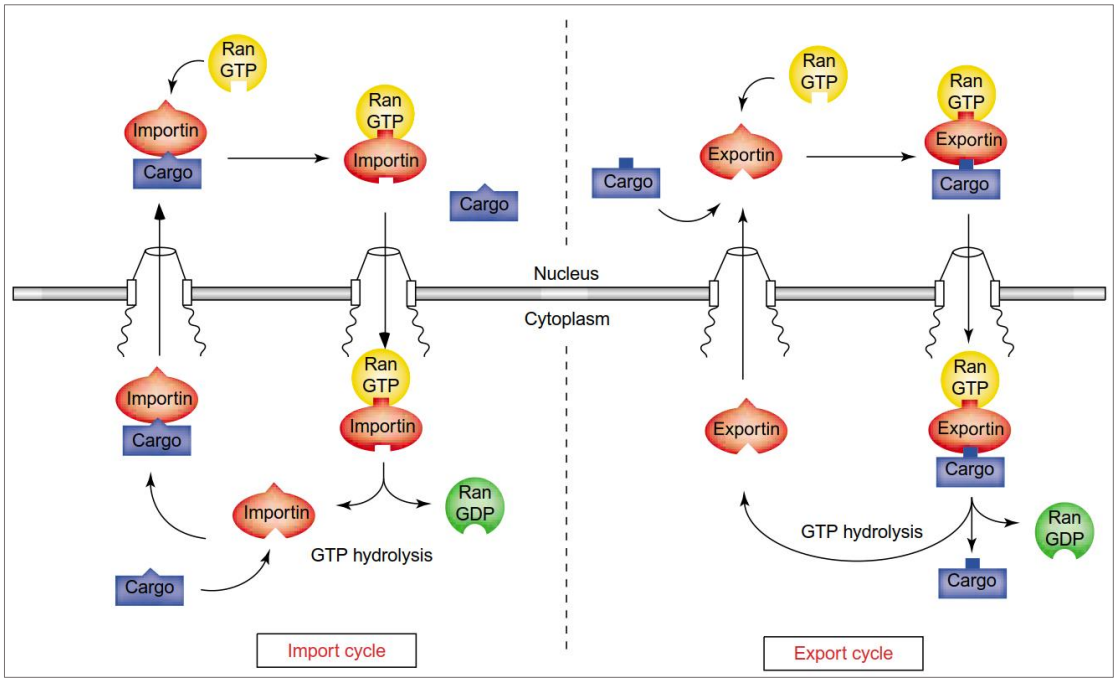


Figure 1: Nuclear import and export of macromolecules.

The left site represents the import cycle of importin ß-like transport receptor. Cargos are bound by importins, which mediate translocation of the cargo through the nuclear pore. In the nucleus, importins associate with RanGTP resulting in the release of the cargo. RanGTP-bound importins are recycled back to the cytoplasm where GTP hydrolysis drives the release of importins. In turn, importins can bind new cargos and the cycle restarts. The export cycle (right) is characterized by the binding of exportin to its cargo in the nucleus in a RanGTP-dependent manner. The bound RanGTP-exportin-cargo complex is shuttled to the cytoplasm interacting with nuclear pore subunits. On the cytoplasmic site, GTP hydrolyses results in release of the cargo and free exportin is recycled back to nucleus. Schematic and figure legend were adapted from reference [4].

Importins, which are karyopherins that bind cargos in the cytoplasm and import them into the nucleus, bind RanGTP in the nucleus, which drives the release of the cargos. In its RanGTP bound state, importin is transported back to the cytoplasm, where GTP hydrolysis results in RanGDP and disposal of importin, which in return can interact with a new cargo. In contrast, karyopherins that passage cargos from the nucleus to the cytoplasm are called exportins. These exportins bind their cargo in a RanGTP bound form in the nucleus and transport them to the cytoplasm, where GTP hydrolysis drives the release of the cargo and unbound exportin can be located back to the nucleus to bind new cargo. All molecules, including proteins and most RNAs (transfer (t), micro (mi), small nuclear (sn) and ribosomal (r) RNA) that are larger than 40 kDa and need to bypass the barrier of the nuclear envelope, follow these general principles. However, one exception is the export of mRNA. The export of this macromolecule is mechanistically different from the others since it uses a karyopherin-unrelated transport receptor and is not dependent on the RanGTP/GDP gradient. The export of mRNA requires the fine-tuned combination of additional transport adaptor and release factors, which makes this export process more sophisticated and controlled. The following section will provide an overview of the major RNA export pathways and their differences.

1.1.2 Insights from small RNAs: tRNA, miRNA, rRNA and snRNA

Commonly, all RNA classes are processed in the nucleus to a mature and thereby export-competent molecule. The classes of small RNAs, including tRNAs, miRNAs, rRNA and snRNAs, follow the general principles of nuclear to cytoplasmic transport dependent on karyopherins and the selective and directional Ran cycle [5]. In this section, the differences in export routes of selective RNAs are introduced.

1.1.2.1 Processing and export of tRNAs

Transfer RNAs are loaded with amino acids (AAs) and are required in the cytoplasm during translation to provide the ribosomal complex with AAs for peptide chain elongation. About 40 different tRNAs (in eukaryotic cells) need to be transported from their place of synthesis in the nucleus to the cytoplasm, where they function in translation. After synthesis by polymerase III, tRNAs are processed by multiple maturation events including trimming of the 3` and 5`trailer sequence, modification of nucleotides, addition of the 3`CCA nucleotide which serves as the AA acceptor stem and removal of introns (when present) (Figure 2a) [6,7]. It is believed that maturation steps occur in the nucleus and are essential for proper export [8,9], which was also shown through the finding that exportin-t (XPOT) only weakly interacts with tRNAs lacking the trimmed 5` or 3`end or required modifications [10–12]. Surprisingly, studies injecting XPOT and intron-containing tRNA in *Xenopus* oocytes showed that XPOT can bind and export tRNAs containing introns [10,12]. However, since intron removal was shown to occur before 5`end processing, in physiological conditions export of in-mature tRNAs should not be likely [13]. Interestingly, cumulative studies carried out in yeast suggest that tRNA can be re-imported to the nucleus in an event called retrograde process [14–21]. It was implicated that the re-import of tRNAs could be an intra-nuclear quality control mechanism or important for tRNA modifications [22,23] and is accelerated during starved cellular conditions [14,18,19,23,24]. Conflicting studies not only in different yeast strains but also in vertebrates are currently challenging if tRNAs are retrograded to the nucleus in response to nutrient stress, which was to date only clearly shown in *Saccharomyces cerevisiae* [24–28]. Following maturation in the nucleus, tRNAs ensue the general export principles using a karyopherin-ß family transport receptor and RanGTP (Figure 2a). The classical tRNA export route includes XPOT as transport receptor that preferentially binds tRNA in the nucleus and is favored by the association of RanGTP [11,29]. It is believed that XPOT interacts with Nups and translocates the trimeric XPOT-tRNA-RanGTP complex through the nuclear pore. Subsequently in the cytoplasm, RanGAP stimulates the hydrolysis of GTP and consequently GDP is produced and the tRNA is released from its receptor, which in turn is recycled back to the nucleus. In addition to XPOT, it

has also been shown that exportin-5 (XPO5), which has its main function in miRNA export (see chapter 1.1.2.2), can also associate with tRNAs and export them in a similar manner to XPOT [30,31]. Furthermore, studies on the yeast XPOT orthologue LOS1 showed that LOS1 is not essential for viability and that certain tRNA species are not impaired in export by the depletion of LOS1 [32]. This suggests that some tRNAs are translocated to the cytoplasm independent of LOS1 indicating that alternative tRNA export receptors or even novel pathways are still unexplored.

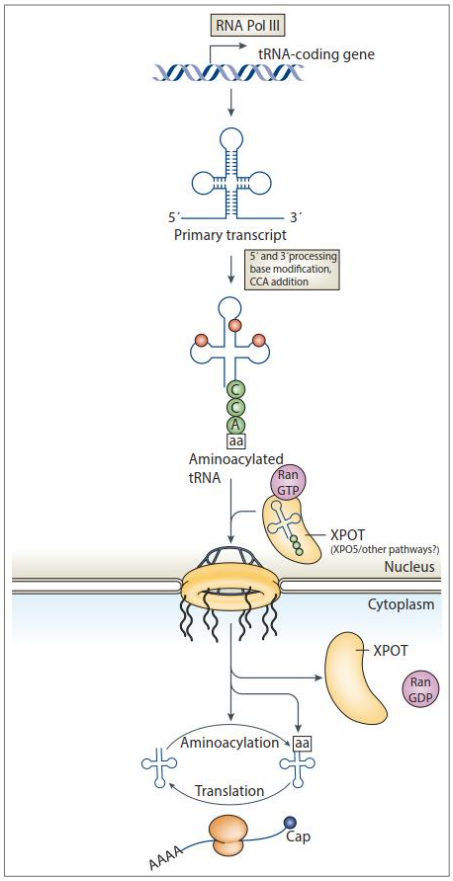


Figure 2: Schematic view of tRNA processing and export.

Genes coding for tRNAs are transcribed by RNA polymerase III. The primary transcript is processed including 5`- and 3`end trimming, base modifications (red circles), 3`end CCA-nucleotide (amino acid acceptor site) addition and intron removal (if present). tRNAs ensue the general export principles using a karyopherin-ß family transport receptor and RanGTP. In this manner, export mature tRNAs bind XPOT and RanGTP and the trimeric XPOT-tRNA-RanGTP complex is translocated to the cytoplasm by interactions with the nuclear pore complex. Other less well characterized pathways may include XPO5, the main export receptor for miRNAs. Upon GTP hydrolysis in the cytoplasm, XPOT is translocated back to the nucleus and the released tRNA is located to the translational machinery. Schematic and figure legend were adapted from reference [5].

1.1.2.2 Processing and export of miRNAs

Micro RNAs play a central role in the regulation of a wide range of biological processes such as cell proliferation, developmental timing, organogenesis, apoptosis and immunity against viruses [33–35]. In that manner, targeted mRNAs are degraded by the activity of the RNA-induced silencing complex (RISC), which is guided to its target by the hybridization of the integrated miRNA [36].

Micro RNAs are encoded by genes that are mainly transcribed by polymerase II and thereby carry a 5`cap structure and a polyadenylated (poly(A)) tail similar to mRNA (see chapter 1.1.3 for more details) [37,38]. In 2006, Borchert et al. showed that polymerase III can also transcribe miRNAs in some cases [39]. Polymerase II transcribed miRNA sequences are found within introns or located at separate genomic loci and are produced as primary miRNAs (pri-miRNA) transcripts. Before function in the RISC complex, pri-miRNAs are processed in both the nucleus and the cytoplasm (Figure 3). Canonically, the pri-miRNAs are first cleaved by the nuclear type III RNase called DROSHA and its adaptor protein DGCR8 producing a 60-80 nucleotide (nt) long stem-looped precursor miRNA (pre-miRNA) with two nt overhangs at the 3`end [40–43]. The unique structure of pre-miRNAs are recognized and bound by the karyopherin ß-family export receptor XPO5 and subsequently exported in a RanGTP-dependent manner [44–47]. Upon GTP hydrolysis and release of the pre-miRNA in the cytoplasm, the stem-loop gets further cleaved by the cytoplasmic type III RNase DICER which results in a ~22 nt RNA duplex product [48]. Finally, one strand of the fully processed miRNA is incorporated via the argonaute (AGO) protein into the RISC complex where it functions in the recognition of the target sequence by base pairing with the 3`untranslated region (3`UTR) (Figure 3) [48–50].

Additional to the canonical pathway, several alternative mechanisms have been described to be involved in the generation of miRNA or miRNA-like RNAs (Figure 3) [51]. The DROSHA/DGCR8-independent pathway was first identified during mRNA splicing in which a small RNA precursor (Mirtron) is generated bypassing the DROSHA/DGCR8-mediated processing step [52,53]. Similar to the canonical pathway, export of this class of miRNAs is thought to occur via XPO5 (Figure 3). In addition, small RNAs which are derived from endogenous short hairpin (sh) RNAs and generated directly through transcription can similarly bypass DROSHA/DGCR8 processing (Figure 3) [51,54,55]. This family of small RNAs carry a 5`cap structure and are thought to be bound and exported by CRM1 (also known as XPO1) through the recruitment of phosphorylated adapter RNA export (PHAX) protein [56]. Given that the interaction between PHAX and the canonical cap-binding-complex (CBC) is required for snRNA export via CRM1, the CBC might be similarly involved.

In addition to the DROSHA/DGCR8-independent pathways, the DICER-independent pathway for miR-451 has been described, which instead uses the catalytic activity of argonaute-2 (AGO2) and poly(A)-specific ribonuclease (PARN) for further processing and trimming, respectively (Figure 3) [57–60]. As well as group I pre-miRNA that have two nt 3`overhangs, group II pre-miRNA with only one nt 3`overhang can be generated. Group II pre-miRNAs involving most of the let-7

family in vertebrates use a TUTase-dependent pathway in which the terminal uridylyl transferases (TUTase) extent the 3`overhang by one nt following DICER processing (Figure 3) [61].

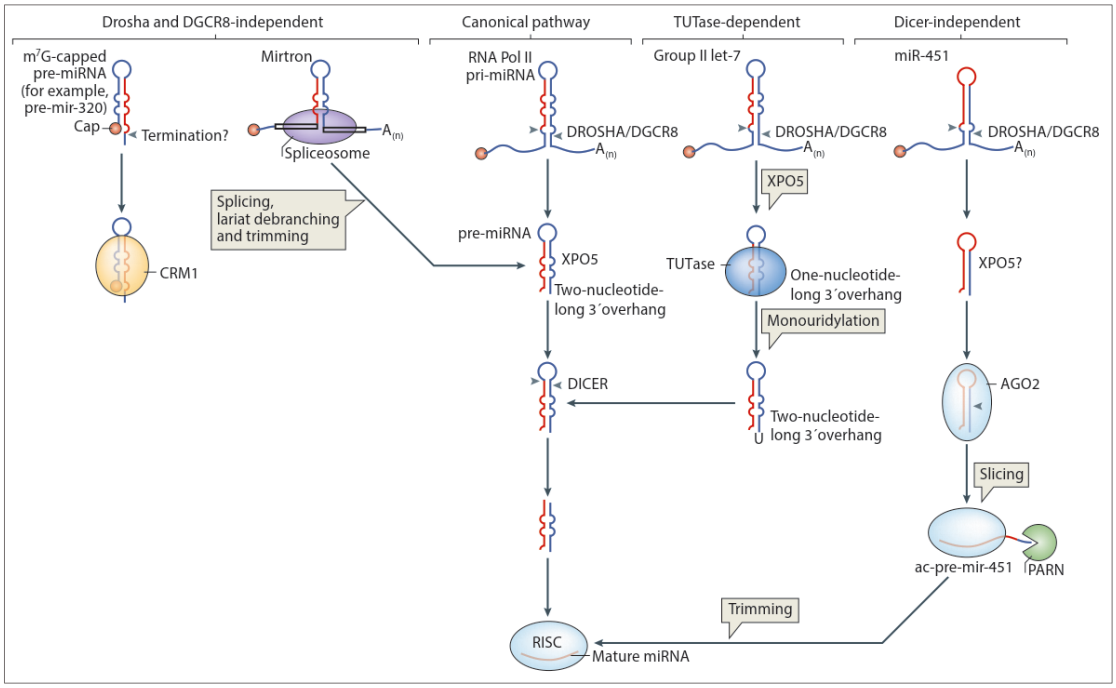


Figure 3: Canonical and non-canonical miRNA processing and export pathways.

Polymerase II transcribed miRNA sequences are found within introns or located at separate genomic loci and are produced as primary miRNAs (pri-miRNA) transcripts. Following the canonical miRNA pathway, the primary transcript is processed by DROSHA/DGCR8 generating a pre-miRNA transcript with a stem-loop structure. Subsequently, the pre-miRNA is bound by XPO5 in a RanGTP dependent manner and exported to the cytoplasm where GTP hydrolysis leads to the release of the pre-miRNA. On the cytoplasmic site, DICER further processes pre-miRNAs before the mature miRNA is integrated into the RISC complex and functions in RNA silencing. The DROSHA/DGCR8-independent pathway bypasses the DROSHA processing. The capped pre-mir-320 product is directly generated through transcription and exported to the cytoplasm bound by CRM1. Mirtron pre-miRNAs are a spliced product, debranched and trimmed before exported to the cytoplasm bound by XPO5. Group II pre-miRNAs follow the terminal uridylyl transferase (TUTase)-dependent export pathway. Pre-miRNA with a shorter 3`overhang are produces and exported to the cytoplasm via XPO5. Due to the shorter overhangs pre-miRNA need to be monouridylated by TUTases before DICER processing in the cytoplasm. Following the DICER-independent pathway, pre-mir-451 transcript generated by DROSHA is possibly exported by the export receptor XPO5. The pre-miRNA is integrated into argonaute 2 (AGO2) without DICER processing. Followed AGO2-dependent splicing and poly(A)-specific ribonuclease PARN-dependent trimming the miRNA is integrated into the RISC complex. Pol II, RNA polymerase II. Schematic and figure legend were adapted from reference [51].

The existence of all non-canonical pathways was shown by several studies, however, only 1% of conserved vertebral miRNAs are generated in a DROSHA/DGCR8- or DICER-independent manner and the majority of miRNAs follow the canonical pathway [51].

1.1.2.3 Processing and export of rRNAs

Ribosomes are RNA-containing particles needed in the process of protein production in the cytoplasm. Four ribosomal RNAs (28S rRNA, 5.8S rRNA, 5S rRNA and 18S rRNA) and more than 70 proteins build up the small (40S) and large (60S) subunit of the ribosome [5]. Before executing their function in the cytoplasm, rRNAs are transcribed, processed and ribosomal subunits are erected in the nucleus and transported to the cytoplasm by several export receptors. This requires the transcription of precursor rRNA (pre-rRNA) and the synthesis and import of

ribosomal proteins. The imported ribosomal proteins are assembled with the pre-rRNAs in the nucleus. In the nucleo- and cytoplasm the pre-ribosomes transiently assemble with more than 150 non-ribosomal proteins [62–64]. Here, they undergo highly regulated and sophisticated mechanisms of processing, maturation and quality control before ribosomal ribonucleic particles (rRNPs) are bound by nuclear export factors in the nucleoplasm. The pre-40S and the pre-60S particles are exported through separate routes to the cytoplasm. The exact mechanism how the pre-40S subunit is transported to the cytoplasm is currently unknown, however, it was known that the karyopherin CRM1 mediates the export in a RanGTP-dependent manner and several pre-40S assembly factors and ribosomal proteins have been implicated to play a role in 40S particle export [65–68]. The export of pre-60S particles has been studied more extensively (mostly in yeast) and depends on the CRM1 export receptor and is therefore RanGTP-dependent [69–74]. In addition, the export adaptor protein NMD3 that contains a NES signal is recruited to the export competent pre-60S particles before nuclear exit through the NPC [69–72,75,76]. In the cytoplasm, RanGTP hydrolysis leads to the dissociation of CRM1 from the NMD3 adaptor. Binding of the cytoplasmic GTPase LSG1 releases NMD3 from the 60S subunit and subsequently the ribosomal protein RPL10 is bound [75,76]. In Yeast, the heterodimer Mex67-Mtr2 (yeast homolog of NXF1-NXT1) was identified as additional export receptor that participate in the export of pre-60S and binds the 60S subunit through a distinct interaction surface [77]. In addition, ARX1 was implicated as an auxiliary export factor which is recruited to the export competent rRNP alongside with NMD3 and MEX67-MTR2 [75,78–80]. In the cytoplasm, binding of REI1, which contributes to the terminal step in pre-60S biogenesis, releases ARX1 and its binding partner ALB1 from the 60S subunit [78,79]. ARX1 was demonstrated to interact with the phenylalanine–glycine (FG) repeats of nucleoporins and thereby shuttling the export receptors through the nuclear pore [81]. The human homolog of ARX1, PA2G4 has, in addition to its function as transcription factor, recently been implicated to play a role in 60S biogenesis by associating with pre-60S particles [82]. In contrast to other RNA classes, the pre-60S particles can associate with several export receptors and thereby potentially increase efficiency of export.

1.1.2.4 Processing and export of snRNAs

Small nuclear RNAs can be divided into two classes, namely the Sm-class and Lsm-class snRNAs. All Sm-class snRNAs are synthesized by polymerase II and after full processing contain a 5`trimethylguanosine cap, a 3`stem-loop and a Sm protein binding site [83,84]. Except for U6 snRNA, all synthesized snRNAs (U1, U2, U4, U5, U7, U11, and U12) belong to the Sm-class snRNAs. U6 snRNA, which is the only member of the Lsm-class snRNA, is characterized by its 5`monomethylphosphate cap, a 3`stem-loop which is followed by a terminating uridine stretch

which forms the platform for Lsm protein binding. U6 snRNA is generated by polymerase III and does not leave the nucleus [85,86]. Apart from U7 snRNP, which is involved in histone pre-mRNA 3`end processing, all other snRNPs function in the spliceosome, which is a multifactorial machinery responsible for the removal of introns from pre-mRNAs in the nucleus [87].

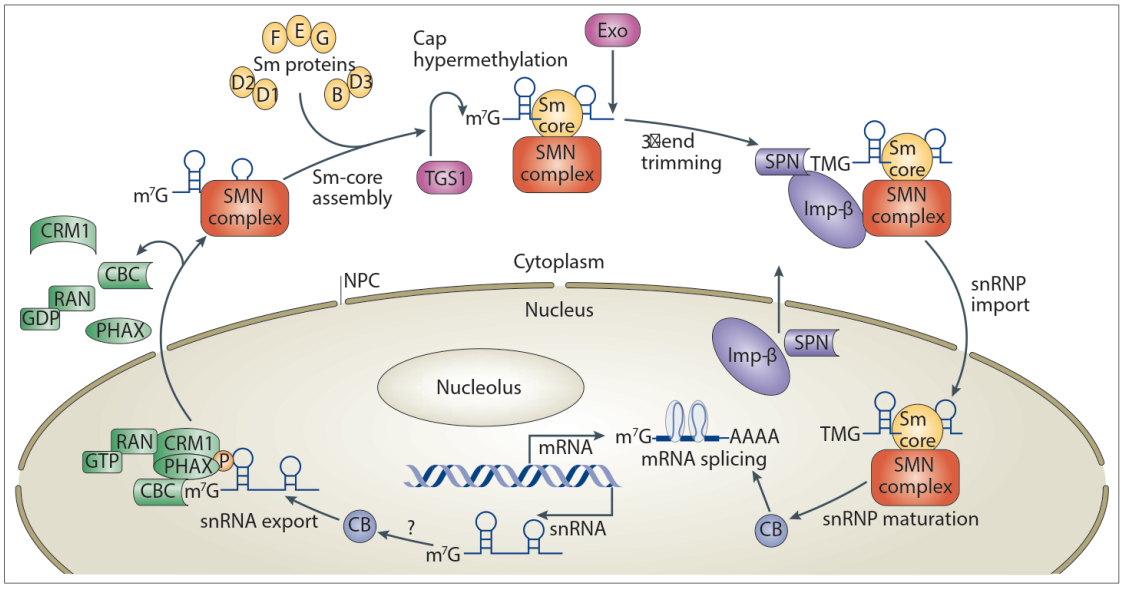


Figure 4: Biogenesis of Sm-class snRNAs.

Sm-class small nuclear RNAs (snRNA) are transcribed by RNA polymerase II and the methylated cap structure (m^7G) is bound by the cap-binding complex (CBC). Additional binding of the hyperphosphorylated form of the export adaptor (PHAX), the export receptor chromosome region maintenance 1 (CRM1) and RanGTP facilitates export to the cytoplasm. On the cytoplasmic site, GTP hydrolysis and dephosphorylation of PHAX drives the dissociation of all factors. Subsequently, the survival of neuron (SMN) complex binds snRNA by a specific sequence element and recruits a set of seven Sm proteins, which from the Sm-core RNP. Trimethylguanosine synthase 1 (TGS1) modifies the m^7G cap with two additional methyl groups forming hypermethylated 2,2,7-trimethylguanosine (TMG) cap structure. In addition, the 3`end is trimmed by an unknown exonuclease (Exo). The hypermethylated cap formation leads to the binding of the import complex including the import adaptor snurportin-1 (SPN) and the import receptor importin-β (Imp-β) leading to re-import to the nucleus. On nuclear site, the re-imported snRNP complex is targeted to cajal bodies (CB) for further snRNP maturation before it functions in mRNA splicing. Schematic and figure legend were adapted from reference [83].

Before functioning in the spliceosome, Sm-class snRNAs are exported to the cytoplasm to undergo maturation and it has been suggested that as part of a quality control mechanism the cytoplasmic phase might be required to prevent nuclear accumulation of dysfunctional snRNAs (Figure 4) [88]. Following transcription by polymerase II, a N7-methylguanosine cap (see chapter 1.1.3.1) is added to the 5`end of snRNAs and the 3`end is processed in the nucleus, however, snRNAs do not get the typical poly(A) tail as found on polymerase II transcripts like miRNAs and mRNAs [89–92]. The 5`cap structure is co-transcriptionally bound by the canonical CBC and the NES-containing adaptor protein PHAX [89,93,94]. Upon phosphorylation of PHAX, CRM1 is recruited to the snRNA and in turn, the snRNA is recruited in a RanGTP-dependent manner. After arrival in the cytoplasm, the snRNA is released by the hydrolysis of GTP and the dephosphorylation of PHAX (Figure 4). Hypophosphorylated PHAX is re-imported to the nucleus by the help of importin-ß [95]. The cytoplasmic snRNA is then assembled with a ring of seven

different Sm proteins, which is enabled by the binding of the survival of motor neuron (SMN) complex [88]. Following the binding of the Sm proteins, the monomethylated 5`cap becomes trimethylated (TMG) by the activity of the trimethylguanosine synthase (TGS1) and the 3`end sequence is trimmed by an exonuclease (Figure 4) [96,97]. The compositions of the snRNP, consisting of the TMG cap and the Sm ring, provides the snRNA with two NLS for re-import to the nucleus. The two NLS are bound by different adaptor proteins [98,99]: The TMG cap NLS is recognized by the import adaptor called snurportin-1, whereas, in case of Sm core NLS the SMN (or a sub-complex) serves as the import adaptor. Both adaptors associate with the import receptor importin-ß and subsequently initiate the translocation of the snRNP to the nucleus [100]. After additional snRNP maturation steps in the cajal bodies of the nucleus, the snRNPs together with a variety of additional splicing factors build up the spliceosome and function in the removal of introns [101–106].

1.1.3 Insights from mRNA: from pre-mRNA processing to export of mature mRNPs

1.1.3.1 The RNA cap structure

The RNA cap structure is the first modification that is co-transcriptionally added to the 5`end of newly generated polymerase II transcripts (Figure 5a) [107]. The phosphorylation state of the C-terminal domain (CTD) of polymerase II, which is composed of 52 heptapeptide repeats (YSPTSPS), drives the timing of RNA modifications. Following transcription initiation, CTD serine 5 becomes phosphorylated which in turn activates the process of 5`end RNA capping by a transcript length of ~20-30 nt [108,109]. The enzymatic activity of three specified enzymes is required for the addition of an N7-methylguanosine to link to the first nucleotide by a 5`-5`triphosphate linkage (Figure 5b) [110].

The first step, which removes the γ-phosphate from the 5`triphosphate and generates a di-phosphate residue, is executed by the RNA triphosphatase (RTPase). Subsequently, the RNA guanylyltransferase (GTase) adds a guanosine monophosphate (Gp) to the RNA 5`diphosphate. Finally, the guanine-N7 methyltransferase (N7MTase) modifies the N7 amine of the guanine with a methyl group. Unlike yeast, metazoan RTPase and GTase are subunits of one protein called capping enzyme (CE) or RNA guanylyltransferase and 5`triphosphatase (RNGTT) which is suggested to perform a more efficient and coordinated reaction [110,111].

The final cap structure builds up a unique feature that not only protects the RNA against 5`-3`degradation but also triggers fundamental events in the biogenesis of polymerase II transcripts by the association with cap-binding proteins such as the CBC in the nucleus and the eukaryotic translation initiation factor 4E (eIF4E) in the cytoplasm.

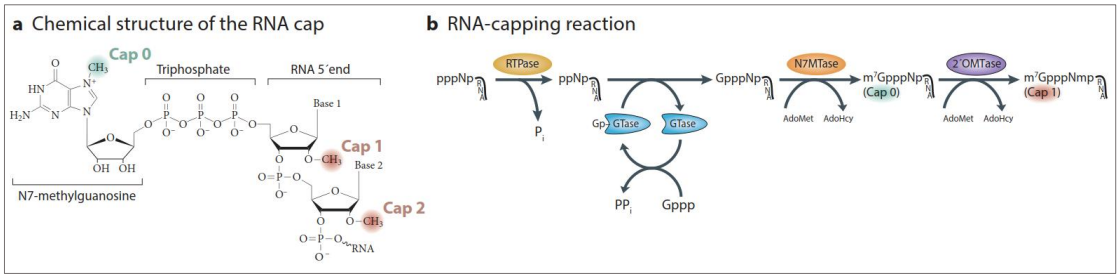


Figure 5: RNA cap structure and the capping mechanism.

(a) The RNA cap consists of a N7-methylguanosine linked by a 5`-5`triphosphate linkage to the 5`nucleoside of the RNA. Three methyl modification can occur at the 5`end of the RNA. Methylation of the N7 position of the guanosine is referred as Cap 0 (green), methylation of the O2 position of the first and second transcribed nucleotide are referred to Cap 1 and Cap 2 (red), respectively. **(b)** The RNA capping reaction generating Cap 0 structures is performed by the action of three enzymes. First, the reaction of RNA triphosphatase (RTPase) removes the γ-phosphate of the nascent RNA (pppNp-RNA) and generates disphosphate RNA (ppNp-RNA) and inorganic phosphate (P_i). The second enzyme, the guanylyltransferase (GTase), adds the α-phosphate of GTP (Gppp) to the disphosphate RNA template via a covalent enzyme-guanylate intermediate (Gp–GTase) releasing pyrophosphate (PP_i) and forming GpppNp-RNA. In the last step, the guanine-N7 methyltransferase (N7MTase) transfers the methyl group from *S*-adenosyl-L-methionine (AdoMet) to the guanine and generates N7-methyl-GpppNp-RNA (m^7GpppNp-RNA; Cap 0) releasing *S*-adenosyl-L-homocysteine (AdoHcy) as a by-product. Cap 1 RNA structure is formed by the action of the AdoMet-dependent 2′*O*-methyltransferase (2′OMTase). Similar to N7MTase, 2′OMTase transfers a methyl group from AdoMet to the O2 position of the first ribose releasing AdoHcy forming the Cap 1 RNA structure (m^7GpppNmp). Schematic and figure legend were adapted from reference [107].

In the nucleus, the CBC instantly binds to the newly synthesized cap structure and orchestrates the recruitment and binding of additional factors that successively lead to the activation of processes such as pre-mRNA splicing and 3`end processing as well as transcription termination, exosomal degradation, RNA export, translation initiation and mRNA pseudo-circulation [88,89]. In addition to the N7-guanosine methylation, metazoan polymerase II transcripts are methylated at the O2 position of the first (and second) transcribed nucleotide, which was reported to be fundamental for the discrimination of self and foreign RNA (e.g. invading viral RNA) by the innate immune system (see chapter 1.2.1) [113].

In the following section, the mechanisms and factors involved in successful mRNA processing and export from the nucleus are described in more detail.

1.1.3.2 Processing and assembly of export-competent mRNPs

Capping of the 5`end of nascent polymerase II transcripts is the hallmark of successful transcription initiation and a prerequisite for elongation of transcripts. Elongated transcripts are bound by heterogeneous nuclear ribonucleoprotein particles (hnRNPs) which are thought to discriminate mRNA from snRNA [114]. RNAs longer than 200-300 nt preferentially bind hnRNPs which displaces PHAX, the snRNA export adaptor, from the transcript (Figure 6a) [115]. hnRNPs predominantly associate with introns in a repeated fashion and wrap ~150-250 nt around its tetrameric core [116–118]. These findings suggested that hnRNPs scaffold pre-mRNA during transcription and act as chaperons and/or regulators [119]. Additionally, one can assume that hnRNPs are able to package RNA similar to histones that bundle DNA and form nucleosomes [120].

Nascent messenger ribonucleoprotein particles (mRNPs) contain hnRNPs as well as spliceosomal components and both protein complexes are assembled co-transcriptionally on mRNA suggesting that hnRNP packaging has a role in mRNA splicing (Figure 6b)[116,120–123]. Packaging could bring distant regions into close proximity enabling cis-interactions such as constitutive or alternative splicing (Figure 6c, d).

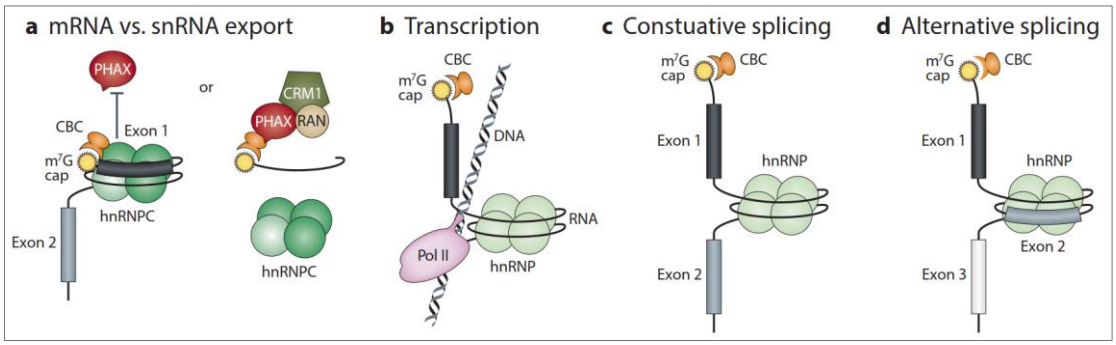


Figure 6: Co-transcriptional mRNA packaging.

(a) Polymerase II transcripts are packed during transcription by heterogeneous nuclear ribonucleoprotein C (hnRNPC) which act as a kind of 'molecular ruler' to sort the transcripts before export to the cytoplasm. This sorting is based on the binding of either hnRNPC or the phosphorylated adaptor for RNA export (PHAX). PHAX binds small nuclear RNAs (snRNA) up to a length of ~200-300 nucleotides by the interaction with the cap-binding complex (CBC). snRNA are escorted to the cytoplasm binding to chromosomal region maintenance protein 1 (CRM1) (right branch). Transcripts longer than ~200-300 nucleotides are subsequently bound by the tetrameric hnRNPC complex (left branch), which inhibits the association of PHAX. Transcripts wrapped around the tetrameric hnRNPC complex are targeted to the nuclear RNA export factor 1 (NXF1) export pathway. **(b)** Transcription: The tetrameric hnRNP complex is assembled co-transcriptionally on mRNA and this packaging is suggested to have a role in mRNA splicing. **(c)** Constitutive splicing: Packaging by the hnRNP tetramer holds consecutive exons in close proximity and ensures correct ordered splicing. **(d)** Alternative splicing: Through packaging by the tetrameric hnRNP complex strong splice sites or alternative exons could be hided and thereby exon skipping could be promoted. Additionally, weak splice sites could be exposed to facilitate exon inclusion. m^7G, 7-methylguanosine. Schematic and figure legend were adapted from reference [120].

The process of splicing which removes introns from pre-mRNA is initiated by the phosphorylation of CTD serine 2 [108]. Splicing is catalyzed by the spliceosome that assembles *de novo* on each pre-mRNA transcript during transcription and is composed of five spliceosomal small nuclear ribonucleoprotein particles (snRNPs) associated with a large number of additional proteins [122]. In pre-mRNA splicing, the importance of the m^7G (N7-methylguanosine) and the binding of the CBC was demonstrated in several studies. The initial study using *Xenopus leavis* oocytes showed that micro-injected transcripts need the m^7G structure to efficiently splice 5′proximal introns [124]. Another study by Izaurralde et al. revealed the importance of the CBC for efficient mRNA splicing by microinjecting transcripts in nuclear cap-binding protein 2 (NCBP2) immuno-depleted *Xenopus leavis* oocytes that resulted in decreased splicing efficiency [94]. Treating total or nuclear HeLa extracts with uncapped or m^7G -capped transcripts confirmed that the m^7G group is essential for efficient splicing in mammalian cells [125–129]. Additionally, inhibition of the CBC in HeLa cells resulted in altered mRNA splicing and decreased recruitment of U1 snRNP to the 5`splice site of 5`proximal introns [125,130]. Recently, it was shown that the spliceosome assembly depends on the interaction of the CBC with protein components of the U4/U6.U5 tri-snRNP complex [131].

Additionally, m^7G and the CBC is not only required for removal of 5`proximal introns but also for splicing of downstream introns [131,132]. However, it is still unclear to what extent the CBC is required for downstream splicing on a transcriptome wide level.

In the wake of splicing, the exon-junction complex (EJC) is deposited ~20-24 nt upstream of each exon-exon junction [133,134]. The EJC couples splicing events to post-transcriptional steps such as mRNA export and surveillance. The best-studied surveillance mechanism is the nonsense-mediated decay (NMD) pathway, which detects premature termination codons in the pioneer round of translation sensing possible downstream EJCs and removing these mRNAs from the translational pool [135]. Early in splicing, the EJC core proteins EIF4A3, MAGOH and RBM8A are recruited transiently to the spliceosome and build up the trimeric pre-EJC showing that EJC assembly and pre-mRNA splicing are two tightly linked processes (Figure 7) [122]. In this regard, the association of EIF4A3 and the pre-mRNA-splicing factor CWC22 homolog (CWC22), which is an abundant and essential factor of the spliceosome is to date the only identified direct interaction of the spliceosome and the pre-EJC [136–141]. It was shown that splicing efficiency as well as EJC assembly is impaired when CWC22 is missing emphasizing the necessity of the interaction of CWC22 and EIF4A3 for EJC assembly during splicing [138,139]. Following splicing termination by exon ligation and release of the spliceosome, the trimeric pre-EJC is joined by barentsz (BTZ; also known as MLN51 or CASC3) protein to form the mature, tetrameric EJC (Figure 7) [142]. CWC22 keeps EIF4A3 in an open state, however, it is not known how the transition of the open to the closed state, where CWC22 releases EIF4A3, happens [143,144]. In its closed form EIF4A3 can bind RNA and ATP and it was shown that the three other core proteins are required to stably keep EIF4A3 in its closed RNA-bound form and to prevent release of hydrolyzed ATP [144,145]. The EJC clamps on the RNA in a sequence-independent mode which was revealed by the crystal structure of the core EJC bound to RNA [143,146]. EIF4A3 RNA-binding is established mainly through the binding of the ribose-phosphate backbone and not by the association of specific bases. Studies combining biochemical and mass spectrometry based analysis showed that more than 30 proteins can be co-purified with the core EJC [147–149]. The four core EJC proteins employ a large number of transiently interacting proteins to fulfill the demanding varying role in post-transcriptional processes such as export of mature mRNA to the cytoplasm, initiation of translation and quality control mechanisms.

The EJC-associated proteins can be classified as components of outer shell and transiently interacting proteins [149–151]. The outer shell is composed of ALYREF, PNN, ACIN1, RNPS1 and SAP18 from which ACIN1/RNPS1/SAP18 form the apoptosis and splicing associated protein

(ASAP) complex and PNN/RNPS1/SAP18 build up an alternative complex named PSAP [149,152–154]. It is currently unknown to which extent the function of the two complexes overlap or how they are recruited to the EJC. It is speculated that the ASAP associates with the EJC during deposition on mRNAs indicated by co-purification with the spliceosome and the EJC and localization to the nucleus [149,152,155,156]. A more transient interaction with the EJC was found for UAP56, NXF1/NXT1 (also known as TAP/p15) as well as SRRM1, UPF1, UPF2 and UPF3a/b.

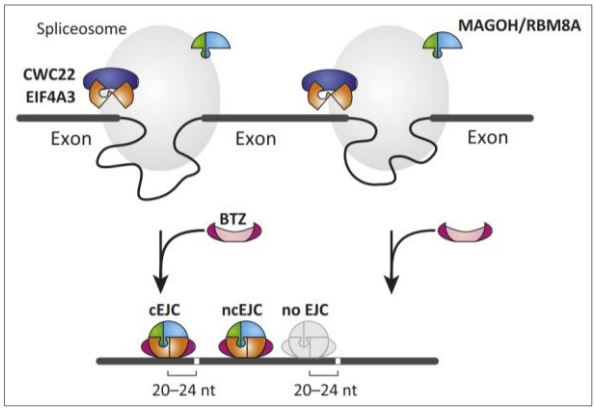


Figure 7: Assembly of the exon-junction complex.

During splicing of pre-mRNAs in the nucleus, the exon-junction complex (EJC) is assembled on mRNAs. One essential factor of the spliceosome, namely CWC22, recruits the core EJC component EIF4A3 to the RNA. The stable MAGOH/RBM8A heterodimer is recruited to the RNA in a separate so far unknown mechanism. The spliceosome releases the spliced mRNA on which the EJC is deposited on exon-exon junctions. After release from the spliceosome, barentsz (BTZ), the fourth EJC subunit, binds to the EJC core. Three potential outcomes of splicing and EJC assembly have been described: (1) binding of a canonical EJC at a position 20-24 nucleotides (nt) upstream of the exon–exon junction, (2) binding of a non-canonical EJC at a position different from the canonical binding site, or (3) absence of an EJC at the canonical binding site. Schematic and figure legend were adapted from reference [156].

The EJC core, as well as its interacting proteins, participate in various pathways such as splicing/alternative splicing (SRRM1, SAP18, RNPS1, ACIN1, PNN), mRNA export (UAP56, ALYREF, NXF1/NXT1) and NMD (UPF1, UPF2, UPF3a/b) and thereby effectively connect these pathways [134,153,157–170]. Splicing was reported to promote transcript kinetics and abundance in the cytoplasm, demonstrating a direct dependence of splicing on mRNA export [171]. Additionally, the export factors ALYREF (also known as THOC4) and UAP56 are recruited to spliced mRNAs via the EJC linking pre-mRNA splicing to mRNA export [165,170,172]. ALYREF and UAP56 are EJC-associated proteins that together with THOC1, THOC2, THOC3, THOC5, THOC6, THOC7 form the transcription-export 1 (TREX-1) complex that couples transcription to mRNA export [165,172–174]. More recently, researchers identified SARNP, POLDIP3, ZC3H11A, CHTOP, DDX39A, UlF ,ERH and LUZP4 as putative new TREX-1 subunits [175–180]. Cheng et al. showed in 2006 that TREX-1 complex is recruited to the 5`end of the mRNA in a splicing- and cap-dependent manner [181]. This recruitment is based on the interaction between ALYREF and the nuclear cap-binding protein 1 (NCBP1; CBC subunit). The interaction of ALYREF with TREX-1 complex is dependent on ATP-binding of UAP56 [175]. The mRNA export receptor NXF1 is

recruited to the mRNA through its interaction with ALYREF and THOC5 [182]. This interaction drives NXF1 into its open conformation revealing its RNA recognition motif (RRM) and subsequently allowing the direct binding of mRNA (Figure 8). In turn, the binding activity of ALYREF to the mRNA is decreased implicating that mRNA is handed from ALYREF to the export receptor NXF1 at this step [183]. This highlights the important role of ALYREF in mRNA biogenesis, through coupling the EJC and TREX complex with mRNA export. Interestingly, knockdown of single TREX components normally results in a moderate export block, whereas double knockdown reveals a severe export block. In contrast, single knockdown in yeast already results in a rapid mRNA export block [174]. This suggests that metazoans have developed a mechanism to ensure processing and export even if one export player is downregulated.

To complete processing of mRNAs, the 3`end is cleaved and subsequently polyadenylated to define the end of the mRNA. Consequently, mRNAs are released from the transcription machinery and are targeted to the export machinery [120]. The cleavage and polyadenylation (CPA) complex is recruited to the poly(A) signal (PAS) in the 3`UTR of pre-mRNA by interacting with serine 2 phosphorylated CTD and RNA binding domains that specifically bind the PAS [184,185]. The poly(A) polymerase (PAP) adds the poly(A) tail which is rapidly bound by poly(A)-binding proteins (PABPs) increasing the processivity of PAPs. The nuclear poly(A)-binding protein PABPN1 is thought to act as "molecular ruler" by determining the length of the poly(A) tail [186]. Studies using electron microscopy revealed that the poly(A) tail is coated by many PABPN1 generating a spherical particle with ~250 accompanied nucleotides [186]. As expected, genome-wide studies have shown that loss of PABPN1 results in overall shortening of 3`UTRs [186]. In addition, PABPN1 is essential for bulk mRNA export and shuttles between the nucleus and the cytoplasm [186]. Interestingly, the CPA complex is associated with ALYREF, CLIP, DDX39A/B and PCF11 and it was hypothesized that these proteins are involved in the PAS selection [187]. In addition, THOC5 was found to interact with the 3`end processing factors CPSF2 and CPSF6 suggesting to be involved in PAS choice for specific genes [188,189]. CPSF6 is a positive regulator of alternative polyadenylation (APA) and known to be important in selecting alternative poly(A) sites [186]. To date, it is believed that APA and mRNA splicing are connected in some way, however, it is not known how APA and mRNA export are linked [190,191]. It was demonstrated that U1 snRNP could suppress APA binding to pre-mRNAs introns or non-canonical PASs in the 3`UTRs and thereby shielding premature cleavage and polyadenylation sites [192,193]. Moreover, CPSF6 has recently been shown to shuttle to the nucleus and associate with the export receptor NXF1 indicating a potential link between APA and mRNA export [194].

Overall, the coupling of 3`end processing and mRNA export is needed at least for canonical polyadenylation to enable the release of fully processed transcripts from the transcription machinery and ensure efficient targeting to the export machinery.

1.1.3.3 Routes of mRNA export

After the three main processing steps of mRNA capping, splicing and 3`end processing, correctly processed mRNPs are targeted for export and translocation through the NPC to the cytoplasm. In all cases, the transport of macromolecules like mRNA needs energy to shuffle through the NPC to the cytoplasm. In case of bulk mRNA, this energy is generated by the proteins GLE1, DDX19 (also known as DBP5) and inositol hexakisphosphat (IP_6) facilitated by an ATP cycle (Figure 8). The translocation through the nuclear pore can be divided into three steps characterized by the docking of the mRNA onto the nuclear basket, the transit through the central channel and the release of the mRNA into the cytoplasm [2]. The export receptor NXF1 bound to NXT1 is the most common export receptor for bulk mRNA export. Through the recruitment of multiple adaptor proteins NXF1 can increase its affinity for its target RNAs. It was also shown that CRM1 can associate with adaptor proteins to bind specific mRNAs and utilize their export to the cytoplasm. In the following paragraphs, the different mRNA export routes with their export receptors, adaptors and their selectivity for targets are described in more detail.

NXF1 mediated export of bulk mRNAs

The first identification of NXF1 was as an export receptor for Mason Pfizer monkey virus where NXF1 interacts with the constitutive transport element (CTE) of viral RNAs and exports them to the cytoplasm [195–197]. Only later, studies revealed the fundamental role of NXF1 in bulk mRNA export [196,198]. NXF1 is crucial for cell proliferation and depletion of NXF1 leads to retention of polyadenylated RNA in the nucleus. NXF1 is composed of multiple domains, namely the arginine-rich RNA-binding domain (RBD), adjacent ψRRM, leucine-rich repeat (LRR), NTF2-like (NTLF2L) and ubiquitin-associated (UBA) domain [2]. THOC5 binds the NTF2L domain of NXF1 and links the association of NXF1 with ALYREF [183,199–201]. NXF1 binds non-specifically to RNA via its N-terminal RDB domain, however, once ALYREF is bound to NXF1, the RNA-binding affinity of NXF1 is enhanced. Upon ALYREF binding, NXF1 exhibits its open form resulting in the transfer of RNA from ALYREF to NXF1 (Figure 8) [183]. NXF1, in its open and mRNA-locked form, escorts the mRNP complexes to the nuclear pore and translocates the mRNPs through the nuclear pore by interacting with nucleoporins via its NTF2L and UBA domains (Figure 8). Simultaneous depletion of ALYREF and THOC5 reduces the amount of NXF1 bound to mRNA highlighting the major role of TREX-1 complex to recruit NXF1 to

mRNAs [182]. In addition, by recruiting NXF1 to the mRNA cargo via TREX-1 it is ensured that only fully processed mRNAs are targeted for export.

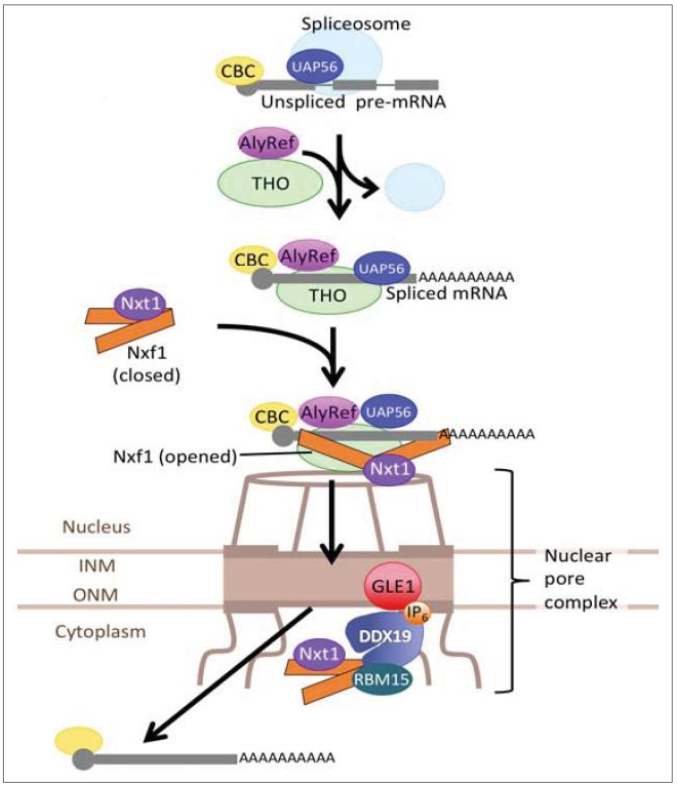


Figure 8: NXF1-dependent mRNA export.

During NXF1-dependent mRNA export, the transcription-export 1 (TREX-1) complex consisting of ALYREF, UAP56 and the THO complex subunits assembles on the spliced mRNA. In turn, the NXF1/NXT1 heterodimer is recruited to the mRNA via interactions with TREX-1 complex subunits. Bound by the open form of NXF1, mRNPs are translocated to the cytoplasm trough interaction with subunits of the nuclear pore complex (NPC). CBC, cap-binding complex; INM, inner nuclear membrane; IP_6, Inositol hexakiphosphate ;ONM, outer nuclear membrane. Schematic and figure legend were adapted from reference [202].

Transcription and processing is thought to take place in nuclear speckles in the nucleus, which means that the mRNP complexes need to travel through the nucleoplasm to reach and dock onto the NPC. In this context, the TREX-2 complex was proposed to assist mRNPs on their way to the nuclear basket [203,204]. TREX-2 is composed of GANP, ENY2, CETN2/CETN3, PCID2 and DSS1. GANP directly interacts with NXF1 and scaffolds the other members of the TREX-2 complex to the mRNA [205]. A fraction of TREX-2 is located within the nucleoplasm, however, the main fraction is located at the NPC [203,204,206]. ENY2 and GANP were shown to also interact with RNA polymerase II and thereby potentially facilitating the recruitment of TREX-2 to a subset of specific genes [204]. TREX-2 was shown to transfer mRNA to NXF1. However, the exact mechanism is not well characterized. In addition, it is still unclear whether TREX-2 is a general export complex or utilizes export of specific transcripts. Studies demonstrated that GANP specific transcripts were exported faster than others suggesting an accelerated pathway enabled by GANP for specific transcripts [207]. In addition, GANP depletion experiments revealed that GANP is

needed for the export of transcripts involved in gene expression (e.g. mRNA processing, splicing and ribosome biogenesis) [207]. These GANP-specific transcripts have in average shorter half-lives and are expressed above-average compared to GANP-independent transcripts. This suggests the potential of GANP to institute an accelerated export route for specific transcripts, which is needed to rapidly adapt to internal or external stimuli.

The transition through the nuclear pore is a better characterized process. After docking to the nuclear basket, the traveling through the nuclear pore occurs via series of interactions with FG Nups [2]. First, RAE1 interacts with NXF1 and it was proposed that RAE1 delivers NXF1/mRNPs to Nup98 [208,209]. Inside the channel, NXF1 interacts with Nup62, however, the exact mechanism how NXF1/mRNPs travel through the central channel is still unclear. Nonetheless, recent studies have revealed insights into the kinetics of mRNPs translocation from the nuclear site of transcription to the cytoplasmic site of translation. The transit through the NPC is a relatively rapid process which takes less than 500 ms, whereas, the diffusion of mRNPs from transcription and processing speckles can take up to several minutes and is seen as the rate-limiting step in this pathway [210–212]. Surprisingly, only 25-35% of initially docked mRNPs are successfully transferred to the cytoplasm [210–212]. From their place of transcription and processing, mRNPs transit trough regions of interchromatin space or zones of heterchromatin exclusions in the nucleoplasm, which build up a channel-like structure that enable the directed diffusion of mRNPs to the NPC [213–215]. The filamentous NPC-associated nucleoprotein TPR is thought to from such channels since depletion of TPR resulted in heterchromatin-covered NPCs [216,217]. Interestingly, TREX-2 associates with TPR at the nuclear basket, which could result in more efficient diffusion of TREX-2 associated mRNPs to the NPCs and a faster export of TREX-2 dependent transcripts [206,207].

Passaged the nuclear pore channel, mRNAs associate with fibrils on the cytoplasmic site of the NPC mainly comprised of Nup358/RanBP2 [218]. RanBP2 associates with the NPCs via Nup88 and Nup214 and can additionally interact with NXF1, RanGAP and Ran [218–220]. DDX19 and GLE1 enable the release of mRNAs to the cytoplasm in an ATP-dependent manner (Figure 8). In that manner, GLE1 associates with the signaling molecule IP_6 stimulating the binding of DDX19 to the mRNA and triggering mRNA release by ATP hydrolysis [2,221]. The release and remodeling of the mRNA in the cytoplasm leaves free export factors that are recycled back to the nucleus. mRNAs are then transferred to the translational machinery where they undergo further remodeling. For instance, the CBC complex is replaced after the pioneered round of translation by the cytoplasmic cap-binding protein eIF4E allowing steady-state translation.

NXF1 mediated export of specific mRNA transcripts

In addition to the above described mechanisms, transcript specificity can be introduced by changing the composition of the mRNPs. In this manner, NXF1 can associate with different adaptor complexes to allow binding of specific transcripts (Figure 9). For instance, components of the TREX complex were demonstrated to be involved in the export of specific transcripts. THOC5 was shown to be dispensable for bulk mRNA, however, in association with ALYREF needed for the export of HSP70 mRNAs which are crucial in heat-shock responses (Figure 9a) [222]. THOC5 together with THOC2 controls mRNA export transcripts involved in pluripotency, e.g. Nanog and Sox2, and thereby contributing to embryonic stem cell self-renewal and differentiation (Figure 9a) [223]. Interestingly, THOC5 is reduced during differentiation leading to decreased interaction of THOC2 with these transcripts, which results in reduced export of transcripts required for pluripotency.

The alternative mRNA export (ALREX) pathway including CIP29, DDX39 and NXF1 exports a subset of transcripts involved in mitosis [180]. Loss of DDX39 results in altered Birc5 and Prc1 mRNA export leading to mitotic defects like chromosome arm resolution defects and failure of cytokinesis. CHTOP is another TREX component, which can compete with ALYREF for NXF1 binding and thus may modulate the export of specific transcripts [180]. The newly identified TREX-component LUZP4, which associates with UAP56 and NXF1 was shown to be upregulated in cancer cells, indicating a dysregulation in mRNA export [180]. LUZP4 was suggested to be involved in export of a subset of transcripts needed for melanoma cells proliferation, however, the subset of transcripts regulated by LUZP4 was thus far not investigated in detail. Organism-wise, it is possible that cell-type specific TREX complexes exist and are involved in export of specific transcripts.

Besides TREX components, SR proteins can also compete with ALYREF for NXF1 binding. In that manner, SRSF1, SRSF3 and SRSF7 were identified to interact with NXF1 and export specific transcripts [224]. SRSF3 and SRSF7 recognize the intronless transport element (ITE), a 22-nucleotide motif in histone H2A mRNA, and target these transcripts to export via NXF1. Following this finding, SRSF3 and SRSF7 were depending on their phosphorylation status demonstrated to be involved in the transport of spliced transcripts to the cytoplasm. In its hyperphosphorylated form, SRSF3 associates with mRNA prior to splicing. After splicing SRSF3 becomes hypophosphorylated and thereby targets its transcripts to NXF1-dependent export [225,226]. In the cytoplasm, re-phosphorylation occurs leading to release of NXF1 and mRNA.

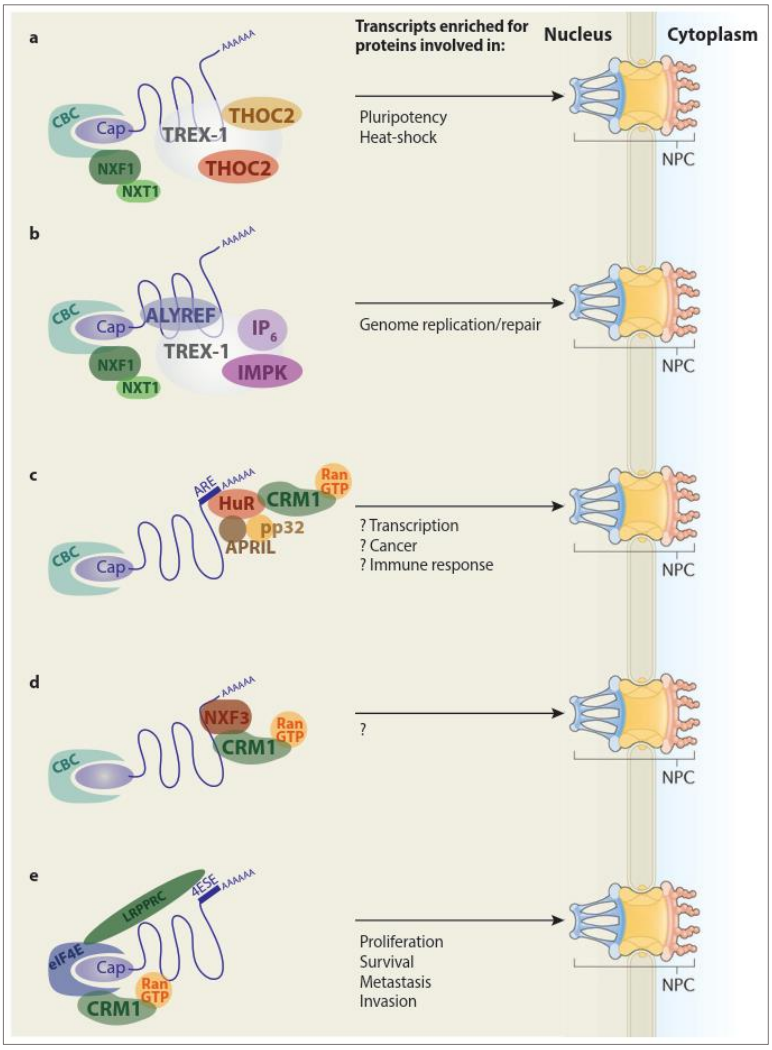


Figure 9: Selective mRNA export pathways.

(a) Selective mRNA export for transcripts involved in pluripotency and heat-shock. Selectively is driven by the control of THOC5 and THOC2. **(b)** Genome replication and repair transcripts are exported selectively by the binding of ALYREF and the enzyme inositol polyphosphate multikinase (IPMK) together with IP_6. **(c)** Hu-antigen R (HuR) in association with pp32 and APRIL binds to a subset of transcripts and exports these transcripts to the cytoplasm in a CRM1-dependent manner. Transcripts were suggested to be involved in transcription, cancer as well as in immune responses. **(d)** NXF3 exports a so far unknown subset of mRNAs in a CRM1-dependent manner. **(e)** Transcripts involved in proliferation, survival, metastasis and invasion are exported in a CRM1-dependend manner bound by eIF4E and LRPPRC, which binds to the 4E-sensitive element (4ESE) in the 3`UTR. CRM1-dependent pathways require the binding of RanGTP in the nucleus and release of the mRNA in the cytoplasm is driven by GTP hydrolysis. CRM1, chromosome region maintenance 1; IPMK, inositol polyphosphate multikinase; LRPPRC, leucine-rich PPR motif-containing protein; NPC, nuclear pore complex. Schematic and figure legend were adapted from reference [202,221].

Another selective mRNA pathway including export of transcripts required for genome replication and repair, e.g. Rad51, Chek1 and Fancd2, depends on the binding of ALYREF and the enzyme inositol polyphosphate multikinase (IPMK) [227]. IPMK is required for the binding of ALYREF to its target transcript in which a sequence motif in the 3`UTR permits ALYREF binding (Figure 9b).

Adaptor proteins mostly preserve specificity by defined sequence elements in the RNA. For example, RNAs containing the RNA transport element (RTE) consisting of a minimum of four internal stem loops are bound by RBM15 and targeted for translocation via NXF1 [228,229]. In some cases, NXF1 can directly bind RNAs by interacting with specific RNA elements. One of such

element is CTE, which was first identified in Mason Pfizer monkey virus to be important to export viral transcripts in a NXF1-dependent manner [196,198,230]. RNA containing the signal sequence-coding element (SSCR) present on transcripts involved in secretion can also be exported in a NXF1-dependent manner [231,232]. SSCR was also found to act as an export sequence for transcripts lacking introns or functional caps being NXF1-dependent but TREX-independent.

Interestingly, Nups can also impact export of specific transcripts. In this context, Nup96 was shown to be involved in the export of transcripts encoding cell cycle and immune responses regulators, namely ß2-microglobulin, CDK6, MHCI and MHCII [233,234].

CRM1 mediated export of specific mRNA transcripts

As described in chapter 1.1.2, CRM1 is the main export receptor for snRNA. Since CRM1 is unable to directly bind RNA, it associates with export adaptors containing NES. It was demonstrated that a subset of mRNA can be exported using the RanGTP-dependent CRM1 export pathway.

One subset of mRNA is selected and exported via CRM1 by the interaction of human antigen R (HuR; also known as ELAVL1) protein with AU-rich elements (AREs) in the 3`UTR of the RNA encoding for oncoproteins, cytokines and transcription proteins [235–237]. In addition to HuR, pp32 and APRIL also associate with CRM1 to export these mRNAs (Figure 9c) [238]. Studies have shown that not all CRM1-dependent transcripts rely on the association with HuR. For example, human type I interferon α is exported in a CRM1-dependent manner but independent of HuR binding [239,240]. Another adapter associating with CRM1 is NXF3 [241]. NXF3 belongs to the NXF protein family but is unable to associate with Nups, which leads to the association and export utilizing CRM1 as export receptor (Figure 9d). Currently, specific sequence elements and the exact mechanism of this pathway is unknown and needs to be further investigated.

Another selective pathways using CRM1 as export receptor is dependent on the binding of the translation initiation factor eIF4E with transcripts involved in proliferation, survival, metastasis and invasion (Figure 9e). The eIF4E-dependent transcripts contain an approximately 50 nucleotide long secondary structure element in their 3`UTR known as 4E-sensitive element (4ESE) [242]. The association of eIF4E with the sequence element is linked by the direct interaction of the export adaptor protein LRPPRPC with 4ESE and eIF4E [243]. The dependency of CRM1 for the export of this subset of genes was demonstrated by the use of the CRM1 inhibitor leptomycin B which resulted in the inhibition of eIF4E-dependent mRNA export[242]. Moreover, depletion of NXF1 did not affect export of eIF4E-depenedent transcripts. Interestingly, RIP-seq experiments showed that eIF4E can bind to around 2300 transcripts in the nucleus suggesting a broader

importance of CRM1 and eIF4E in the export of mRNAs [244]. To date, it is known that factors like UAP56, hnRNPA1 and DDX3 are common to both, the NXF1 and eIF4E/CRM1 export pathways, whereas, factors like NXF1, ALYREF and CBC are not required for mRNA export via eIF4E/CRM1 [243]. Since 4ESE-containing mRNAs have long 3`UTRs with several export sequence motifs and are targets for both, bulk and eIF4E-depedent export, it can be hypothesized that the eIF4E/CRM1 competes with bulk mRNA export factors to boost export of specific transcripts [2]. Studies propose that RanBP2 reduces release and/or recycling of eIF4E-depentent export factors to the nucleus [245,246]. To overcome this inhibitory mechanism and enhance export, eIF4E can indirectly suppress RanBP2 function. This might be important in a wide range of cancers since eIF4E is upregulated in these tissues influencing their proliferative capacity by promoting eIF4E/CRM1 export pathway [245,246].

Other pathways

Two other mRNA export pathways were described to transport mRNPs to the cytoplasm, namely nuclear envelope budding and NPC enlargement [2]. The first mechanism was shown to happen during synapse development and enables the export of large mRNP granules by budding out the NE similar to the mechanism of nuclear egress used by herpes viruses [247–251]. Specifically, recruitment of the atypical protein kinase C to the nuclear envelope leads to phosphorylation of the nuclear lamina, which prompts the invagination of the inner nuclear membrane into the NE lumen. Consequently, fusion with the outer nuclear membrane enables the delivery of large mRNPs into the cytoplasm [2]. The second mechanism of envelope enlargement was first seen during influenza A virus infection (IAV) [252]. It was shown that viral infection could enlarge the nuclear pores in a caspase-dependent manner to approximately 50 nm (normally approximately 30 nm) facilitating the translocation of large protein-RNA complexes [253–255]. Viral mRNPs are usually exported in a CRM1-dependent manner bound by the viral export adaptor protein nucleoprotein (NP). However, at late stages of infection the enlargement of the nuclear pore occurs, which enables the passive diffusion of proteins with a size of ~125 kDa and viral RNPs. Whether this mechanism plays a role to host mRNP translocation needs further investigation [254].

1.2 RNA EXPORT DURING VIRAL INFECTION

Exporting mRNAs is a critical step in the lifetime of a cell to regulate and maintain cellular functions. However, the nucleo-cytoplasmic transport of RNA becomes even more important when the cell encounters environmental challenges. This includes the appropriate response to pathogens such as viruses. During a viral infection, the invading pathogen is recognized by cytoplasmic or endosomal pattern recognition receptors (PRR) via pathogen associated molecular patterns (PAMPs). Following recognition of the virus, the cell initiates innate immune responses by expressing cytokines and chemokines that establish a pro-inflammatory and anti-proliferative antiviral state. This process involves the expression of type I (IFNα, IFNβ), II (IFNγ), III (IFNλ) interferons (IFN) to defend viral infection by initiating the transcription of mRNAs that encode antiviral factors.

1.2.1 Discrimination between self and non-self RNA

The molecular basis of discriminating host appearance from foreign PAMPs concerning the recognition of invading foreign RNA during virus infection by PRRs was the topic of a review included in the special issue "*Discrimination of Viral and Self Nucleic Acids and IFN Signature Diseases*" of the *Journal of Interferon & Cytokine Research*. I wrote the review together with Beatrice T. Laudenbach und Andreas Pichlmair.

Gebhardt, A. *, Laudenbach, B.T. *, and Pichlmair, A. (2017). Discrimination of Self and Non-Self Ribonucleic Acids. *J. Interf. Cytokine Res.* 37, 184–197. **these authors contributed equally*

JOURNAL OF INTERFERON & CYTOKINE RESEARCH
Volume 37, Number 5, 2017

DOI: 10.1089/jir.2016.0092

REVIEWS

Discrimination of Self and Non-Self Ribonucleic Acids

Anna Gebhardt,* Beatrice T. Laudenbach,* and Andreas Pichlmair

Most virus infections are controlled through the innate and adaptive immune system. A surprisingly limited number of so-called pattern recognition receptors (PRRs) have the ability to sense a large variety of virus infections. The reason for the broad activity of PRRs lies in the ability to recognize viral nucleic acids. These nucleic acids lack signatures that are present in cytoplasmic cellular nucleic acids and thereby marking them as pathogen-derived. Accumulating evidence suggests that these signatures, which are predominantly sensed by a class of PRRs called retinoic acid-inducible gene I (RIG-I)-like receptors and other proteins, are not unique to viruses but rather resemble immature forms of cellular ribonucleic acids generated by cellular polymerases. RIG-I-like receptors, and other cellular antiviral proteins, may therefore have mainly evolved to sense non-processed nucleic acids typically generated by primitive organisms and pathogens. This capability has not only implications on induction of antiviral immunity but also on the function of cellular proteins to handle self-derived RNA with stimulatory potential.

Keywords: ribonucleic acid sensing, antiviral mechanisms, interferon, MDA5, RIG-I, PRR

Introduction

THE HOST IMMUNE SYSTEM is constantly encountering pathogen invasion. Viruses, the most abundant pathogens on earth, can infect eukaryotes and prokaryotes and require the supportive environment of the cell to proliferate and spread. For this reason, organisms evolved barriers including the innate and adaptive immune system to suppress growth of pathogens. Protection against viral infections is to a large extent mastered by the innate immune system, which is able to sense incoming virus particles, viral proteins, viral replication products, and changes in the general integrity of the cell (Isaacs and Lindenmann 1957; Rassa and Ross 2003; Sancho and Reis e Sousa 2013). As a result, the organism mounts an appropriate antiviral response that impairs virus growth and allows virus clearance.

Activation of the antiviral innate immune system is characterized by secretion of antiviral cytokines, particularly type I interferons (IFN-α/β). These cytokines were identified by Isaacs and Lindenmann (1957) who demonstrated that cells secrete antiviral factors upon exposure to heat inactivated viruses. The main stimulatory agent was identified to be nucleic acids that are delivered through the viral infection process (Goubau and others 2013). Viral replication greatly enhances the abundance of stimulatory nucleic acids and thereby regulates the magnitude of the response (Rehwinkel and others 2010; Weber and others 2013). Both, viral RNA and DNA (vRNA/DNA) can stimulate pattern recognition receptors (PRRs) and recently molecular principles underlying the basis for detecting viruses and immune-stimulatory nucleic acids were discovered (Schlee 2013; Ahmad and Hur 2015). Besides activation of the innate immune system, viral nucleic acids modulate the activity of general cellular machineries such as protein translation, RNA degradation, or cell death (Fig. 1). At the same time viral nucleic acids are directly targeted by cellular proteins with antiviral functions. Thus, it becomes more and more evident that the innate immune system is not only a blunt alarm system that reacts to a single stimulus. It rather consists of a highly sophisticated network of proteins that target pathogen-derived nucleic acids. Importantly, the innate immune system has dramatic impact on the cellular and organismal level and its activation has to be modulated in a very tight manner. In this regards, the discrimination between ''pathogen derived'' and ''host'' nucleic acids is of central importance.

In this review we will focus on the features of RNAs that are in place to discriminate between self and non-self nucleic acids. Viral nucleic acids can be sensed by their localization in nucleic acid free organelles (such as endosomes), and by their chemical modifications and secondary structures. Considering current knowledge, it appears that large parts of the cellular sensing mechanisms are targeting ''missing-self'' modifications rather than nucleic acids specific to a certain pathogen. This concept is reminiscent of the adaptive immune system that is also predominantly

Innate Immunity Laboratory, Max-Planck Institute of Biochemistry, Munich, Germany.
*Both these authors contributed equally to this work.

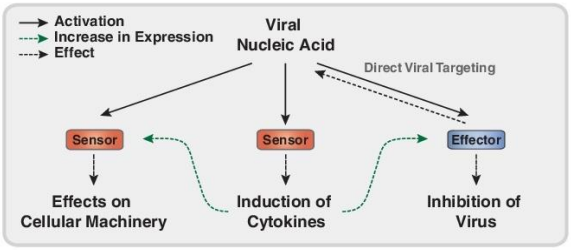


FIG. 1. Viral nucleic acids trigger a variety of events that are governed by a variety of specific cellular sensor proteins. Despite that these sensor proteins can identify the same type of viral nucleic acid, the antiviral and cellular effects are diverse. A key function of viral nucleic acids is the induction of cytokines, which regulate expression of many antiviral proteins, including sensor or effector proteins with affinity for the same viral nucleic acids. Engagement of these proteins with viral nucleic acid leads to changes in biological activity of cellular functions or in direct viral inhibition.

targeting missing-self structures and has proven to work well to protect against a wide range of invading pathogens.

Generation and Properties of Cellular RNA

To explain how intruding vRNAs are sensed by the innate immune system it is important to consider basic processes that are in place to generate cellular nucleic acids in higher eukaryotes. Generation of cellular RNAs is limited to the nucleus and mitochondria. The cytoplasm, however, is the compartment in which most RNAs are active and it is the site that is best surveyed by PRRs. Three cellular RNA polymerases (Pol-I, Pol-II, and Pol-III) are responsible for the generation of different types of RNAs. In a first step these polymerases generate precursor transcripts that carry a 5′ triphosphate (5′ PPP) group on their RNA (Fig. 2). However, before translocation to the cytoplasm, the 5′ end of these transcripts gets variably modified. In case of Pol-II transcripts, such as messenger RNA (mRNA) and most small nuclear RNAs (snRNAs), a guanosine nucleotide is coupled to the 5′ PPP group of the nascent RNA and forms the 5′ cap structure (McCracken and others 1997; Furuichi and Shatkin 2000). Moreover, the 5′ cap structure is further modified by a methylation mark at the N7-position of the guanosine (N7 methylation), the 2′O-position of the first (to generate Cap1 mRNA) and possibly also on the second ribose (Cap2 mRNA) (Byszewska and others 2014). All these modifications are co-transcriptionally added to the newly generated transcript (Topisirovic and others 2011) and are important for further processing and transport to the cytoplasm (Kohler and Hurt 2007; Muller-McNicoll and Neugebauer 2013). Lower eukaryotes, such as yeast, which appear not to use an IFN-based antiviral defense system, lack 2′O methylation on capped RNA (Byszewska and others 2014).

Ribosomal RNAs (rRNAs) are transcribed by Pol-I as a 45S-pre-rRNA with a 5′ PPP-RNA structure. After cleavage of the precursor into 28S, 18S, and 8S, rRNA, a 5′ monophosphate (5′ P) is obtained on the individual RNAs (Fig. 2) (Drygin and others 2010). Similarly, Pol-III generates a subset of additional RNAs comprising transfer RNA (tRNA), some small nucleolar RNA (snoRNAs), 5S rRNA, 7SK RNA, and U6 snRNA (Hopper 2013; Kirchner and Ignatova 2015). Like Pol-I transcribed rRNAs, tRNAs bear a 5′ P after cleavage of a 5′ oligonucleotide. U6 snRNA and 7SK RNA are not cleaved but bear a 5′ gamma-monomethyltriphosphate (5′ P_mPP-RNA) after processing (Singh and Reddy 1989).

In addition to the 5′ end modifications, host RNAs are highly modified on internal nucleotides (Sarin and Leidel 2014; Rosenthal 2015). However, the function of only a few modifications has been elucidated to date. Deamination of adenosines to inosines (A-to-I) by the RNA editing enzyme adenosine deaminase acting on RNA 1 (ADAR1), for instance, was recently shown to destabilize stem loop structures of Alu elements located in the 3′ untranslated region (UTR) of some mRNAs (George and others 2014). Repetitive Alu elements form double-stranded (ds) RNA structures that can stimulate IFN responses (Athanasiadis and others 2004; Levanon and others 2004). Deamination of such Alu elements in the 3′ UTR of mRNA leads to destabilization of dsRNA and reduced activation of the innate immune system (Hartner and others 2009). It still has to be determined how specificity to certain adenines is mediated and to what extent such deamination events also affect coding regions of proteins by accumulating mutations, potentially leading to protein misfolding. However, genetic evidence clearly shows that lack of ADAR1 is embryonically lethal in mice presumably due to elevated levels of dsRNA (Rice and others 2012). Deleting critical innate immune signaling molecules [eg, mitochondrial antiviral-signaling (MAVS) protein (also called IPS-1, VISA, or Cardif)] in these mice reduces this phenotype clearly suggesting that inability to deaminate RNAs on internal residues results in an IFN-dependent pathology (Liddicoat and others 2015; Pestal and others 2015; George and others 2016).

Another type of RNA that is prominently modified on internal residues are tRNAs. Internal modifications, such as ribose 2′O methylation on guanosine at position 18 and 34, have been shown to be important to dampen potential immune-stimulatory activity of tRNAs (Gehrig and others 2012; Jockel and others 2012; Kaiser and others 2014). It is quite likely that other not yet defined chemical modifications exist to prevent activation of the innate immune system.

Besides chemical modifications of RNA it is important to note that most RNAs are bound by RNA-binding proteins under steady-state conditions. Proteins associating to RNA can either contribute to activation or inhibition of the innate immune system. Viral proteins such as the E3L protein of Vaccinia virus, the nonstructural protein 1 (NS1) protein of Influenza A virus (IAV) or B2 of Flock house virus are binding dsRNA and can reduce the potential of stimulatory RNA likely by restricting accessibility to PRRs (Lingel and others 2005; Ayllon and Garcia-Sastre 2015). Similarly, the

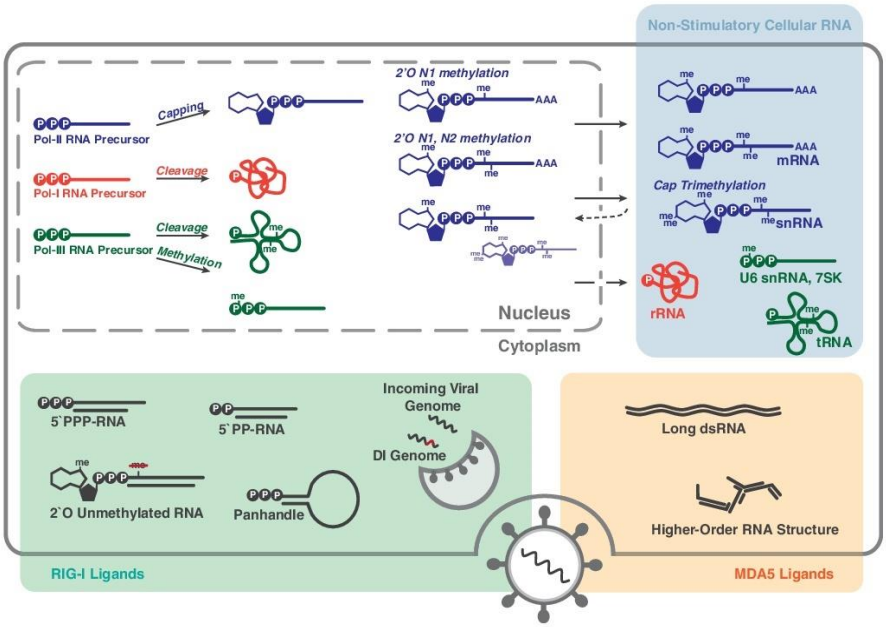


FIG. 2. Stimulatory and nonstimulatory RNAs in cells. This schematic provides an overview of the main proportion of cellular RNAs and cytoplasmic RIG-I and MDA5 stimulatory RNAs. Host RNAs are synthesized by RNA polymerase I–III (Pol-I, Pol-II, and Pol-III) in the nucleus. The RNAs are generated as a precursor RNA bearing a 5′ triphosphate group, which is extensively modified in the nucleus before getting transported to the cytoplasm where they perform their biological function. snRNAs are further modified in the cytoplasm and reimported to the nucleus. Nonstimulatory cytoplasmic cellular RNA (*blue box*) are not activating cytoplasmic PRRs. Specific viruses can introduce different types of stimulatory RNA during the infection process. These RNAs often resemble premature forms of cellular RNA and can be classified into RIG-I (*green box*) and MDA5 ligands (*orange box*). AAA, poly(A) tail; DI genomes, defective interfering genomes; dsRNA, double-stranded ribonucleic acid; MDA5, melanoma differentiation-associated protein 5; me, methyl group; mRNA, messenger RNA; P, phosphate group; PRR, pattern recognition receptor; RIG-I, retinoic acid-inducible gene I; rRNA, ribosomal RNA; snRNA, small nuclear RNA; tRNA, transfer RNA.

cellular RNA binding protein laboratory of genetics and physiology 2 (LGP2) reduces the stimulatory potential of certain RNAs (Venkataraman and others 2007), presumably by steric interference with activation of cellular PRRs (Yoneyama and others 2005; Saito and others 2007). However, LGP2 has also been shown to promote induction of IFN by a subset of viruses (Venkataraman and others 2007). Cellular RNA helicases can prepare viral ligands to be better sensed by effector proteins of the innate immune system (Ahmad and Hur 2015; Yao and others 2015). This is likely happening through displacement of other RNA binding proteins from the stimulatory RNA or due to changes in the secondary structure of the RNA. In recent years, a surprisingly high number of cellular helicases including DEAD-box protein (DDX) DDX1 (Zhang and others 2011a), DDX3 (Oshiumi and others 2010; Thulasi Raman and others 2016), DEAH-box helicase DHX9 (Zhang and others 2011b), DDX17 (Moy and others 2014), DDX60 (Miyashita and others 2011), and others, have been shown to be involved in induction of type I IFN. These proteins mostly do not directly bind to signaling molecules of the canonical IFN pathway but appear to have important accessory functions to properly activate the innate immune system. However, the contribution of RNA binding proteins to modulate innate immune responses can be diverse and needs to be characterized on an individual level.

Features of vRNA

Viruses often use relatively simple replication machineries resulting in RNAs that are only partially processed and therefore resemble a premature state of cellular RNAs. A number of PRRs evolved to sense such unprocessed RNA. The simplest evidence for this is the ability of cells to sense 5′ PPP-RNA, which constitutes the most basic product of RNA polymerases, through the cellular PRR retinoic acid-inducible gene I (RIG-I) (Hornung and others 2006; Pichlmair and others 2006). Delivery of 5′ PPP-RNA into the

cytoplasm activates type I IFN expression in an RIG-I-dependent manner. Many negative strand RNA viruses such as orthomyxo-, paramyxo-, and most bunyaviruses generate 5′ PPP-RNA constituting either the genomic RNA, replication by-products [eg, complementary RNA (cRNA)], or short subgenomic RNAs (Pichlmair and others 2009; Goubau and others 2013; Habjan and Pichlmair 2015). Interestingly, most viruses that generate 5′ PPP-RNA express auxiliary factors (eg, NS1 of IAV, V protein of Measles virus, etc.) that actively impair activation of the innate immune system (Versteeg and Garcia-Sastre 2010). Viruses that do not express such dedicated viral proteins, commonly process the 5′ RNA terminus to escape immune surveillance. Bornaviruses and a subset of Bunyaviruses including Crimean–Congo hemorrhagic fever virus, for instance, employ a 5′ trimming event as part of their replication strategy (Schneider and others 2007; Habjan and others 2008). Thereby the terminal nucleotides are cleaved by a viral nuclease leaving a 5′ P on the genomic RNA. Picorna- and caliciviruses shield their genomic 5′ PPP-end with the covalently linked VPg protein (Flanegan and others 1977; Lee and others 1977; Habjan and others 2008), which impairs binding of cellular PRRs to the 5′-end of vRNAs.

Viruses, such as flavi-, corona-, pox-, and reoviruses encode their own capping enzymes to generate 5′ capped mRNA (Decroly and others 2012). In addition to capping, these viruses independently evolved proteins that methylate the first ribose at the 2′O-position to generate RNA that resembles cellular mRNA modifications, indicating a strong selective pressure to favor this modification. Interestingly, viruses that lack the ability to methylate the 2′O position on the first ribose are sensed by the innate immune system (Daffis and others 2010; Habjan and others 2013; Schuberth-Wagner and others 2015). Orthomyxo- and bunyaviruses employ an alternative approach to gain a fully processed cap structure called ''cap-snatching.'' The first 10–13 nucleotides of a cellular mRNA are fused to the 5′ end of a vRNA, which thereby acquire cellular marks that allow evasion from host sensing mechanisms (Dias and others 2009; Decroly and others 2012; Reich and others 2014; Koppstein and others 2015).

In addition to 5′-end modifications vRNA has structural properties that allow discrimination from cellular RNAs. In particular, RNA double-strandedness of certain length is a feature that is often sensed by PRRs. Such long dsRNAs are generated by genome replication of dsRNA viruses or as replication by-products of many single-stranded (ss) RNA viruses. Even DNA viruses such as pox- and herpes viruses generate dsRNA through convergent transcription (Weber and others 2006; Feng and others 2012).

In the last years, genomic viral secondary structure elements generating portions of double-strandedness on ssRNA genomes like panhandle structures or stem loops have been shown to be recognized by antiviral mechanisms (Schlee and others 2009; Resa-Infante and others 2011; Moy and others 2014; Xu and others 2015). Although the underlying data are very compelling, it is still challenging to explain how the cell distinguishes viral from cellular dsRNA, particularly given the well-documented existence of natural dsRNA (Portal and others 2015) and high frequency of stem loops commonly found in RNAs of human origin. Different explanations for the increased immunogenicity of viral nucleic acids are possible: A certain threshold of dsRNA may have to be reached to activate the innate immune system. The abundance of vRNAs present in infected cells by far exceeds the abundance of cellular dsRNA. A threshold model is also supported by experiments using the poxvirus Modified Vaccinia Virus Ankara (MVA): Although poxviruses naturally generate dsRNA, a genetically engineered MVA strain that expresses increased levels of dsRNA shows superior immunogenicity (Wolferstatter and others 2014). Another explanation could be the subcellular localization of dsRNA since the majority of cellular dsRNA should remain in the nucleus or membrane covered cytoplasmic virus factories (Mackenzie 2005; Paul and Bartenschlager 2015), which cannot be surveyed by PRRs. Additionally, specific sequences associated to viral dsRNAs may contribute to signaling strength. Interestingly, the presence of double-strand portions within RNA is not always beneficial for virus sensing. Alphaviruses exhibit a secondary structural motif within the 5′ UTR, which prevents binding and activation of innate immune restriction proteins (Hyde and others 2014). Mechanistically, this high affinity dsRNA portion may not be accessible to cellular proteins and therefore most likely evolved as virus countermeasure against immune surveillance by the immune system (Hyde and others 2014).

Sensors of Viral or Nonprocessed RNA

A number of germline-encoded receptors have the ability to sense the presence of viral nucleic acids and initiate a variety of downstream events aiming at clearance of the invading pathogen. These receptors can be grouped in respect to their subcellular localization (endosomal or cytoplasmic), their ligand specificity (DNA, dsRNA, and ssRNA) and their effect after vRNA engagement (regulators of transcription, translation, or direct effect on vRNA). Here, we focus on cytoplasmic sensors of viral-derived RNAs and refer to other excellent reviews dealing with sensing in endosomes or of other viral ligands (O'Neill and others 2013; Cai and others 2014; Hornung and others 2014; Pelka and others 2016; Roers and others 2016).

RIG-I-Like Receptor-Mediated Recognition

The discovery of the PRR RIG-I by Takashi Fujita's laboratory opened a new era in the understanding of RNA virus sensing (Yoneyama and others 2004) and led to the identification of a class of receptors named RIG-I-like receptors (RLRs) (Onoguchi and others 2011). The highly conserved family of RLRs is comprised of three structurally related proteins named RIG-I, melanoma differentiation-associated protein-5 (MDA5), and LGP2 (Yoneyama and others 2005; Onoguchi and others 2011). All members belong to the Asp-Glu-Ala-Asp (DEAD) box family and consist of a RNA helicase domain with ATPase activity and a C-terminal domain (CTD), which is important to mediate specificity to virus-derived nucleic acid sensing (Takahasi and others 2008). RIG-I and MDA5 accommodate two N-terminal caspase activation and recruitment domains (CARDs), which initiate downstream signaling through CARD-dimerization with the MAVS protein.

Genetic deletion of RIG-I and MDA5 showed that RIG-I and MDA5 exhibit specificity for certain viruses indicating that distinct nucleic acids are sensed by these receptors. RIG-I is able to detect viruses of the rhabdoviridae family

such as Vesicular stomatitis virus and viruses of the paramyxoviridae family including Newcastle disease virus, Sendai virus (SeV), and Measles virus. MDA5 detects viruses of the Picornaviridae family such as Encephalomyocarditis virus and Polio virus (Gitlin and others 2006; Kato and others 2006). Some flaviviruses (eg, Dengue virus, West Nile virus, and Semliki forest virus) activate both, RIG-I and MDA5 (Fredericksen and others 2008; Akhrymuk and others 2016).

In contrast to RIG-I and MDA5, the role of LGP2 is less well understood particularly since this protein is missing a CARD. Depending on the experimental system, it was shown that LGP2 could serve as a negative regulator of RLR signaling (Saito and others 2007; Venkataraman and others 2007) or activate induction of IFN (Venkataraman and others 2007). However, more recent studies strengthen the finding that LGP2 operates as a positive regulator of MDA5 (Satoh and others 2010; Bruns and others 2014; Uchikawa and others 2016).

An important property of all RLRs is their ability to interact with dsRNA. However, although it became evident that binding to dsRNA is important to activate RLRs, RNA double-strandedness is not always sufficient to induce signaling. Additional requirements on stimulatory RNA, particularly chemical modifications, may serve as safeguard to reduce accidental activation of type I IFN signaling, particularly if the double-stranded stretch on RNA is only short.

Activation and Downstream Signaling of RLRs

Activation of RIG-I requires ligand binding, a cascade of regulatory post-translational modifications and binding of proteins resulting in the exposure of the CARDs. In uninfected cells, RIG-I CARDs and the CTD are constitutively phosphorylated at Ser8 and Thr170 by Protein Kinase C alpha/beta and Casein Kinase II (CKII) (Sun and others 2011; Maharaj and others 2012). In the phosphorylated state RIG-I adopts a ''closed'' conformation sequestering the CARDs from signaling due to interactions with the C-terminal repressor domain. The ATPase activity of the helicase domain allows RIG-I to constantly scan RNAs for the presence of viral motifs and was shown to be a key element to prevent recognition of self-RNA (Lassig and others 2015). ATP hydrolysis facilitates the release of RIG-I from self-RNA, while presence of viral motifs detected through the CTD leads to stalling and activation of RIG-I. Subsequently, the constitutive Ser/Thr phosphorylation sites in the CARDs are removed by serine/threonine-protein phosphatase α and γ (PP1α and PP1γ) resulting in Riplet-mediated ubiquitination of the CTD. This is followed by dimerization and a conformational rearrangement of RIG-I leading to the exposure of CARDs (Oshiumi and others 2013; Peisley and others 2013), binding of ubiquitin/ISG15-conjugating enzyme (TRIM25) to CARD1, and subsequent K63-linked ubiquitination in CARD2 (Gack and others 2007). After activation, the CARDs of RIG-I form a helical tetramer in a lock washer conformation. Multiple lock washer tetramers form a helical trajectory that allows MAVS filament assembly along this structure (Wu and others 2014). This results in clustering of multiple MAVS molecules and activation of downstream signaling.

In contrast to RIG-I, MDA5 oligomerizes as filaments in a head-to-tail fashion along dsRNA to reach high-affinity interactions with long dsRNA ligands (Peisley and others 2011; Berke and others 2012). After oligomerization, the CARDs of MDA5 point outward of the filament and oligomerize in structures that can bind to MAVS and activate downstream signaling (Wu and others 2013). Interestingly, MDA5 activation is significantly increased by LGP2, which does not contain a CARD itself. Structural analysis showed that LGP2 binds the end of dsRNAs (very much like RIG-I) and thereby initiates MDA5 filament oligomerization on dsRNA (Uchikawa and others 2016). LGP2 and MDA5 have been shown to bind similar stimulatory RNAs, which is in line with cooperative activity of LGP2 in MDA5 activation. LGP2-precipitated RNAs induce IFN-α/β in a MDA5-dependent manner providing a functional link between LGP2 and MDA5 (Deddouche and others 2014).

A key uniting feature of all RLRs is signaling through MAVS. Under physiological conditions, MAVS is kept inactive by an autoinhibitory mechanism. Upon binding to oligomerized CARDs of RLRs, the regions responsible for downstream signaling including TANK-binding kinase 1 (TBK1)/Interferon regulatory factor 3 (IRF3) and IκB kinase (IKK)/nuclear factor kappa-light-chain-enhancer of activated B cells (NF-κB) activation sites are exposed and induce signaling events that result in expression of type I IFNs and proinflammatory cytokines (Shi and others 2015).

Sensing of 5′ Modifications by RIG-I

RIG-I activation requires two features that co-occur on one RNA molecule. The first essential feature for RIG-I activation is a chemical modification on the 5′ terminus of the RNA. The best described 5′ modifications that activate RIG-I are 5′ tri- and 5′ di-phosphates (Hornung and others 2006; Pichlmair and others 2006; Goubau and others 2014). However, more recently it has been shown that RIG-I is also activated by capped RNA that lack a methylation mark at the 2′O ribose position of the first nucleotide (Schuberth-Wagner and others 2015; Devarkar and others 2016). A conserved residue (Histidine 830) within the CTD of RIG-I sterically prevents binding of cellular 2′O-methylated RNA, and therefore serves as molecular gatekeeper to pertain activation by cellular mRNAs. Silencing of the endogenous cap-specific mRNA (nucleoside-2′-O-)-methyltransferase 1 (MTr1) converts nonstimulatory into stimulatory mRNA and triggers a spontaneous RIG-I-dependent type I IFN response (Schuberth-Wagner and others 2015). The second essential feature for RIG-I activation is a short stretch of blunt-ended dsRNA. While chemically synthesized 5′ PPP-ssRNA is not sufficient to prominently activate RIG-I, 5′ PPP-dsRNA molecules with the same sequence show very strong activity (Schlee and others 2009; Schmidt and others 2009). Reports regarding the minimum length of this dsRNA stretch range from at least 10 base pairs (bp) (Schmidt and others 2009; Kohlway and others 2013) to 19 bp (Schlee and others 2009). Experiments transfecting differently modified RNA clearly show that 3′ overhangs at the 5′-end are sufficient to impair activation of RIG-I (Schlee and others 2009) and may even serve as dominant negative decoy substrate (Marq and others 2011).

Besides dsRNA with 5′ PPP modifications, short 5′ hydroxyl (5′ OH) and 3′ monophosphoryl (3′ P) dsRNA cleavage products of 2-5A-dependent ribonuclease (RNaseL) have been proposed to serve as RIG-I ligands (Malathi

and others 2007, 2010). Although this notion is supported by many functional data, the exact mechanism of RIG-I activation is not clear to date, particularly since this appears to happen in a cell type-dependent manner (Banerjee and others 2014).

Sensing of RNA in Virus-Infected Cells

Despite formidable progress that allowed defining optimal activation of RIG-I by synthetic ligands much less is known about the nature of the physiological ligand generated during virus infections. Many elegant studies aimed to define the physiological RIG-I stimulus. Compelling evidence suggests that viral genomic RNA delivered by virus infection activates RIG-I. Here, we discuss data that support this notion and also consider questions that still remain to be answered. Evidence for vRNA being the physiological ligand for RIG-I was already provided by Isaacs and Lindenmann (1957), who found that cells treated with high amounts of heat-inactivated viruses activate type I IFN. More recently this was supported by data that show that delivery of replication incompetent viral particles activate RIG-I (Weber and others 2013). Furthermore, shortly after infection, vRNA of infectious viral particles closely associates with mitochondria, which are serving as signaling hubs for the induction of type I IFN (Liedmann and others 2014). However, the magnitude of an antiviral response triggered by incoming vRNA appears to be relatively low and viral replication appears to be required to elicit high amounts of IFN (Crotta and others 2013; Killip and others 2014). One reason could be insufficient abundance of stimulatory RNAs, which questions whether incoming viral nucleic acids are significantly contributing to type I IFN production under physiological conditions.

Genomic vRNA isolated from virus particles and transfected into cells potently stimulates RIG-I (Hornung and others 2006; Pichlmair and others 2006). This is in agreement with the notion that vRNA isolated from IAV particles spontaneously forms the so-called panhandle structure. This structure is formed due to base-pairing of terminal ssRNA sequences of many negative ssRNA viruses and serves as promoter for the viral polymerase complex (Hsu and others 1987; Fodor and others 1994; Tiley and others 1994). The panhandle structure resembles *in vitro* synthesized blunt ended 5′ PPP RNA. In line with RIG-I activation by synthetic ligands, the panhandle thus constitutes a perfect stimulus for RIG-I activation and explains the very strong stimulatory activity of isolated vRNA. Interestingly, mismatches in the panhandle structure of some IAV strains disrupt RNA complementarity and results in reduced activation of RIG-I (Anchisi and others 2016). Such an adaptation may therefore represent a viral strategy to evade RIG-I activation (Anchisi and others 2016). Supporting evidence that viral genomic RNA stimulates RIG-I comes from experiments using minireplicon systems. vRNA of defined length generated by the IAV polymerase complex activates RIG-I (Rehwinkel and others 2010). Interestingly, PPP-RNA is not only generated by viruses but also by bacteria. mRNA in bacteria is not capped and intracellular bacteria such as Listeria and Legionella have been shown to activate RIG-I (Monroe and others 2009; Abdullah and others 2012).

Although the notion that genomic vRNA is the major ligand activating RIG-I in virally infected cells is elegant, a number of additional aspects linked to the replication process of viruses complicate this simple model: During viral replication, vRNA is constantly bound by viral proteins, which theoretically prevent activation of RIG-I. In case of many negative strand RNA viruses the viral polymerase complex is located at the 5′-end of vRNAs potentially shielding the terminal PPP group from being sensed by RIG-I. Indeed, for IAV it has been shown that association of the polymerase complex to the vRNA polymerase protects from innate sensing by RIG-I (Weber and others 2013; Liedmann and others 2014). Variants of the viral polymerase complex featuring weaker affinity for viral genomic RNAs have reduced ability to impair activation of the innate immune system (Weber and others 2013) suggesting that under certain conditions vRNAs can be a physiological activator of RIG-I and that the presence of proteins actively impair RIG-I activation. Besides a potentially inaccessible 5′-end, the relative contribution of the panhandle structure to provide a dsRNA platform for RIG-I binding is not so clear: Although isolated genomic RNA of IAV clearly generates blunt double-stranded ends, the structure of the panhandle bound to the viral polymerase complex, as it exists in virally infected cells, adopts a partially single-stranded conformation (Tiley and others 1994). This notion is supported by low affinity binding of the negative strand RNA virus polymerases to the double-stranded panhandle structure compared to binding to ssRNA (Pflug and others 2014; Gerlach and others 2015). Thus, the panhandle structure most likely either exists as a ''corkscrew'' (Neumann and Hobom 1995; Flick and others 1996) or as a ''fork'' (Fodor and others 1994, 1995; Kim and others 1997) structure in virally infected cells. Both structures presumably only allow suboptimal activation of RIG-I and it is therefore still not finally solved what type of RNA prominently activates RIG-I in virus infected cells.

A possibility is the activation of RIG-I by viral replication intermediates or by defective interfering (DI) particles, which are commonly generated during infections with viruses including SeV, IAV, and Measles virus (Strahle and others 2006; Baum and others 2010; Runge and others 2014). Although DI particles are not replicating, they prominently activate the innate immune system (Killip and others 2012).

Besides generating stimulatory nucleic acids from viral templates, RNA-dependent RNA polymerases (RdRP) also have the ability to use cellular RNA as templates resulting in the generation of host-derived stimulatory RNAs. Expression of Simian foamy virus RdRP, for instance, leads to induction of IFN-β in the absence of viral templates (Nikonov and others 2013). In addition, expression of RdRP in transgenic animals induces a constant IFN response that leads to increase virus resistance in these animals (Painter and others 2015). The ability of viral polymerases to generate cellular RNA copies with IFN inducing capability indicates that RdRP transcripts feature stimulatory modifications rather than the actual sequence being sensed by innate immune sensors. Interestingly, the ability to generate stimulatory RNA from cellular templates has so far only been shown for RdRPs that are active in the cytoplasm. A possible explanation for this could be the absence of RNA-modifying enzymes in the cytoplasm, which does not allow RNA processing to minimize the stimulatory potential of RdRP generated RNA.

Sensing of dsRNAs by MDA5

In contrast to ligands for RIG-I, requirements for MDA5 activation are less well understood. MDA5 recognizes long dsRNA (greater than ~0.5 kB) that is normally not present in the cytoplasm of uninfected cells (Kato and others 2008). A commonly used synthetic activator for MDA5 is poly-I:C, which is generated through annealing of enzymatically synthesized poly-I and poly-C homopolymers of undefined length. Although poly-I:C is regarded as dsRNA analog, the structure of poly-I:C is most likely not a uniform double-strand but may rather adopt a web-like structure (Pichlmair and others 2009). Interestingly, similarly synthesized poly-A:U or poly-G:C have very little stimulatory potential suggesting features associated with poly-I:C that are not yet well understood.

The best characterized viruses leading to MDA5 activation are picornaviruses [eg, Encephalomyocardidits virus (EMCV), Theiler's murine encephalitis virus, Poliovirus, or Norovirus] featuring a positive ssRNA genome (Gitlin and others 2006; Kato and others 2006; Loo and others 2008; McCartney and others 2008). While RIG-I activating viruses have the ability to induce PRR signaling through viral genome recognition, activation of MDA5 by EMCV strictly requires transcription of the viral genome and generation of a dsRNA intermediate resulting in a 7.5 kb replicative form of EMCV (Feng and others 2012; Triantafilou and others 2012). Furthermore, it is likely that additional replication intermediates representing higher order structural RNAs are generated and sensed by MDA5 (Pichlmair and others 2009).

Unbiased Approaches to Identify PRR-Associated vRNA

More recently, next generation sequencing was used to characterize the RNA ligand bound to MDA5 or RIG-I. An issue with such approaches is that the helicase domain of RLRs can generally associate with dsRNA and that this affinity does not directly lead to activation of RIG-I or MDA5. However, next generation sequencing helped to identify commonalities and differences in RNA binding between MDA5 and RIG-I. In Measles virus-infected cells, for instance, both, MDA5 and RIG-I preferentially bind to viral AU-rich sequences, particularly in the Measles virus L-region (Runge and others 2014). However, MDA5 shows superior enrichment for the Measles virus (+) sense RNA, while RIG-I preferentially bind to RNA of negative polarity. In a similar approach, Sanchez David and colleagues used next-generation sequencing analysis to investigate RNA precipitated with RIG-I, MDA5, and LGP2 in Measles and Chikungunya virus-infected cells. The authors showed that each of the RLRs binds distinct regions of the viral genome. RIG-I bound specifically the 3′ UTR of the Chikungunya virus genome and DI genomes of Measles virus, whereas, MDA5 and LGP2 sensed nucleoprotein coding regions of Measles virus (Sanchez David and others 2016). The shared RNAs targeted by MDA5 and LGP2 strongly support a functional relationship between these two PRRs, which is also confirmed by alternative functional and structural approaches (Goubau and others 2014) (Uchikawa and others 2016).

In sum, it is evident that distinct RNA species trigger RIG-I and MDA5 activation. These RNA species may even be expressed at different time points during the infection process. Indeed, it was shown that vRNA products generated at different time points after West Nile virus infection sequentially stimulate RIG-I and MDA5, whereby, RIG-I triggers an early and MDA5 a later response (Errett and others 2013).

Direct Effectors of Viral Nucleic Acids

Besides RLRs, which have the ability to sense nucleic acids and regulate transcriptional programs, a set of additional proteins exist that are binding to specific viral nucleic acids (Habjan and Pichlmair 2015). These proteins either have the ability to regulate cellular processes that are unrelated to transcription or directly bind and thereby impair the activity of viral nucleic acids. Often these additional vRNA binding proteins are inducible by type I IFNs underlining their involvement in antiviral processes. This concept of multiple cellular proteins associating to a given type of stimulatory RNA is best illustrated by dsRNA binding proteins (Fig. 3). dsRNA activates the PRR MDA5, which is activating a transcriptional program culminating in expression of type I IFN. Besides MDA5, dsRNA-dependent Protein Kinase R (PKR) and 2′-5′-oligoadenylate synthetase (OAS1) directly bind dsRNAs. PKR regulates a plethora of antiviral processes after binding to dsRNA including inhibition of translation through continuous phosphorylation of eukaryotic initiation factor 2 alpha (eIF2-a), induction of apoptosis and autophagy (Kang and Tang 2012), and activation of NF-kB through interacting with the IKK complex (Zamanian-Daryoush and others 2000). PKR has additionally been linked to regulation of type I IFN expression (Diebold and others 2003; Gilfoy and Mason 2007). While this accessory function of PKR is dispensable for RIG-I-dependent responses, IFN induction by viruses that activate MDA5 appears to be critically relying on additional activity by PKR (Schulz and others 2010; Wolferstatter and others 2014; Pham and others 2016). After binding to dsRNA OAS1 generates 2′-5′-linked oligoadenylates (2′-5′-OA) that serve as second messengers and have the ability to activate the latent RNAseL. Only after activation, RNAseL cleaves RNA eventually resulting in cell death (Chakrabarti and others 2011). Furthermore, vRNAs bind to a number of DEAD- and DEAH-box helicases, which often appear to have auxiliary functions to regulate IFN-α/β expression and also bear direct virus inhibitory function. DDX3, for instance, has been shown to be important for activation of the IFN signaling pathway (Oshiumi and others 2010; Gu and others 2013). DDX3 was also shown to restrict HBV replication and, thereby, acting as an antiviral effector protein (Ko and others 2014).

As for dsRNA, 5′ PPP-RNA, which activates RIG-I, can associate with additional cellular proteins. Unbiased AP-MS experiments and follow-up studies using mutational approaches identified the interferon-induced protein with tetratricopeptide repeats (IFIT) 1 and -5, which directly associate with 5′ PPP-RNA (Pichlmair and others 2011; Fensterl and Sen 2015). Depending on the exact experimental setup, IFIT1-deficient mice appear to be more susceptible to 5′ PPP-RNA generating viruses such as Vesicular stomatitis virus (Pichlmair and others 2011). However, while RIG-I activation requires dsRNA, IFIT proteins specifically bind only ssRNAs in a helical positively charged binding cleft (Abbas and others 2013). Besides associating with 5′ PPP-RNA, IFIT1 has the ability to directly associate with capped RNA that lacks a methylation mark at the first ribose (2′O N1 unmethylated RNA) (Habjan and others 2013; Kumar and

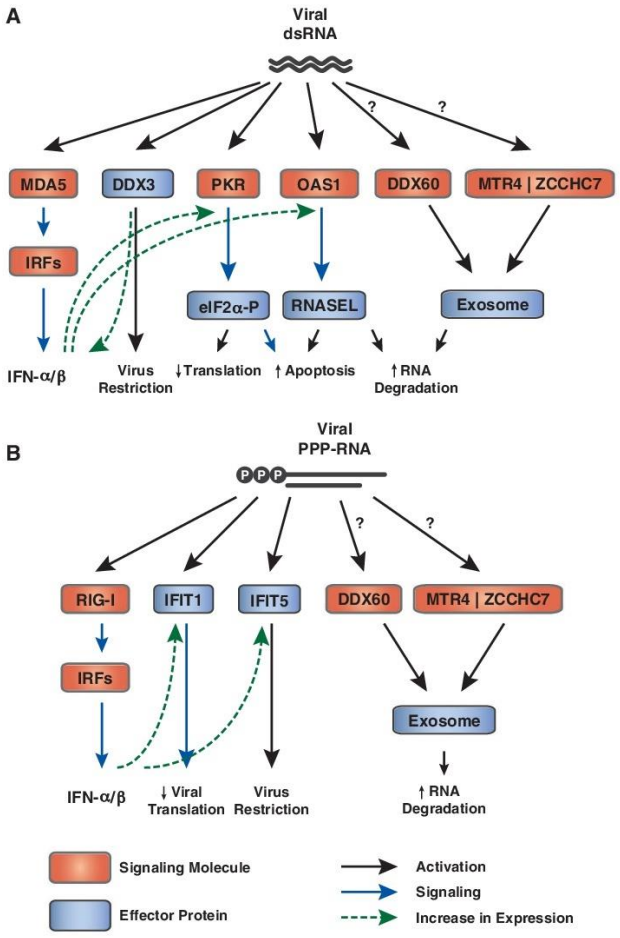


FIG. 3. Cellular sensor and effector proteins binding viral nucleic acids. **(A)** Sensing of dsRNAs by MDA5 results in expression of IFN-α/β. These in turn upregulate additional sensors including dsRNA-dependent PKR and OAS1. Binding of PKR to dsRNA phosphorylates the translation initiation factor eIF2-leading to an inhibition of translation and an induction of apoptosis. OAS1 synthesizes 2′-5′-oligoadenylates activating RNASEL. Activation of RNASEL results in RNA degradation and apoptosis. DDX60, ATP-dependent RNA helicase SKIV2L2 (MTR4) and ZCCHC7 promote vRNA degradation via the exosome complex upon virus infection. DDX3 activates IFN signaling and restricts virus replication. **(B)** Engagement of 5′-triphosphorylated-RNA by RIG-I leads to the expression of type I IFNs, thereby inducing the expression of the effector proteins IFIT1 and IFIT5. Sensing of PPP-RNA by IFIT1 and IFIT5 leads to decreased viral translation and restriction of virus, respectively. vRNA can be recognized and degraded by the exosome cofactors DDX60, MTR4, and ZCCHC7. DDX3, DEAD box protein 3; DDX60, DEAD box protein 60; eIF2, eukaryotic initiation factor 2; IFIT, interferon-induced protein with tetratricopeptide repeats; IFN, interferon; IRFs, interferon regulatory factors; MTR4, ATP-dependent RNA helicase SKIV2L2; OAS1, 2′-5′-oligoadenylate synthetase 1; PKR, double-stranded RNA-dependent protein Protein Kinase R; PPP-RNA, triphosphorylated RNA; vRNA, viral RNA; ZCCHC7, Zinc finger CCHC domain-containing protein 7.

others 2014). Viruses that fail to methylate the 2′O position of the first ribose are attenuated in wild-type mice but are highly pathogenic in IFIT-deficient animals (Daffis and others 2010; Leung and Amarasinghe 2016). It is yet not entirely clear how IFIT1 and -5 impair virus growth but the high amounts of IFIT proteins expressed after IFN treatment suggests stoichiometric interference with viral nucleic acids rather than enzymatic activity of IFIT proteins (Habjan and others 2013; Kumar and others 2014). Pathogenic alphaviruses that generate high affinity dsRNA secondary structures on the 5′-end of their genome evade surveillance by IFIT proteins (Hyde and others 2014). IFIT1 appears to show surprisingly little efficiency against negative strand RNA viruses *in vivo* despite that these viruses are known to generate 5′ PPP-RNA (Pinto and others 2015). Potential explanations may be evasion strategies of these viruses, including replication in the nucleus that is not surveyed by IFIT proteins (eg, for IAV) or generation of secondary structures of the RNA 5′-end as has been shown for alphaviruses (Pinto and others 2015). IFITs have also been proposed to be (Berchtold and others 2008; Zhang and others 2013; Imaizumi and others 2016) or not to be (Pichlmair and others 2011; Habjan and others 2013) involved in regulation of antiviral gene expression. Thus, the activity of IFIT proteins is not yet fully understood and highlights the importance of functional studies that will give further mechanistic insights.

While the induction of type I IFNs and cellular restriction mechanisms are partially well understood, relatively little is known about the cellular machinery that specifically degrades vRNA. A prominent machinery responsible

for RNA degradation is the exosome, a large molecular complex with 3′-5′ exonuclease activity. However, this complex requires accessory factors that mark RNA for degradation. Recently, the Ski2-like protein DDX60 was identified as cellular protein that promotes degradation of vRNA of diverse viruses through the exosome (Oshiumi and others 2015). Furthermore, DDX60 has been proposed to support RIG-I and MDA5-dependent induction of type I IFNs (Oshiumi and others 2015). However, the role of DDX60 appears complex since other laboratories found little effect of DDX60 in antiviral immunity (Goubau and others 2015; Grunvogel and others 2015). More recently, superkiller viralicidic activity 2-like 2 (SKIV2L2; also called Mtr4) and Zinc finger CCHC domain-containing protein 7 (ZCCHC7), components of the Trf4/Air2/Mtr4p polyadenylation (TRAMP) complex, were identified to colocalize with vRNA and the exosome in the cytoplasm of infected cells (Molleston and others 2016) and the helicase SKI2W (SKIV2L) RNA exosome has been shown to prevent autoimmunity by regulating the abundance of RIG-I ligands (Eckard and others 2014) suggesting that vRNAs are specifically targeted for decay.

Concluding Remarks

The knowledge on virus sensing and restriction has dramatically increased in recent years. It became evident that the innate immune system particularly senses RNA that is insufficiently processed and therefore lacks motifs commonly found on cellular RNA. The ability of cytoplasmic PRRs to sense such missing-self motifs allows them to be broadly active and to detect nucleic acids generated by viruses and other pathogens such as bacteria. In case of failure to properly process cellular RNA this ability also bears the risk of inducing unwanted immune responses that can lead to autoimmune disorders. Besides nucleic acid sensors, a set of cellular proteins exists that directly restricts viruses or leads to changes in cellular machineries ranging from translational control to cell death. More mechanistic insights into the regulation and function of nucleic acid binding proteins are important to understand antiviral immunity and to exploit this knowledge for therapeutic interventions.

Acknowledgments

We want to thank the members of the innate immunity laboratory, particularly Philipp Hubel and Pietro Scaturro, for input and critical discussions. The work in the authors' laboratory is funded by a Max-Planck free floater program, ERC starting grant (StG 311339, iVIP), the German Research Foundation (PI1084/2-1, PI1084/3-1, and TRR179), and the Federal Ministry for Education and Research (ERA-Net grant ERASe).

Author Disclosure Statement

No competing financial interests exist.

References

Abbas YM, Pichlmair A, Gorna MW, Superti-Furga G, Nagar B. 2013. Structural basis for viral 5′-PPP-RNA recognition by human IFIT proteins. Nature 494(7435):60–64.

Abdullah Z, Schlee M, Roth S, Mraheil MA, Barchet W, Bottcher J, Hain T, Geiger S, Hayakawa Y, Fritz JH, Civril F, Hopfner KP, Kurts C, Ruland J, Hartmann G, Chakraborty T, Knolle PA. 2012. RIG-I detects infection with live Listeria by sensing secreted bacterial nucleic acids. EMBO J 31(21): 4153–4164.

Ahmad S, Hur S. 2015. Helicases in antiviral immunity: dual properties as sensors and effectors. Trends Biochem Sci 40(10):576–585.

Akhrymuk I, Frolov I, Frolova EI. 2016. Both RIG-I and MDA5 detect alphavirus replication in concentration-dependent mode. Virology 487:230–241.

Anchisi S, Guerra J, Mottet-Osman G, Garcin D. 2016. Mismatches in the influenza A virus RNA panhandle prevent retinoic acid-inducible gene I (RIG-I) sensing by impairing RNA/RIG-I complex formation. J Virol 90(1):586–590.

Athanasiadis A, Rich A, Maas S. 2004. Widespread A-to-I RNA editing of Alu-containing mRNAs in the human transcriptome. PLoS Biol 2(12):e391.

Ayllon J, Garcia-Sastre A. 2015. The NS1 protein: a multitasking virulence factor. Curr Top Microbiol Immunol 386: 73–107.

Banerjee S, Chakrabarti A, Jha BK, Weiss SR, Silverman RH. 2014. Cell-type-specific effects of RNase L on viral induction of beta interferon. MBio 5(2):e00856-14.

Baum A, Sachidanandam R, Garcia-Sastre A. 2010. Preference of RIG-I for short viral RNA molecules in infected cells revealed by next-generation sequencing. Proc Natl Acad Sci U S A 107(37):16303–16308.

Berchtold S, Manncke B, Klenk J, Geisel J, Autenrieth IB, Bohn E. 2008. Forced IFIT-2 expression represses LPS induced TNF-alpha expression at posttranscriptional levels. BMC Immunol 9:75.

Berke IC, Yu X, Modis Y, Egelman EH. 2012. MDA5 assembles into a polar helical filament on dsRNA. Proc Natl Acad Sci U S A 109(45):18437–18441.

Bruns AM, Leser GP, Lamb RA, Horvath CM. 2014. The innate immune sensor LGP2 activates antiviral signaling by regulating MDA5-RNA interaction and filament assembly. Mol Cell 55(5):771–781.

Byszewska M, Smietanski M, Purta E, Bujnicki JM. 2014. RNA methyltransferases involved in 5′ cap biosynthesis. RNA Biol 11(12):1597–1607.

Cai X, Chiu YH, Chen ZJ. 2014. The cGAS-cGAMP-STING pathway of cytosolic DNA sensing and signaling. Mol Cell 54(2):289–296.

Chakrabarti A, Jha BK, Silverman RH. 2011. New insights into the role of RNase L in innate immunity. J Interferon Cytokine Res 31(1):49–57.

Crotta S, Davidson S, Mahlakoiv T, Desmet CJ, Buckwalter MR, Albert ML, Staeheli P, Wack A. 2013. Type I and type III interferons drive redundant amplification loops to induce a transcriptional signature in influenza-infected airway epithelia. PLoS Pathog 9(11):e1003773.

Daffis S, Szretter KJ, Schriewer J, Li J, Youn S, Errett J, Lin TY, Schneller S, Zust R, Dong H, Thiel V, Sen GC, Fensterl V, Klimstra WB, Pierson TC, Buller RM, Gale M, Jr, Shi PY, Diamond MS. 2010. 2′-O methylation of the viral mRNA cap evades host restriction by IFIT family members. Nature 468(7322):452–456.

Decroly E, Ferron F, Lescar J, Canard B. 2012. Conventional and unconventional mechanisms for capping viral mRNA. Nat Rev Microbiol 10(1):51–65.

Deddouche S, Goubau D, Rehwinkel J, Chakravarty P, Begum S, Maillard PV, Borg A, Matthews N, Feng Q, van Kuppe-

veld FJ, Reis e Sousa C. 2014. Identification of an LGP2-associated MDA5 agonist in picornavirus-infected cells. Elife 3:e01535.
Devarkar SC, Wang C, Miller MT, Ramanathan A, Jiang F, Khan AG, Patel SS, Marcotrigiano J. 2016. Structural basis for m7G recognition and 2′-O-methyl discrimination in capped RNAs by the innate immune receptor RIG-I. Proc Natl Acad Sci U S A 113(3):596–601.
Dias A, Bouvier D, Crepin T, McCarthy AA, Hart DJ, Baudin F, Cusack S, Ruigrok RW. 2009. The cap-snatching endonuclease of influenza virus polymerase resides in the PA subunit. Nature 458(7240):914–918.
Diebold SS, Montoya M, Unger H, Alexopoulou L, Roy P, Haswell LE, Al-Shamkhani A, Flavell R, Borrow P, Reis e Sousa C. 2003. Viral infection switches non-plasmacytoid dendritic cells into high interferon producers. Nature 424(6946):324–328.
Drygin D, Rice WG, Grummt I. 2010. The RNA polymerase I transcription machinery: an emerging target for the treatment of cancer. Annu Rev Pharmacol Toxicol 50:131–156.
Eckard SC, Rice GI, Fabre A, Badens C, Gray EE, Hartley JL, Crow YJ, Stetson DB. 2014. The SKIV2L RNA exosome limits activation of the RIG-I-like receptors. Nat Immunol 15(9):839–845.
Errett JS, Suthar MS, McMillan A, Diamond MS, Gale M, Jr. 2013. The essential, nonredundant roles of RIG-I and MDA5 in detecting and controlling West Nile virus infection. J Virol 87(21):11416–11425.
Feng Q, Hato SV, Langereis MA, Zoll J, Virgen-Slane R, Peisley A, Hur S, Semler BL, van Rij RP, van Kuppeveld FJ. 2012. MDA5 detects the double-stranded RNA replicative form in picornavirus-infected cells. Cell Rep 2(5):1187–1196.
Fensterl V, Sen GC. 2015. Interferon-induced Ifit proteins: their role in viral pathogenesis. J Virol 89(5):2462–2468.
Flanegan JB, Petterson RF, Ambros V, Hewlett NJ, Baltimore D. 1977. Covalent linkage of a protein to a defined nucleotide sequence at the 5′-terminus of virion and replicative intermediate RNAs of poliovirus. Proc Natl Acad Sci U S A 74(3):961–965.
Flick R, Neumann G, Hoffmann E, Neumeier E, Hobom G. 1996. Promoter elements in the influenza vRNA terminal structure. RNA 2(10):1046–1057.
Fodor E, Pritlove DC, Brownlee GG. 1994. The influenza virus panhandle is involved in the initiation of transcription. J Virol 68(6):4092–4096.
Fodor E, Pritlove DC, Brownlee GG. 1995. Characterization of the RNA-fork model of virion RNA in the initiation of transcription in influenza A virus. J Virol 69(7):4012–4019.
Fredericksen BL, Keller BC, Fornek J, Katze MG, Gale M, Jr. 2008. Establishment and maintenance of the innate antiviral response to West Nile Virus involves both RIG-I and MDA5 signaling through IPS-1. J Virol 82(2):609–616.
Furuichi Y, Shatkin AJ. 2000. Viral and cellular mRNA capping: past and prospects. Adv Virus Res 55:135–184.
Gack MU, Shin YC, Joo CH, Urano T, Liang C, Sun L, Takeuchi O, Akira S, Chen Z, Inoue S, Jung JU. 2007. TRIM25 RING-finger E3 ubiquitin ligase is essential for RIG-I-mediated antiviral activity. Nature 446(7138):916–920.
Gehrig S, Eberle ME, Botschen F, Rimbach K, Eberle F, Eigenbrod T, Kaiser S, Holmes WM, Erdmann VA, Sprinzl M, Bec G, Keith G, Dalpke AH, Helm M. 2012. Identification of modifications in microbial, native tRNA that suppress immunostimulatory activity. J Exp Med 209(2):225–233.
George CX, John L, Samuel CE. 2014. An RNA editor, adenosine deaminase acting on double-stranded RNA (ADAR1). J Interferon Cytokine Res 34(6):437–446.
George CX, Ramaswami G, Li JB, Samuel CE. 2016. Editing of cellular self-RNAs by adenosine deaminase ADAR1 suppresses innate immune stress responses. J Biol Chem 291(12):6158–6168.
Gerlach P, Malet H, Cusack S, Reguera J. 2015. Structural insights into bunyavirus replication and its regulation by the vRNA promoter. Cell 161(6):1267–1279.
Gilfoy FD, Mason PW. 2007. West Nile virus-induced interferon production is mediated by the double-stranded RNA-dependent protein kinase PKR. J Virol 81(20):11148–11158.
Gitlin L, Barchet W, Gilfillan S, Cella M, Beutler B, Flavell RA, Diamond MS, Colonna M. 2006. Essential role of mda-5 in type I IFN responses to polyriboinosinic:polyribocytidylic acid and encephalomyocarditis picornavirus. Proc Natl Acad Sci U S A 103(22):8459–8464.
Goubau D, Deddouche S, Reis e Sousa C. 2013. Cytosolic sensing of viruses. Immunity 38(5):855–869.
Goubau D, Schlee M, Deddouche S, Pruijssers AJ, Zillinger T, Goldeck M, Schuberth C, Van der Veen AG, Fujimura T, Rehwinkel J, Iskarpatyoti JA, Barchet W, Ludwig J, Dermody TS, Hartmann G, Reis e Sousa C. 2014. Antiviral immunity via RIG-I-mediated recognition of RNA bearing 5′-diphosphates. Nature 514(7522):372–375.
Goubau D, van der Veen AG, Chakravarty P, Lin R, Rogers N, Rehwinkel J, Deddouche S, Rosewell I, Hiscott J, Reis ESC. 2015. Mouse superkiller-2-like helicase DDX60 is dispensable for type I IFN induction and immunity to multiple viruses. Eur J Immunol 45(12):3386–3403.
Grunvogel O, Esser-Nobis K, Reustle A, Schult P, Muller B, Metz P, Trippler M, Windisch MP, Frese M, Binder M, Fackler O, Bartenschlager R, Ruggieri A, Lohmann V. 2015. DDX60L is an interferon-stimulated gene product restricting hepatitis C virus replication in cell culture. J Virol 89(20):10548–10568.
Gu L, Fullam A, Brennan R, Schroder M. 2013. Human DEAD box helicase 3 couples IkappaB kinase epsilon to interferon regulatory factor 3 activation. Mol Cell Biol 33(10):2004–2015.
Habjan M, Andersson I, Klingstrom J, Schumann M, Martin A, Zimmermann P, Wagner V, Pichlmair A, Schneider U, Muhlberger E, Mirazimi A, Weber F. 2008. Processing of genome 5′ termini as a strategy of negative-strand RNA viruses to avoid RIG-I-dependent interferon induction. PLoS One 3(4):e2032.
Habjan M, Hubel P, Lacerda L, Benda C, Holze C, Eberl CH, Mann A, Kindler E, Gil-Cruz C, Ziebuhr J, Thiel V, Pichlmair A. 2013. Sequestration by IFIT1 impairs translation of 2′O-unmethylated capped RNA. PLoS Pathog 9(10):e1003663.
Habjan M, Pichlmair A. 2015. Cytoplasmic sensing of viral nucleic acids. Curr Opin Virol 11:31–37.
Hartner JC, Walkley CR, Lu J, Orkin SH. 2009. ADAR1 is essential for the maintenance of hematopoiesis and suppression of interferon signaling. Nat Immunol 10(1):109–115.
Hopper AK. 2013. Transfer RNA post-transcriptional processing, turnover, and subcellular dynamics in the yeast *Saccharomyces cerevisiae*. Genetics 194(1):43–67.
Hornung V, Ellegast J, Kim S, Brzozka K, Jung A, Kato H, Poeck H, Akira S, Conzelmann KK, Schlee M, Endres S, Hartmann G. 2006. 5′-Triphosphate RNA is the ligand for RIG-I. Science 314(5801):994–997.
Hornung V, Hartmann R, Ablasser A, Hopfner KP. 2014. OAS proteins and cGAS: unifying concepts in sensing and re-

sponding to cytosolic nucleic acids. Nat Rev Immunol 14(8):521–528.
Hsu MT, Parvin JD, Gupta S, Krystal M, Palese P. 1987. Genomic RNAs of influenza viruses are held in a circular conformation in virions and in infected cells by a terminal panhandle. Proc Natl Acad Sci U S A 84(22):8140–8144.
Hyde JL, Gardner CL, Kimura T, White JP, Liu G, Trobaugh DW, Huang C, Tonelli M, Paessler S, Takeda K, Klimstra WB, Amarasinghe GK, Diamond MS. 2014. A viral RNA structural element alters host recognition of nonself RNA. Science 343(6172):783–787.
Imaizumi T, Yoshida H, Hayakari R, Xing F, Wang L, Matsumiya T, Tanji K, Kawaguchi S, Murakami M, Tanaka H. 2016. Interferon-stimulated gene (ISG) 60, as well as ISG56 and ISG54, positively regulates TLR3/IFN-beta/STAT1 axis in U373MG human astrocytoma cells. Neurosci Res 105: 35–41.
Isaacs A, Lindenmann J. 1957. Virus interference. I. The interferon. Proc R Soc Lond B Biol Sci 147(927):258–267.
Jockel S, Nees G, Sommer R, Zhao Y, Cherkasov D, Hori H, Ehm G, Schnare M, Nain M, Kaufmann A, Bauer S. 2012. The 2′-O-methylation status of a single guanosine controls transfer RNA-mediated Toll-like receptor 7 activation or inhibition. J Exp Med 209(2):235–241.
Kaiser S, Rimbach K, Eigenbrod T, Dalpke AH, Helm M. 2014. A modified dinucleotide motif specifies tRNA recognition by TLR7. RNA 20(9):1351–1355.
Kang R, Tang D. 2012. PKR-dependent inflammatory signals. Sci Signal 5(247):pe47.
Kato H, Takeuchi O, Mikamo-Satoh E, Hirai R, Kawai T, Matsushita K, Hiiragi A, Dermody TS, Fujita T, Akira S. 2008. Length-dependent recognition of double-stranded ribonucleic acids by retinoic acid-inducible gene-I and melanoma differentiation-associated gene 5. J Exp Med 205(7): 1601–1610.
Kato H, Takeuchi O, Sato S, Yoneyama M, Yamamoto M, Matsui K, Uematsu S, Jung A, Kawai T, Ishii KJ, Yamaguchi O, Otsu K, Tsujimura T, Koh CS, Reis e Sousa C, Matsuura Y, Fujita T, Akira S. 2006. Differential roles of MDA5 and RIG-I helicases in the recognition of RNA viruses. Nature 441(7089):101–105.
Killip MJ, Smith M, Jackson D, Randall RE. 2014. Activation of the interferon induction cascade by influenza a viruses requires viral RNA synthesis and nuclear export. J Virol 88(8):3942–3952.
Killip MJ, Young DF, Precious BL, Goodbourn S, Randall RE. 2012. Activation of the beta interferon promoter by paramyxoviruses in the absence of virus protein synthesis. J Gen Virol 93(Pt 2):299–307.
Kim HJ, Fodor E, Brownlee GG, Seong BL. 1997. Mutational analysis of the RNA-fork model of the influenza A virus vRNA promoter in vivo. J Gen Virol 78(Pt 2):353–357.
Kirchner S, Ignatova Z. 2015. Emerging roles of tRNA in adaptive translation, signalling dynamics and disease. Nat Rev Genet 16(2):98–112.
Ko C, Lee S, Windisch MP, Ryu WS. 2014. DDX3 DEAD-box RNA helicase is a host factor that restricts hepatitis B virus replication at the transcriptional level. J Virol 88(23):13689–13698.
Kohler A, Hurt E. 2007. Exporting RNA from the nucleus to the cytoplasm. Nat Rev Mol Cell Biol 8(10):761–773.
Kohlway A, Luo D, Rawling DC, Ding SC, Pyle AM. 2013. Defining the functional determinants for RNA surveillance by RIG-I. EMBO Rep 14(9):772–779.
Koppstein D, Ashour J, Bartel DP. 2015. Sequencing the cap-snatching repertoire of H1N1 influenza provides insight into the mechanism of viral transcription initiation. Nucleic Acids Res 43(10):5052–5064.
Kumar P, Sweeney TR, Skabkin MA, Skabkina OV, Hellen CU, Pestova TV. 2014. Inhibition of translation by IFIT family members is determined by their ability to interact selectively with the 5′-terminal regions of cap0-, cap1- and 5′ppp-mRNAs. Nucleic Acids Res 42(5):3228–3245.
Lassig C, Matheisl S, Sparrer KM, de Oliveira Mann CC, Moldt M, Patel JR, Goldeck M, Hartmann G, Garcia-Sastre A, Hornung V, Conzelmann KK, Beckmann R, Hopfner KP. 2015. ATP hydrolysis by the viral RNA sensor RIG-I prevents unintentional recognition of self-RNA. Elife 4:e10859.
Lee YF, Nomoto A, Detjen BM, Wimmer E. 1977. A protein covalently linked to poliovirus genome RNA. Proc Natl Acad Sci U S A 74(1):59–63.
Leung DW, Amarasinghe GK. 2016. When your cap matters: structural insights into self vs non-self recognition of 5′ RNA by immunomodulatory host proteins. Curr Opin Struct Biol 36:133–141.
Levanon EY, Eisenberg E, Yelin R, Nemzer S, Hallegger M, Shemesh R, Fligelman ZY, Shoshan A, Pollock SR, Sztybel D, Olshansky M, Rechavi G, Jantsch MF. 2004. Systematic identification of abundant A-to-I editing sites in the human transcriptome. Nat Biotechnol 22(8):1001–1005.
Liddicoat BJ, Piskol R, Chalk AM, Ramaswami G, Higuchi M, Hartner JC, Li JB, Seeburg PH, Walkley CR. 2015. RNA editing by ADAR1 prevents MDA5 sensing of endogenous dsRNA as nonself. Science 349(6252):1115–1120.
Liedmann S, Hrincius ER, Guy C, Anhlan D, Dierkes R, Carter R, Wu G, Staeheli P, Green DR, Wolff T, McCullers JA, Ludwig S, Ehrhardt C. 2014. Viral suppressors of the RIG-I-mediated interferon response are pre-packaged in influenza virions. Nat Commun 5:5645.
Lingel A, Simon B, Izaurralde E, Sattler M. 2005. The structure of the flock house virus B2 protein, a viral suppressor of RNA interference, shows a novel mode of double-stranded RNA recognition. EMBO Rep 6(12):1149–1155.
Loo YM, Fornek J, Crochet N, Bajwa G, Perwitasari O, Martinez-Sobrido L, Akira S, Gill MA, Garcia-Sastre A, Katze MG, Gale M, Jr. 2008. Distinct RIG-I and MDA5 signaling by RNA viruses in innate immunity. J Virol 82(1):335–345.
Mackenzie J. 2005. Wrapping things up about virus RNA replication. Traffic 6(11):967–977.
Maharaj NP, Wies E, Stoll A, Gack MU. 2012. Conventional protein kinase C-alpha (PKC-alpha) and PKC-beta negatively regulate RIG-I antiviral signal transduction. J Virol 86(3):1358–1371.
Malathi K, Dong B, Gale M, Jr., Silverman RH. 2007. Small self-RNA generated by RNase L amplifies antiviral innate immunity. Nature 448(7155):816–819.
Malathi K, Saito T, Crochet N, Barton DJ, Gale M, Jr, Silverman RH. 2010. RNase L releases a small RNA from HCV RNA that refolds into a potent PAMP. RNA 16(11):2108–2119.
Marq JB, Hausmann S, Veillard N, Kolakofsky D, Garcin D. 2011. Short double-stranded RNAs with an overhanging 5′ ppp-nucleotide, as found in arenavirus genomes, act as RIG-I decoys. J Biol Chem 286(8):6108–6116.
McCartney SA, Thackray LB, Gitlin L, Gilfillan S, Virgin HW, Colonna M. 2008. MDA-5 recognition of a murine norovirus. PLoS Pathog 4(7):e1000108.

McCracken S, Fong N, Rosonina E, Yankulov K, Brothers G, Siderovski D, Hessel A, Foster S, Shuman S, Bentley DL. 1997. 5′-Capping enzymes are targeted to pre-mRNA by binding to the phosphorylated carboxy-terminal domain of RNA polymerase II. Genes Dev 11(24):3306–3318.

Miyashita M, Oshiumi H, Matsumoto M, Seya T. 2011. DDX60, a DEXD/H box helicase, is a novel antiviral factor promoting RIG-I-like receptor-mediated signaling. Mol Cell Biol 31(18):3802–3819.

Molleston JM, Sabin LR, Moy RH, Menghani SV, Rausch K, Gordesky-Gold B, Hopkins KC, Zhou R, Jensen TH, Wilusz JE, Cherry S. 2016. A conserved virus-induced cytoplasmic TRAMP-like complex recruits the exosome to target viral RNA for degradation. Genes Dev 30(14):1658–1670.

Monroe KM, McWhirter SM, Vance RE. 2009. Identification of host cytosolic sensors and bacterial factors regulating the type I interferon response to *Legionella pneumophila*. PLoS Pathog 5(11):e1000665.

Moy RH, Cole BS, Yasunaga A, Gold B, Shankarling G, Varble A, Molleston JM, tenOever BR, Lynch KW, Cherry S. 2014. Stem-loop recognition by DDX17 facilitates miRNA processing and antiviral defense. Cell 158(4):764–777.

Muller-McNicoll M, Neugebauer KM. 2013. How cells get the message: dynamic assembly and function of mRNA-protein complexes. Nat Rev Genet 14(4):275–287.

Neumann G, Hobom G. 1995. Mutational analysis of influenza virus promoter elements in vivo. J Gen Virol 76(Pt 7):1709–1717.

Nikonov A, Molder T, Sikut R, Kiiver K, Mannik A, Toots U, Lulla A, Lulla V, Utt A, Merits A, Ustav M. 2013. RIG-I and MDA-5 detection of viral RNA-dependent RNA polymerase activity restricts positive-strand RNA virus replication. PLoS Pathog 9(9):e1003610.

O'Neill LA, Golenbock D, Bowie AG. 2013. The history of Toll-like receptors – redefining innate immunity. Nat Rev Immunol 13(6):453–460.

Onoguchi K, Yoneyama M, Fujita T. 2011. Retinoic acid-inducible gene-I-like receptors. J Interferon Cytokine Res 31(1):27–31.

Oshiumi H, Miyashita M, Matsumoto M, Seya T. 2013. A distinct role of Riplet-mediated K63-Linked polyubiquitination of the RIG-I repressor domain in human antiviral innate immune responses. PLoS Pathog 9(8):e1003533.

Oshiumi H, Miyashita M, Okamoto M, Morioka Y, Okabe M, Matsumoto M, Seya T. 2015. DDX60 is involved in RIG-I-dependent and independent antiviral responses, and its function is attenuated by virus-induced EGFR activation. Cell Rep 11(8):1193–1207.

Oshiumi H, Sakai K, Matsumoto M, Seya T. 2010. DEAD/H BOX 3 (DDX3) helicase binds the RIG-I adaptor IPS-1 to up-regulate IFN-beta-inducing potential. Eur J Immunol 40(4):940–948.

Painter MM, Morrison JH, Zoecklein LJ, Rinkoski TA, Watzlawik JO, Papke LM, Warrington AE, Bieber AJ, Matchett WE, Turkowski KL, Poeschla EM, Rodriguez M. 2015. Antiviral protection via RdRP-mediated stable activation of innate immunity. PLoS Pathog 11(12):e1005311.

Paul D, Bartenschlager R. 2015. Flaviviridae replication organelles: oh, what a tangled web we weave. Annu Rev Virol 2(1):289–310.

Peisley A, Lin C, Wu B, Orme-Johnson M, Liu M, Walz T, Hur S. 2011. Cooperative assembly and dynamic disassembly of MDA5 filaments for viral dsRNA recognition. Proc Natl Acad Sci U S A 108(52):21010–21015.

Peisley A, Wu B, Yao H, Walz T, Hur S. 2013. RIG-I forms signaling-competent filaments in an ATP-dependent, ubiquitin-independent manner. Mol Cell 51(5):573–583.

Pelka K, Shibata T, Miyake K, Latz E. 2016. Nucleic acid-sensing TLRs and autoimmunity: novel insights from structural and cell biology. Immunol Rev 269(1):60–75.

Pestal K, Funk CC, Snyder JM, Price ND, Treuting PM, Stetson DB. 2015. Isoforms of RNA-editing enzyme ADAR1 independently control nucleic acid sensor MDA5-driven autoimmunity and multi-organ development. Immunity 43(5):933–944.

Pflug A, Guilligay D, Reich S, Cusack S. 2014. Structure of influenza A polymerase bound to the viral RNA promoter. Nature 516(7531):355–360.

Pham AM, Santa Maria FG, Lahiri T, Friedman E, Marie IJ, Levy DE. 2016. PKR transduces MDA5-dependent signals for type I IFN induction. PLoS Pathog 12(3):e1005489.

Pichlmair A, Lassnig C, Eberle CA, Gorna MW, Baumann CL, Burkard TR, Burckstummer T, Stefanovic A, Krieger S, Bennett KL, Rulicke T, Weber F, Colinge J, Muller M, Superti-Furga G. 2011. IFIT1 is an antiviral protein that recognizes 5′-triphosphate RNA. Nat Immunol 12(7):624–630.

Pichlmair A, Schulz O, Tan CP, Naslund TI, Liljestrom P, Weber F, Reis e Sousa C. 2006. RIG-I-mediated antiviral responses to single-stranded RNA bearing 5′-phosphates. Science 314(5801):997–1001.

Pichlmair A, Schulz O, Tan CP, Rehwinkel J, Kato H, Takeuchi O, Akira S, Way M, Schiavo G, Reis e Sousa C. 2009. Activation of MDA5 requires higher-order RNA structures generated during virus infection. J Virol 83(20):10761–10769.

Pinto AK, Williams GD, Szretter KJ, White JP, Proenca-Modena JL, Liu G, Olejnik J, Brien JD, Ebihara H, Muhlberger E, Amarasinghe G, Diamond MS, Boon AC. 2015. Human and murine IFIT1 proteins do not restrict infection of negative-sense RNA viruses of the orthomyxoviridae, bunyaviridae, and filoviridae families. J Virol 89(18):9465–9476.

Portal MM, Pavet V, Erb C, Gronemeyer H. 2015. Human cells contain natural double-stranded RNAs with potential regulatory functions. Nat Struct Mol Biol 22(1):89–97.

Rassa JC, Ross SR. 2003. Viruses and toll-like receptors. Microbes Infect 5(11):961–968.

Rehwinkel J, Tan CP, Goubau D, Schulz O, Pichlmair A, Bier K, Robb N, Vreede F, Barclay W, Fodor E, Reis e Sousa C. 2010. RIG-I detects viral genomic RNA during negative-strand RNA virus infection. Cell 140(3):397–408.

Reich S, Guilligay D, Pflug A, Malet H, Berger I, Crepin T, Hart D, Lunardi T, Nanao M, Ruigrok RW, Cusack S. 2014. Structural insight into cap-snatching and RNA synthesis by influenza polymerase. Nature 516(7531):361–366.

Resa-Infante P, Jorba N, Coloma R, Ortin J. 2011. The influenza virus RNA synthesis machine: advances in its structure and function. RNA Biol 8(2):207–215.

Rice GI, Kasher PR, Forte GM, Mannion NM, Greenwood SM, Szynkiewicz M, Dickerson JE, Bhaskar SS, Zampini M, Briggs TA, Jenkinson EM, Bacino CA, Battini R, Bertini E, Brogan PA, Brueton LA, Carpanelli M, De Laet C, de Lonlay P, del Toro M, Desguerre I, Fazzi E, Garcia-Cazorla A, Heiberg A, Kawaguchi M, Kumar R, Lin JP, Lourenco CM, Male AM, Marques W, Jr., Mignot C, Olivieri I, Orcesi S, Prabhakar P, Rasmussen M, Robinson RA, Rozenberg F, Schmidt JL, Steindl K, Tan TY, van der Merwe WG, Vanderver A, Vassallo G, Wakeling EL, Wassmer E, Whittaker

E, Livingston JH, Lebon P, Suzuki T, McLaughlin PJ, Keegan LP, O'Connell MA, Lovell SC, Crow YJ. 2012. Mutations in ADAR1 cause Aicardi-Goutieres syndrome associated with a type I interferon signature. Nat Genet 44(11):1243–1248.

Roers A, Hiller B, Hornung V. 2016. Recognition of endogenous nucleic acids by the innate immune system. Immunity 44(4):739–754.

Rosenthal JJ. 2015. The emerging role of RNA editing in plasticity. J Exp Biol 218(Pt 12):1812–1821.

Runge S, Sparrer KM, Lassig C, Hembach K, Baum A, Garcia-Sastre A, Soding J, Conzelmann KK, Hopfner KP. 2014. In vivo ligands of MDA5 and RIG-I in measles virus-infected cells. PLoS Pathog 10(4):e1004081.

Saito T, Hirai R, Loo YM, Owen D, Johnson CL, Sinha SC, Akira S, Fujita T, Gale M, Jr. 2007. Regulation of innate antiviral defenses through a shared repressor domain in RIG-I and LGP2. Proc Natl Acad Sci U S A 104(2):582–587.

Sanchez David RY, Combredet C, Sismeiro O, Dillies MA, Jagla B, Coppee JY, Mura M, Guerbois Galla M, Despres P, Tangy F, Komarova AV. 2016. Comparative analysis of viral RNA signatures on different RIG-I-like receptors. Elife 5:e11275.

Sancho D, Reis e Sousa C. 2013. Sensing of cell death by myeloid C-type lectin receptors. Curr Opin Immunol 25(1): 46–52.

Sarin LP, Leidel SA. 2014. Modify or die?—RNA modification defects in metazoans. RNA Biol 11(12):1555–1567.

Satoh T, Kato H, Kumagai Y, Yoneyama M, Sato S, Matsushita K, Tsujimura T, Fujita T, Akira S, Takeuchi O. 2010. LGP2 is a positive regulator of RIG-I- and MDA5-mediated antiviral responses. Proc Natl Acad Sci U S A 107(4):1512–1517.

Schlee M. 2013. Master sensors of pathogenic RNA – RIG-I like receptors. Immunobiology 218(11):1322–1335.

Schlee M, Roth A, Hornung V, Hagmann CA, Wimmenauer V, Barchet W, Coch C, Janke M, Mihailovic A, Wardle G, Juranek S, Kato H, Kawai T, Poeck H, Fitzgerald KA, Takeuchi O, Akira S, Tuschl T, Latz E, Ludwig J, Hartmann G. 2009. Recognition of 5′ triphosphate by RIG-I helicase requires short blunt double-stranded RNA as contained in panhandle of negative-strand virus. Immunity 31(1):25–34.

Schmidt A, Schwerd T, Hamm W, Hellmuth JC, Cui S, Wenzel M, Hoffmann FS, Michallet MC, Besch R, Hopfner KP, Endres S, Rothenfusser S. 2009. 5′-Triphosphate RNA requires base-paired structures to activate antiviral signaling via RIG-I. Proc Natl Acad Sci U S A 106(29):12067–12072.

Schneider U, Martin A, Schwemmle M, Staeheli P. 2007. Genome trimming by Borna disease viruses: viral replication control or escape from cellular surveillance? Cell Mol Life Sci 64(9):1038–1042.

Schuberth-Wagner C, Ludwig J, Bruder AK, Herzner AM, Zillinger T, Goldeck M, Schmidt T, Schmid-Burgk JL, Kerber R, Wolter S, Stumpel JP, Roth A, Bartok E, Drosten C, Coch C, Hornung V, Barchet W, Kummerer BM, Hartmann G, Schlee M. 2015. A conserved histidine in the RNA sensor RIG-I controls immune tolerance to N1-2′O-methylated self RNA. Immunity 43(1):41–51.

Schulz O, Pichlmair A, Rehwinkel J, Rogers NC, Scheuner D, Kato H, Takeuchi O, Akira S, Kaufman RJ, Reis e Sousa C. 2010. Protein kinase R contributes to immunity against specific viruses by regulating interferon mRNA integrity. Cell Host Microbe 7(5):354–361.

Shi Y, Yuan B, Qi N, Zhu W, Su J, Li X, Qi P, Zhang D, Hou F. 2015. An autoinhibitory mechanism modulates MAVS activity in antiviral innate immune response. Nat Commun 6:7811.

Singh R, Reddy R. 1989. Gamma-monomethyl phosphate: a cap structure in spliceosomal U6 small nuclear RNA. Proc Natl Acad Sci U S A 86(21):8280–8283.

Strahle L, Garcin D, Kolakofsky D. 2006. Sendai virus defective-interfering genomes and the activation of interferon-beta. Virology 351(1):101–111.

Sun Z, Ren H, Liu Y, Teeling JL, Gu J. 2011. Phosphorylation of RIG-I by casein kinase II inhibits its antiviral response. J Virol 85(2):1036–1047.

Takahasi K, Yoneyama M, Nishihori T, Hirai R, Kumeta H, Narita R, Gale M, Jr, Inagaki F, Fujita T. 2008. Nonself RNA-sensing mechanism of RIG-I helicase nd activation of antiviral immune responses. Mol Cell 29(4):428–440.

Thulasi Raman SN, Liu G, Pyo HM, Cui YC, Xu F, Ayalew LE, Tikoo SK, Zhou Y. 2016. DDX3 interacts with influenza A virus NS1 and NP proteins and exerts antiviral function through regulation of stress granule formation. J Virol 90(7): 3661–3675.

Tiley LS, Hagen M, Matthews JT, Krystal M. 1994. Sequence-specific binding of the influenza virus RNA polymerase to sequences located at the 5′ ends of the viral RNAs. J Virol 68(8):5108–5116.

Topisirovic I, Svitkin YV, Sonenberg N, Shatkin AJ. 2011. Cap and cap-binding proteins in the control of gene expression. Wiley Interdiscip Rev RNA 2(2):277–298.

Triantafilou K, Vakakis E, Kar S, Richer E, Evans GL, Triantafilou M. 2012. Visualisation of direct interaction of MDA5 and the dsRNA replicative intermediate form of positive strand RNA viruses. J Cell Sci 125(Pt 20):4761–4769.

Uchikawa E, Lethier M, Malet H, Brunel J, Gerlier D, Cusack S. 2016. Structural analysis of dsRNA binding to anti-viral pattern recognition receptors LGP2 and MDA5. Mol Cell 62(4):586–602.

Venkataraman T, Valdes M, Elsby R, Kakuta S, Caceres G, Saijo S, Iwakura Y, Barber GN. 2007. Loss of DExD/H box RNA helicase LGP2 manifests disparate antiviral responses. J Immunol 178(10):6444–6455.

Versteeg GA, Garcia-Sastre A. 2010. Viral tricks to grid-lock the type I interferon system. Curr Opin Microbiol 13(4):508–516.

Weber F, Wagner V, Rasmussen SB, Hartmann R, Paludan SR. 2006. Double-stranded RNA is produced by positive-strand RNA viruses and DNA viruses but not in detectable amounts by negative-strand RNA viruses. J Virol 80(10):5059–5064.

Weber M, Gawanbacht A, Habjan M, Rang A, Borner C, Schmidt AM, Veitinger S, Jacob R, Devignot S, Kochs G, Garcia-Sastre A, Weber F. 2013. Incoming RNA virus nucleocapsids containing a 5′-triphosphorylated genome activate RIG-I and antiviral signaling. Cell Host Microbe 13(3):336–346.

Wolferstatter M, Schweneker M, Spath M, Lukassen S, Klingenberg M, Brinkmann K, Wielert U, Lauterbach H, Hochrein H, Chaplin P, Suter M, Hausmann J. 2014. Recombinant modified vaccinia virus Ankara generating excess early double-stranded RNA transiently activates protein kinase R and triggers enhanced innate immune responses. J Virol 88(24):14396–14411.

Wu B, Peisley A, Richards C, Yao H, Zeng X, Lin C, Chu F, Walz T, Hur S. 2013. Structural basis for dsRNA recognition, filament formation, and antiviral signal activation by MDA5. Cell 152(1–2):276–289.

Wu B, Peisley A, Tetrault D, Li Z, Egelman EH, Magor KE, Walz T, Penczek PA, Hur S. 2014. Molecular imprinting as a signal-activation mechanism of the viral RNA sensor RIG-I. Mol Cell 55(4):511–523.

Xu J, Mercado-Lopez X, Grier JT, Kim WK, Chun LF, Irvine EB, Del Toro Duany Y, Kell A, Hur S, Gale M, Jr, Raj A, Lopez CB. 2015. Identification of a natural viral RNA motif that optimizes sensing of viral RNA by RIG-I. MBio 6(5): e01265-15.

Yao H, Dittmann M, Peisley A, Hoffmann HH, Gilmore RH, Schmidt T, Schmid-Burgk JL, Hornung V, Rice CM, Hur S. 2015. ATP-dependent effector-like functions of RIG-I-like receptors. Mol Cell 58(3):541–548.

Yoneyama M, Kikuchi M, Matsumoto K, Imaizumi T, Miyagishi M, Taira K, Foy E, Loo YM, Gale M, Jr., Akira S, Yonehara S, Kato A, Fujita T. 2005. Shared and unique functions of the DExD/H-box helicases RIG-I, MDA5, and LGP2 in antiviral innate immunity. J Immunol 175(5):2851–2858.

Yoneyama M, Kikuchi M, Natsukawa T, Shinobu N, Imaizumi T, Miyagishi M, Taira K, Akira S, Fujita T. 2004. The RNA helicase RIG-I has an essential function in double-stranded RNA-induced innate antiviral responses. Nat Immunol 5(7): 730–737.

Zamanian-Daryoush M, Mogensen TH, DiDonato JA, Williams BR. 2000. NF-kappaB activation by double-stranded-RNA-activated protein kinase (PKR) is mediated through NF-kappaB-inducing kinase and IkappaB kinase. Mol Cell Biol 20(4):1278–1290.

Zhang B, Liu X, Chen W, Chen L. 2013. IFIT5 potentiates anti-viral response through enhancing innate immune signaling pathways. Acta Biochim Biophys Sin (Shanghai) 45(10): 867–874.

Zhang Z, Kim T, Bao M, Facchinetti V, Jung SY, Ghaffari AA, Qin J, Cheng G, Liu YJ. 2011a. DDX1, DDX21, and DHX36 helicases form a complex with the adaptor molecule TRIF to sense dsRNA in dendritic cells. Immunity 34(6):866–878.

Zhang Z, Yuan B, Lu N, Facchinetti V, Liu YJ. 2011b. DHX9 pairs with IPS-1 to sense double-stranded RNA in myeloid dendritic cells. J Immunol 187(9):4501–4508.

Address correspondence to:
Dr. Andreas Pichlmair
Innate Immunity Laboratory
Max-Planck Institute of Biochemistry
Am Klopferspitz 18
82152 Martinsried/Munich
Germany

E-mail: apichl@biochem.mpg.de

Received 19 September 2016/Accepted 5 January 2017

1.2.2 Viral strategies to inhibit host gene expression and hijack host mRNA export factors

All DNA and RNA viruses are prone to the IFN-induced cellular responses. Therefore, viruses established strategies to escape the cellular innate immune responses by masking their PAMPs (see chapter 1.2.1) and additionally they acquired the ability to interfere with almost all steps in gene expression in order to globally suppress host protein expression known as "host shutoff". This includes inhibition of cellular events like transcription, mRNA processing, mRNA export from the nucleus, regulation of mRNA stability as well as multiple steps in translation. Through this mechanism, viruses not only dampen antiviral responses but also transfer cellular resources towards viral gene expression. In addition, nuclear replicating viruses including almost all DNA viruses (except poxviruses, asfarviruses and phycoviruses) and a few RNA viruses (bornaviruses, orthomycoviruses and retroviruses) have to compete with host mRNA for effective export of their viral mRNAs to successfully replicate and assemble in the host. In this section, I will focus on the viral manipulation of the host cell through two key viruses since they were predominantly used in the thesis research: Influenza A virus (*Orthomyxoviridae*) and Vesicular stomatitis virus (*Rabhdoviridae*).

1.2.2.1 Influenza A virus

Influenza A virus is a nuclear replicating, negative-orientated, single-stranded (ss) RNA virus belonging to the *Orthomyxoviridae* family. The influenza genome is segmented into eight fragments encoding for 12-14 structural and nonstructural proteins depending on the strain. IAV is a rapidly mutating virus resulting in the existence of a large number of different strains with varying non-essential functions, mostly involved in the modulation of virus-host interactions. However, a global block of host protein expression was generally observed in IAV infected cells. In this regard, three main mechanisms of host shutoff by IAV infection have been described: inhibition of cellular mRNA processing and nuclear export, degradation of host RNA polymerase II and wide-ranging degradation of host mRNA [256,257].

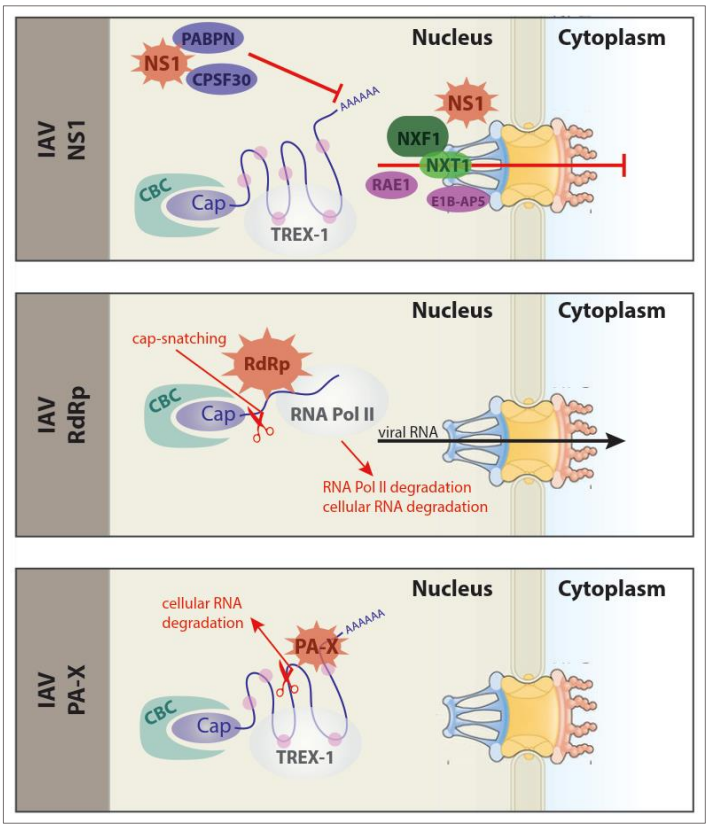


Figure 10: Mechanisms of cellular manipulation by influenza A virus.

Influenza A virus (IAV) NS1 binds and inhibits factors involved in mRNA processing and export (top box). The functions of the viral RNA-dependent RNA polymerase (RdRp) leads to alteration of host gene expression by the combination of cap-snatching and cellular RNA polymerase II degradation (middle box). Through the nuclease activity of PA-X and its selectivity for polymerase II transcripts, cellular RNA is degraded committing to cellular host shutoff (bottom box). Cap, N7-methylguaosine; CBC, cap-binding complex; RNA Pol II, RNA polymerase II; TREX-1, transcription-export complex 1.Schematic and figure legend were adapted from references [221,258].

NS1 mediated inhibition of polyadenylation and host mRNA export

The non-structural protein 1 (NS1) is composed of a N-terminal RNA-binding domain and a C-terminal effector domain by which NS1 exhibits its function in several processes to inhibit type I IFN responses [259,260]. The inhibition of innate immune responses by NS1 was underlined by the usage of IAV with a deleted NS1 protein (IAV-ΔNS1) [261,262]. This virus replicates to a greater extent in IFN-α/β deficient cells (Vero cells) than in type I IFN competent cells (MCDK cells), in which virus infection results in elevated levels of IFN and IFN-stimulated genes. Additionally, in contrast to IAV wild-type, IAV-ΔNS1 is not lethal for mice as the innate immune system is not inhibited by viral NS1 protein. Therefore, the crucial role of NS1 is to counteract and down-regulate the host innate immune responses. In that manner, NS1 blocks nuclear processing of RNA polymerase II transcripts by forming a complex with CPSF30. CPSF30 is an important component of the cellular 3`end processing machinery, which is known as the cellular and polyadenylation specificity factor (CPSF) complex (Figure 10) [263]. The CPSF complex cuts

pre-mRNA downstream of the polyadenylation signal and recruits poly(A) polymerase to polyadenylate the 3`end of the RNA. By binding CPSF30, NS1 prevents polyadenylation and consequently inhibits nuclear export of these incompletely processed mRNAs [264]. In addition, NS1 effector domain associates with the nuclear poly(A) binding protein (PABPN) and inhibits its function in synthesis of long poly(A) tails driven by poly(A) polymerase (Figure 10). The combination of blocking the CPSF complex activity and the synthesis of long poly(A) tails results in the accumulation of short poly(A) tailed mRNA that are less efficiently transported to the cytoplasm [264]. A more direct effect on inhibiting mRNA export was shown by the interaction of NS1 with proteins such as NXF1-NXT1, RAE1 and E1B-AP5, proteins that all bind to cellular mRNA and nucleoporins to guide mRNAs through the nuclear pore (Figure 10) [265,266]. Interestingly, the inhibition of mRNA export by NS1 could be reversed by the overexpression of NXF1. Additionally, it was shown that NS1 weakly binds Nup98 and initiates its degradation to further achieve a block in mRNA export from the nucleus. Surprisingly, not all IAV strains have the ability to carry out NS1-mediated host shutoff. Avian and swine IAV strains as well as the mouse-adapted laboratory strain A/PuertoRico/8/1934 (PR8) and the human pandemic 2009 H1N1 strain carry NS1 proteins that do not inhibit mRNA 3`end processing [267–271]. Several studies indicated that specific amino acids in NS1 are required for CPSF30 binding and that these residues vary between CPSF30-inhibiting and non-inhibiting strains. By mutating these residues in the non-inhibiting strains (avian and swine origin) to the NS1-consensus sequence, these strains showed an increased pathogenicity.

Surprisingly, although NS1 has such a strong inhibitory effect on host mRNA export, a solid poly(A) signal can still be identified in the nucleus of IAV infected or NS1 transfected cells. This indicates that subsets of mRNAs may have the ability to bypass NS1-mediated mRNA processing inhibition. Indeed, the Stern-Ginossar group has recently shown that a subset of genes contributing to oxidative phosphorylation can escape degradation and are constantly expressed during general IAV host shutoff [257]. Since oxidative phosphorylation is a cellular housekeeping function to generate biological usable energy and viruses are unable to produce their own energy, the authors hypothesized that bypassing these genes is one way to maintain or even increase energy production during host shutoff.

In contrast to cellular mRNAs, viral mRNAs are not affected by the general shutoff mediated by NS1 since the polyadenylation of these transcripts is carried out by the viral RNA-dependent RNA polymerase (RdRp) by shuttering on a polyuridine stretch of the viral genome [257,272]. Additionally,

some studies indicated that at least a few viral transcripts are exported in an NXF1-independent manner, thus, escaping the virus induced shut off mechanisms.

RdRp mediated degradation of host RNA polymerase II

The viral RdRp consists of the viral PB1, PB2 and PA subunits and associates with the cellular RNA polymerase II on actively transcribed genes including protein-coding mRNAs and snRNAs. The interaction occurs during the initiation of transcription and is dependent on serine 5 phosphorylation of the CTD. In this state, the RNA polymerase II recruits factors to synthesize the 5`cap structure on nascent mRNA transcripts. Following the interaction of RdRp with this state of actively transcribed RNA, viral RdRp cleaves nascent transcripts close to the 5`end and uses this 5`capped RNA as primer for viral mRNA synthesis. This process, known as "cap-snatching", leaves uncapped cellular RNA fragments for degradation (Figure 10). In addition, this process leads to ubiquitination of the large subunit of the RNA polymerase II and subsequently to its degradation by the proteasome. The combination of cap-snatching and RNA polymerase II degradation is an effective mechanism to alter host gene expression. However, by degrading RNA polymerase II the viral transcriptional activity also decreases since it is dependent on the interaction with RNA Pol II and cap-snatching. Nevertheless, the degradation of RNA polymerase II may support to shift from viral transcription to replication by favoring viral genomic synthesis instead of viral mRNA transcription.

PA-X mediated RNA degradation

Until now, it was commonly accepted that RNA degradation was an effect of cap-snatching by the RdRp. However, more recently evidence occurred that the viral protein PA-X is actively required for host shutoff (Figure 10). PA-X is the product of a +1 ribosomal frameshift after amino acid 191 in PA protein ending in a short C-terminal domain (called X-ORF), which is responsible for host shutoff [273,274]. The common N-terminal domain of PA and PA-X encodes a RNA endonuclease domain [274–276]. X-ORF domain varies from 41 to 61 amino acids in different IAV strains, whereby only the first 15 amino acid are needed for host shutoff but the extended 61 amino acid versions have revealed stronger activity [277,278]. RdRp subunit PA and its ssRNA cleaving potential may only be needed to cut pre-mRNA during cap-snatching, whereas, the endonuclease activity of PA-X with less specificity might be committed to host shutoff [277]. A common feature of PA and PA-X endonuclease activity is that both proteins preferentially cleave RNA polymerase II transcripts and spare transcripts arising from viral RdRp transcription and other cellular RNA polymerases [277].

Interestingly, IAV infections with decreased PA-X levels induced by mutating the frameshift-promoting sequence resulted in elevated immune responses in mice leading to a higher

pathogenicity due to increased lung immunopathology [274,279–281]. Cells infected with the PA-X mutant IAV virus showed similar levels of viral mRNAs and vRNAs compared to IAV wild-type infection. However, viral proteins were less expressed indicating that PA-X reduces the pool of cellular mRNAs that access the translational machinery and thereby viral transcripts do not have to compete for translation with host mRNAs [282].

Since PA-X production by ribosomal frameshift transcription is conserved among the various IAV strains [283], it indicates that PA-X triggered RNA degradation might be a general IAV mechanism for host shutoff and additional strain-dependent adaptions were added through evolution to compete with increased host complexity.

1.2.2.2 Vesicular stomatitis virus

Vesicular stomatitis virus (VSV) is a ssRNA virus with a negative-oriented linear genome belonging to the *Rhabdoviridae* family. In contrast to IAV, VSV replicates in the cytoplasm and therefore has no need to hijack the cellular mRNA export machinery. However, host mRNA export is rapidly repressed by the nuclear localizing viral matrix (M) protein to primarily dampen cellular innate immune responses [258,266]. In addition, by blocking cellular mRNA export VSV decreases the competition with cellular mRNAs for the use of the translational machinery. Nuclear VSV M protein interacts with the RNA binding protein RAE1, which is in a complex with the nucleoporin Nup98 (Figure 11) [284,285]. This interaction leads to the nuclear retention of mRNAs and snRNAs. Using crystallographic and biochemical studies, Quan et al. revealed the molecular mechanism of M-RAE1-Nup98 binding and inhibition of cellular mRNA export [286]. Viral M proteins mimics the phosphate backbone by which it hijacks the nucleic acid binding site of RAE1 and sabotages cellular mRNA export.

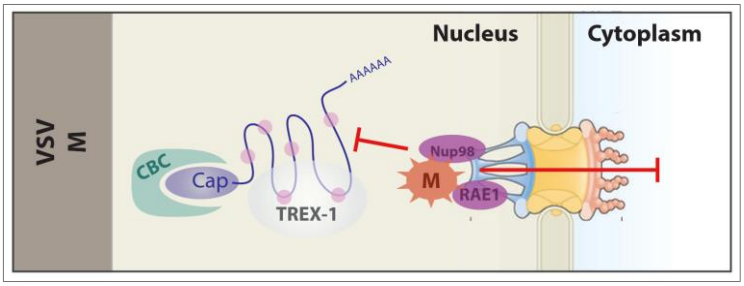


Figure 11: VSV disruption of cellular mRNA export.

Vesicular stomatitis virus (VSV) matrix (M) protein interacts with RAE1 and Nup98 and thereby blocks cellular mRNA export to the cytoplasm. CBC, cap-binding complex; Cap, N7-methylguaosine; TREX-1, transcription-export complex 1. Schematic and figure legend were adapted from reference [258].

The most prominent feature of protein M mimicking the phosphate backbone is the methionine on position 51 (M51) with its upstream and downstream acidic residues. Using VSV with a mutation in M51 to alanine (M51A) did not block mRNA export and immune responses were

mounted upon virus infection [287]. Interestingly, M-mediated mRNA export block can be reverted by IFN since IFN stimulation increases the levels of RAE1 and Nup98/96 [285,288,289]. It has also been shown that M-RAE1-Nup98 interaction inhibits transcription [290]. However, several studies demonstrated that high levels of bulk poly(A) RNA are retained in the nucleus in the presence of VSV-M expression indicating that the cellular transcripts are properly polyadenylated and that VSV-M uses post-transcriptional mechanisms to block host gene expression [258]. Since nucleoplasmic Nup98 has been indicated to regulate a subset of RNA polymerase II genes [291,292], it might be possible that the VSV M protein thereby specifically inhibits transcription of certain Nup98-dependent genes. RAE1 and Nup98 were shown to regulate spindle assembly during mitosis [293,294], which has a high impact in cancer cells due to their high mitotic index. The interaction of VSV M and RAE1-Nup98 complex inhibits mitotic progression and leads to cell death [295], which likely contributes to VSV`s oncological potential. The potential of VSV to preferentially infect cancer cells and inhibiting mitotic progression leading to cell death is used in the development of new cancer therapeutics [296].

In summary, inhibition of cellular mRNA export by the interaction of VSV-M protein with RAE1-Nup98 helps the virus to exclusively access the translational machinery and avoid the competition with host mRNAs for translation. In addition, the block of mRNA export inhibits the expression of immune regulated genes.

1.3 IMPORTANCE OF VIRAL RNA-BINDING PROTEINS FOR VIRAL REPLICATION

Single-stranded RNA viruses are classified according to their positively (+) or negatively (-) encoded genome. Depending on the orientation, the viral genetic information is translated directly into functional proteins ((+) RNA genome) or it is transcribed into positively orientated RNA ((-) RNA genome) by the viral RdRp. Translation of viral proteins produce structural and non-structural proteins that are involved in the architecture of the viral particle or participate in the transcription and replication of the viral genome, respectively. Assembly of functional viral particles includes genome replication. For both, (+) and (-) ssRNA viruses, the viral RNA polymerase generates a complementary RNA (cRNA) intermediate as template for the replication of the RNA genome. In this process, viral non-structural proteins with RNA-binding ability are involved and crucial for successful execution of the replication process. In the following section, I will highlight the importance of two viral RNA-binding proteins, namely the NP protein of influenza A virus ((-) RNA virus) and the non-structural 5A (NS5A) protein of hepatitis C virus (HCV; (+) RNA virus), in virus replication.

1.3.1 Nucleoprotein of influenza A virus

Each segment of the negative-orientated influenza virus genome is coated by multiple copies of NP that form together with one RdRp complex the vRNP, the minimal functional unit for viral transcription and replication (Figure 12, box) [297,298]. Comparing NPs of different negative-stranded RNA viruses has revealed two shared features of NP: a positively charged RNA-binding groove and broad contacts between neighboring NP molecules [298]. Influenza A virus NP was shown to bind RNA with high affinity in a sequence-independent manner potentially through the positively charged cleft between the head and the body domain [299,300]. In addition, crystal structures of NP revealed a tail loop structure which enables NP molecules to form oligomeric structures [299,300]. These oligomeric structures were shown to be required for RNP activity [301]. Studies using artificial mini vRNA of 81 nt length demonstrated that NP can melt secondary RNA structures proposing that NP functions in genome replication by supporting transcript elongation [300]. Indeed, for full-length genome replication NP is absolutely required, whereas, the viral RdRp consisting of PB1, PB2 and PA was shown *in vitro* and *in vivo* to be adequate to replicate short RNA templates in the absence of NP [302–307]. Moreover, NP stimulates polymerase activity [305,308]. This highlights the important of NP for successful and complete replication of the viral genomes. The current model of viral RNA transcription and replication (stabilization model) proposes that the synthesis of cRNA and mRNA from virion-derived RNPs

is stochastic [309]. However, for stabilization and replication of cRNA the active expression of polymerase and NP molecules is important. This was supported by the finding that the RNA-binding activity of NP is needed to stabilize cRNA and that both, RNA-binding and NP oligomerization, promote replication [309]. It was suggested by the model that newly synthesized RdRp and NP molecules are required to stabilize and protect cRNA from cellular degradation. Newly synthesized NP molecules were shown to be kept in a monomeric form before being recruited to nascent cRNAs by the interaction with RdRp, which is bound first to emerging 5′ends of cRNAs [301,307,310,311].

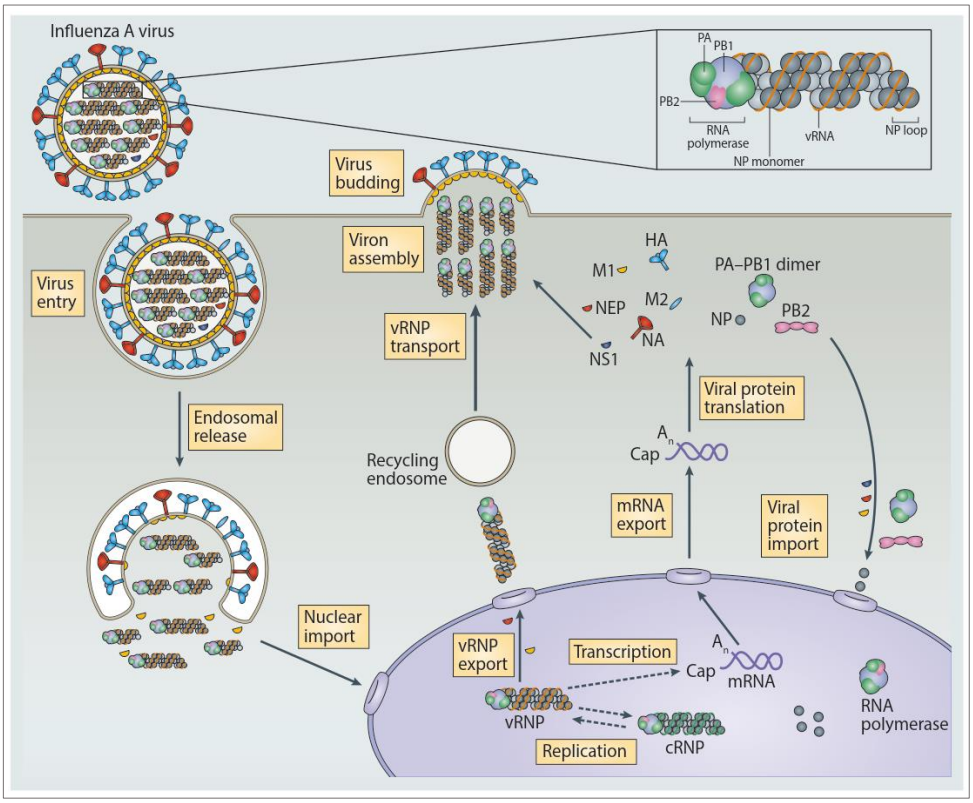


Figure 12: Life cycle of influenza A virus.

The life cycle of Influenza A virus includes the following steps: (1) Virion binding to cell surface receptors containing sialic acid, (2) Endocytosis of the virion and release of the vRNP to the cytoplasm, (3) Nuclear import of vRNP, (4) Cap-snatching and transcription of viral segments into mRNA, (5) Export of viral mRNAs, (6) Translation and nuclear re-import of viral proteins, (7) Replication of vRNPs by the viral RNA polymerase producing an intermediate complementary RNA (cRNA), (8) Export of vRNP, (9) vRNP transport and virion assembly and (10) Budding of the viral particles. The box schematically indicates the architecture of a vRNP containing RdRp (PB1, PB2 and PA), NP proteins and the viral RNA (vRNA). Cap, N7-methylguanosine; RdRp, RNA-dependent RNA polymerase; vRNP, viral ribonucleoprotein particles. Schematic and figure legend were adapted from reference [297].

Subsequently, additional monomeric NP molecules are bound through tail loop interactions forming oligomeric structures. Phosphorylation and the interaction with cellular proteins like importin α5 were shown to be required to keep NP as monomer before it is assembled into oligomeric structures incorporating viral RNA into the coiled vRNP structures [312–315].Altogether, RNA-associated NP is required during viral genome replication for full functionality of viral RdRp during elongation of the cRNA (Figure 12). The high affinity to RNA and the regulation through

post-translational modifications like phosphorylation of NP molecules play an important role in the function of NP.

1.3.2 Non-structural 5A protein of hepatitis C virus

The positively orientated HCV genome is translated at the rough endoplasmic reticulum (ER) in a cap-independent manner by the existence of an internal ribosomal entry site (IRES) in the 5′non-coding region (5′NCR) of the viral RNA (Figure 13) [316]. This yields a polyprotein that is subsequently cleaved into ten structural and non-structural (C, E1, E2, p7, NS2, NS3, NS4A, NS4B, NS5A and NS5B) proteins by cellular (e.g. signal peptidase) and viral proteases (NS2, NS3-4A). After cleavage from the polyprotein, the non-structural proteins function in several steps of the viral life cycle (Figure 13).

One of this process is the RNA replication of HCV genomes, which occurs in membranous derived structures from the ER known as "membranous web" that contain cellular and non-structural viral proteins, HCV RNA, ER membranes and lipid droplets [317,318]. The synthesis of negative-strand RNA is thought to be initiated through the recognition of a sequence motif in the 3`UTR of HCV RNA by the RdRp NS5B (Figure 13) [319–321]. Template specificity and replication activity of NS5B are dependent on the presence of NS2 and NS3 as well as the multifunctional NS5A protein [319]. NS5A was shown to be unconditionally required for viral replication and additionally involved in viral particle assembly (specific function is unknown) [322–326]. NS5A is a proline-rich phosphoprotein consisting of an N-terminal amphipathic α-helix anchoring NS5A to the ER lumen and a cytoplasmic portion organized into three domains (domain I, II and III) separated by low-complexity regions (LCS) I and II. Domain I is highly conserved among HCV genotypes and contains a zinc-binding motif required for dimerization of the N-terminal ends [327]. Mutations in the four cysteines of the zinc-binding domain showed that this domain is critical for viral replication and NS5A dimerization [328]. Through dimerization, a basic groove at the dimeric interface is exposed by which NS5A associates with newly synthesized viral RNA during replication [328–332]. In addition, it was demonstrated that NS5A preferentially binds the polypyrimidine regions of the 3`UTR, but it can also associate with uridylate- and guanylate-rich (U/G) stretches in the IRES of the 5′UTR [329,330]. The latter association was suggested to influence HCV translation, while binding to the conserved polypyrimidine stretch of the 3′UTR is essential for replication. Additionally, it was suggested that through binding of the viral RNA, NS5A protects the RNA from degradation by cellular RNAses as well as from recognition by the innate immune system[316]. This hypothesis was supported by the fact that only a small proportion of the expressed NS5A is required for RNA replication, suggesting that the excess of protein could

participate in shielding the viral RNA from sensing by the host [333,334]. In this regard, a recent study revealed that formation of the "membranous web" limits the access of PRRs to viral replication sites and protects from innate immune sensing [335]. NS5A domains II and III are less well conserved among HCV genotypes and are highly disordered in solution [327]. This flexibility was indicated to be important to interact with a variety of cellular proteins that are required for NS5A function and virus persistence. Domain III was shown to be non-relevant for RNA replication, however, required for successful assembly of viral particles [323,336,337]. Importantly, regions within domain I and II interact with NS5B and stimulates viral RNA polymerase activity in-vitro [338,339].

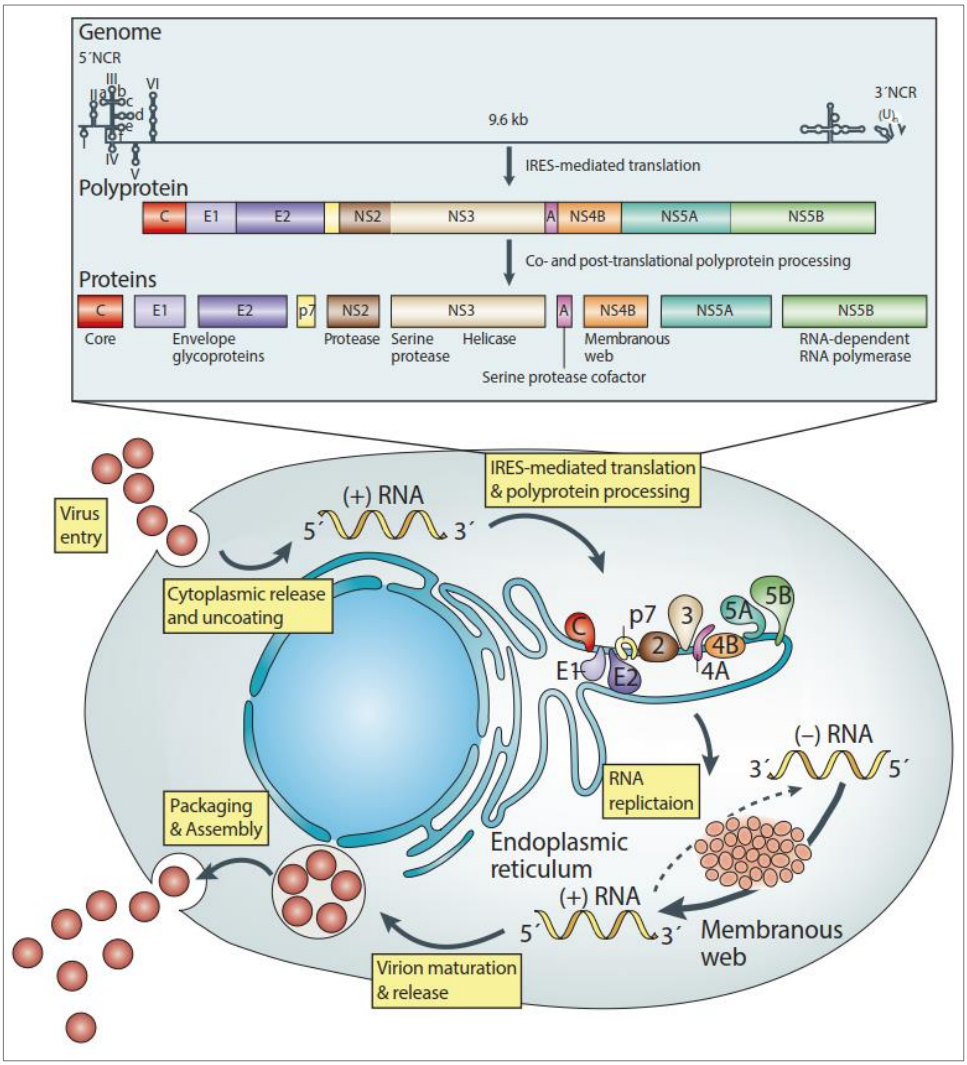


Figure 13: Life cycle of hepatitis C virus.

The top box shows the hepatitis C virus (HCV) internal ribosome entry site (IRES) mediated translation and polyprotein processing. The HCV genome (9.6 kDa positive-stranded RNA) organization is depicted. Secondary structure is shown in a simplified manner representing the 5′- and 3′-non-coding regions (NCRs) and the core gene. IRES-mediated translation results in a polyprotein precursor that is co- and post-translationally processed into the mature structural and non-structural proteins (C, E1, E2, p7, NS2, NS3, NS4A, NS4B, NS5A and NS5B). The bottom depicts a schematic of the viral life cycle with the following steps: (1) Viral entry, (2) Cytoplasmic release and uncoating, (3) IRES-mediated translation and polyprotein processing, (4) RNA replication, (5) Virion maturation and release and (6) Viral particle packaging and assembly. The topology of HCV structural and non-structural proteins at the endoplasmic reticulum is indicated schematically. Schematic and figure legend were adapted from reference [316].

Intracellularly, NS5A is present as a basally phosphorylated (56 kDa) and hyperphosphorylated (58 kDa) form [316]. Mass spectrometry based studies have revealed that the LCS I region is extensively phosphorylated during HCV replication [340–342]. In addition, phosphorylation sites within domain I, II and LCS II have been detected. Several studies showed that serine residues within LCS I are important for viral replication. However, these phosphorylation marks are HCV genotype dependent. Serine 225 phosphorylation within LCS I was demonstrated to be important for correct localization of viral and cellular proteins involved in viral replication. Consequently, mutations in S225 resulted in reduction in viral replication [343]. In sum, NS5A is a highly regulated protein and a plethora of studies have revealed it's important for successful HCV replication. In particular, phosphorylation within LCS I was demonstrated to be dynamically regulated and fundamental for genome replication.

1.4 OBJECTIVE OF THE THESIS

In eukaryotes, the transfer of genetic information from DNA to RNA intermediates and finally to functional proteins is fundamental for every single cell. The separation of the cell into the nucleus and the cytoplasm is one important barrier to separate transcription from translation and to precisely regulate and control gene expression. One essential step in the transmission from DNA to a functional protein is the active transport of mRNA from the nucleus to the cytoplasm. The central player, which is co-transcriptionally bound to the RNA cap structure and triggers processing and export of mRNA, is the canonical CBC. This complex consisting of NCBP2 and its adapter protein NCBP1 was already identified almost 30 years ago by two independent studies of Mutsuhito Ohno and Elisa Izaurralde [125,344].

During my thesis, I challenged the long-standing dogma that the canonical CBC is critically required for mRNA processing and export. This interest was sparked by very elegant inhibition experiments that had shown that loss of NCBP2 impairs snRNA export but not mRNA export [94]. More recently, siRNA-based screens supported this data and draw my awareness to study whether additional components are required for mRNA processing [345–347]. I first sought to investigate the eukaryotic repertoire of cap-binding proteins to identify potential candidates that could compensate for the loss of NCBP2. The aim of my thesis was to characterize the role of the identified candidate/s in mRNA processing and export. I managed to identify and functionally characterize an alternative CBC that is assembled by NCBP1 and -3 and functions in a redundant manner with the canonical CBC under steady-state conditions. I could show that the alternative CBC is important during conditions of cellular stress, which I demonstrated in virus infection experiments using loss of function assays.

During viral infections, mRNA processing is regulated by the function of viral proteins that are themselves target of post-translational regulatory mechanisms. In this manner, serine, threonine or tyrosine phosphorylation plays a fundamental role in regulating the function of proteins. For viral proteins, regulation of phosphorylation-sites may be required for proper viral replication. In collaboration with the lab of Marc Harris (University of Leeds, United Kingdom), we investigated the effect of post-translational modifications of the viral protein NS5A on viral replication. In this manner, I analysed the influence of NS5A S225 phosphorylation on the interaction with cellular proteins.

2 RESULTS

2.1 PUBLICATION 1: mRNA export through an additional cap-binding complex consisting of NCBP1 and NCBP3

Gebhardt, A.*, Habjan, M.*, Benda, C., Meiler, A., Haas, D.A., Hein, M.Y., Mann, A., Mann, M., Habermann, B., and Pichlmair, A. (2015). mRNA export through an additional cap-binding complex consisting of NCBP1 and NCBP3. *Nat. Commun.* 6, 8192. **these authors contributed equally*

The manuscript "mRNA export through an additional cap-binding complex consisting of NCBP1 and NCBP3" describes the first identification and characterization of the cap-binding protein NCBP3. Mass spectrometry based identification of the human and mouse cap-binding protein repertoire revealed the existence of a till then poorly characterized protein encoded by the open reading frame 85 of chromosome 17 (C17orf85). Biochemical analysis confirmed the direct binding of C17orf85 to the mRNA cap structure in an N7-methylguanosine-dependent manner. We renamed C17orf85 to nuclear cap-binding protein 3 (NCBP3). Additional experiments demonstrated the participation of NCBP3 together with NCBP1 in an alternative cap-binding complex. Depletion experiments showed that for appropriate cell viability and distribution of cellular polyadenylated RNA the function of either NCBP2 or NCBP3 is required demonstrating that the alternative (NCBP1/3) and the canonical (NCBP1/2) cap-binding complex execute redundant function under physiological conditions.

The main text including material and methods and supplementary information are displayed in the following. Supplementary tables can be downloaded from the website of *Nature Communications* (https://www.nature.com/articles/ncomms9192).

nature COMMUNICATIONS

ARTICLE

Received 2 Mar 2015 | Accepted 28 Jul 2015 | Published 18 Sep 2015

DOI: 10.1038/ncomms9192 OPEN

mRNA export through an additional cap-binding complex consisting of NCBP1 and NCBP3

Anna Gebhardt[1,*], Matthias Habjan[1,*], Christian Benda[2], Arno Meiler[1], Darya A. Haas[1], Marco Y. Hein[3], Angelika Mann[1], Matthias Mann[3], Bianca Habermann[4] & Andreas Pichlmair[1]

The flow of genetic information from DNA to protein requires polymerase-II-transcribed RNA characterized by the presence of a 5′-cap. The cap-binding complex (CBC), consisting of the nuclear cap-binding protein (NCBP) 2 and its adaptor NCBP1, is believed to bind all capped RNA and to be necessary for its processing and intracellular localization. Here we show that NCBP1, but not NCBP2, is required for cell viability and poly(A) RNA export. We identify C17orf85 (here named NCBP3) as a cap-binding protein that together with NCBP1 forms an alternative CBC in higher eukaryotes. NCBP3 binds mRNA, associates with components of the mRNA processing machinery and contributes to poly(A) RNA export. Loss of NCBP3 can be compensated by NCBP2 under steady-state conditions. However, NCBP3 becomes pivotal under stress conditions, such as virus infection. We propose the existence of an alternative CBC involving NCBP1 and NCBP3 that plays a key role in mRNA biogenesis.

[1] Innate Immunity Laboratory, Max-Planck Institute of Biochemistry, Martinsried, Munich D-82152, Germany. [2] Department of Structural Cell Biology, Max-Planck Institute of Biochemistry, Martinsried, Munich D-82152, Germany. [3] Department of Proteomics and Signal Transduction, Max-Planck Institute of Biochemistry, Martinsried, Munich D-82152, Germany. [4] Bioinformatics Core Facility, Max-Planck Institute of Biochemistry, Martinsried, Munich D-82152, Germany. * These authors contributed equally to this work. Correspondence and requests for materials should be addressed to A.P. (email: apichl@biochem.mpg.de).

Expression of all germline-encoded genetic information in eukaryotes requires RNA transcription through polymerase complexes and subsequent RNA processing, export and translation. Polymerase II transcripts, such as messenger (mRNA), antisense (asRNA), long intergenic non-coding (lincRNA) and small nuclear RNA (snRNA) are marked by an N7-methylated guanine (m^7G) 'cap structure' that is co-transcriptionally attached to the 5′-end of the RNA and serves as a signal for engaging proteins required for downstream processing[1]. Consistent with this notion, splicing and export of snRNA and mRNA can be inhibited by exogenously providing cap analogues[2–4]. The RNA cap structure is bound by the highly conserved nuclear cap-binding complex (CBC), a central factor, known to orchestrate most downstream RNA biogenesis processes such as pre-mRNA splicing, 3′-end processing, nonsense-mediated decay, nuclear–cytoplasmic transport and recruitment of translation factors in the cytoplasm[1,5–8]. The CBC consists of a heterodimer formed by nuclear cap-binding protein 2 (NCBP2, also known as CBP20), which directly associates with the RNA cap, and NCBP1 (also known as CBP80), which stabilizes NCBP2 and serves as an adaptor for other RNA processing factors[9–12]. The central role of the CBC is demonstrated by short interfering RNA (siRNA)-mediated depletion of NCBP1, which results in deregulated expression of several hundred genes, a reduction in the cell proliferation rate[13] and reduction of co-transcriptional spliceosome assembly[14]. NCBP1 directly binds the mRNA export factor ALYREF, and 'CBC competition experiments' using excess of capped RNA led to the conclusion that NCBP1 is involved in mRNA and U snRNA export from the nucleus[15]. However, despite the apparent requirement of NCBP1 for export of capped RNA, antibody-mediated inhibition of NCBP2 *in vivo* only impairs export of U snRNA, but not mRNA[8]. In addition to these data, several genome-wide RNA interference (RNAi)-based screens in human cells that allow assessment of loss-of-function phenotypes in an unbiased manner found that depletion of NCBP1 negatively affects cell growth and viability, whereas depletion of NCBP2 showed only weak phenotypes[16–18]. Collectively, these data suggest that the two CBC subunits only in part share the same biological function. Thus, we questioned whether an additional protein exists that has partially redundant activity to NCBP2 and associates with NCBP1 to form an alternative CBC.

Here, we identify the largely uncharacterized protein C17orf85 (NCBP3) as a novel genuine cap-binding protein that directly interacts with NCBP1 and binds cellular mRNA. Similar to NCBP2, NCBP3 is non-essential under steady-state conditions. However, simultaneous depletion of NCBP2 and -3 mimics the phenotype of NCBP1 knockdown. Notably, NCBP3 becomes pivotal under cellular stress conditions, such as virus infections. We propose the existence of a canonical and an alternative CBC that is fundamental for mRNA biogenesis of higher eukaryotes.

Results

Loss of NCBP1 and NCBP2 leads to different phenotypes. To study the individual requirement of CBC components NCBP1 and NCBP2 for cell viability, we evaluated cell growth after their transient siRNA-mediated depletion in HeLa cells. As expected, depletion of NCBP1 or the Nuclear RNA export factor 1 (NXF1, also known as TAP) severely affected cell growth (Fig. 1a). Surprisingly, after selective depletion of NCBP2, we did not observe a similar effect on cell viability. Likewise, depletion of NCBP2 did not affect intracellular distribution of poly-adenylated (poly(A)) mRNA, as tested by RNA fluorescence *in situ* hybridization (RNA-FISH) (Fig. 1b). In contrast, loss of NCBP1 resulted in accumulation of poly(A) RNA in the nucleus, confirming a critical role of NCBP1 in mRNA export[15]. Since NCBP1 cannot directly associate with capped RNA, we hypothesized on the existence of a protein with a redundant function to NCBP2, that is, a protein with the ability to bridge the association between capped RNA and NCBP1. Contribution of NCBP1 to additional protein complexes is supported by protein expression data based on quantitative mass spectrometry (MS) that suggest about three times higher abundance of NCBP1 as compared with NCBP2 (Fig. 1c).

Identification of C17orf85 as cap-binding protein NCBP3. A protein with redundant function to NCBP2 should have the ability to associate with the RNA cap structure and to bind to NCBP1 to engage factors required for RNA biogenesis. To identify potential candidate proteins, we performed affinity purification followed by liquid chromatography tandem MS (AP–LC–MS/MS) using lysates from human and mouse cell lines and capped and non-capped RNA as baits (Supplementary Fig. 1a). Capped RNA reliably enriched proteins known to associate directly with the RNA cap structure, among them the CBC (consisting of NCBP1 and -2)[6], the genuine cap-binding protein EIF4E[1] and the antiviral protein IFIT1 (ref. 19) from human THP-1 and RAW 264.7 cell lysates (Fig. 2a,b; Supplementary Fig. 1; Supplementary Data 1). Among the proteins identified in human and murine cell lysates was only one protein, C17orf85 (also known as ELG), for which we could not explain its association with RNA in a cap-dependent manner[20,21]. C17orf85 is a highly conserved protein, and orthologues can be found in the fungal and metazoan kingdoms (Supplementary Fig. 2; Supplementary Data 2), suggesting evolutionary conservation for more than one billion years. However, based on currently available whole-genome sequencing data, it seems to have been lost at least by some *Muscomorpha* and *Saccharomycetaceae*, which include *Drosophila melanogaster* and *Saccharomyces cerevisiae*, respectively. Recent data on distribution of proteins in human tissues suggest that C17orf85 is expressed in all tissues[22]. Consistent with a potential redundant role to NCBP2, endogenous and green fluorescent protein (GFP)-tagged C17orf85 showed predominantly nuclear localization (Fig. 2c; Supplementary Fig. 3a). Subcellular fractionation confirmed the nuclear localization, but also identified a cytoplasmic proportion (Supplementary Fig. 3b), a phenotype often found for proteins involved in RNA export[23,24]. Here, we name C17orf85 NCBP3 for nuclear cap-binding protein 3.

To confirm cap-dependent binding of NCBP3, we used capped and non-capped RNA as bait for AP, followed by western blotting. NCBP3 precipitated with capped RNA in human THP-1 and murine RAW 264.7 cells comparably well as the cap-binding protein EIF4E (Fig. 2d). The canonical CBC has been mainly studied in human HeLa cells. We therefore confirmed the presence and binding of NCBP3 to capped RNA in HeLa cells (Fig. 2d). In line with specific affinity for the cap structure, NCBP3 was precipitated from cell lysates using N7-methylated guanosine-5′-triphosphate (m^7GTP)-coupled beads (Fig. 2e). m^7GTP also precipitated the CBC component NCBP1 but not poly-A-binding protein 1 (PABP1), confirming specificity of this assay.

To functionally assess the potential contribution of NCBP3 to compensate for loss of NCBP2, we employed siRNA-based knockdown and tested for cell growth in HeLa cells. Depletion of either NCBP2 or NCBP3 alone did not considerably affect cell growth (Fig. 2f) despite efficient and specific knockdown as tested by quantitative reverse transcription (qRT)–PCR, western blotting and quantitative proteomics (Fig. 2g; Supplementary Fig. 4; Supplementary Data 3). Although we cannot rule out that

NATURE COMMUNICATIONS | DOI: 10.1038/ncomms9192 ARTICLE

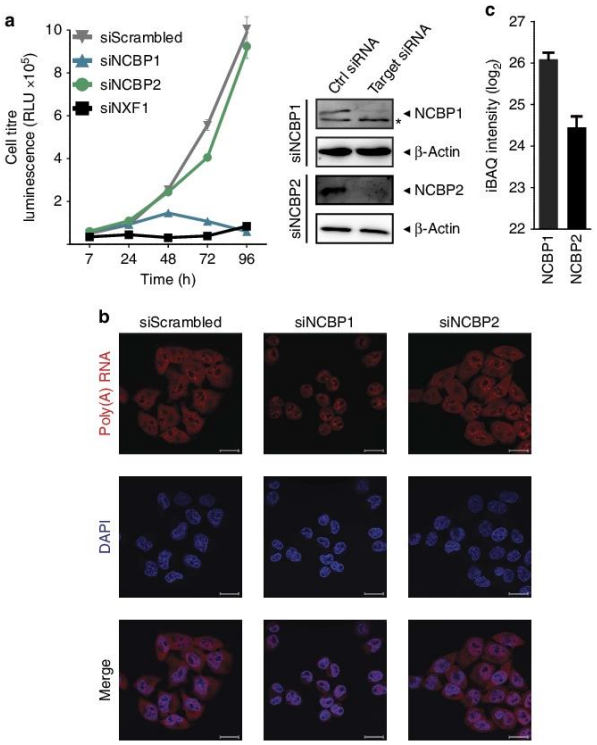


Figure 1 | Cell growth and poly(A) RNA distribution after knockdown of NCBP1 or NCBP2. (**a**) Growth of HeLa cells after RNAi-mediated knockdown (left panel). Cells were treated twice with siRNAs against NCBP1, NCBP2, NXF1 or non-targeting siRNA control (siScrambled), and cell titre determined by a luminescence-based cell viability assay at the indicated time points. The graph shows the mean ± s.d. of two individual treatments measured in triplicates. One representative experiment of three is shown. RLU, relative light units. Knockdown efficiency of NCBP1 and NCBP2 was confirmed by western blotting against indicated proteins (right panel). Asterisk: nonspecific band. (**b**) Poly(A) RNA distribution in HeLa cells 48 h after repeated RNAi-mediated knockdown. HeLa cells were transfected with the indicated siRNAs, and localization of poly(A) RNA (red) was stained by RNA fluorescence *in situ* hybridization (RNA-FISH) using fluorescently labelled oligo (dT) as probe. DAPI (blue) was used to visualize nuclei. Shown confocal images are representative for three independent experiments. Scale bar, 20 μm. (**c**) Abundance of NCBP1 and NCBP2 in HeLa cells. Complete HeLa cell lysates were analysed by LC–MS/MS. Shown are average iBAQ intensities for NCBP1 and NCBP2 of 11 measurements ± s.d.

residual amounts of the proteins are sufficient for biological activity, simultaneous depletion of NCBP2 and NCBP3 with comparable efficiency markedly reduced cell growth and mimicked loss of NCBP1 (Fig. 2f,g; Supplementary Fig. 4c; Supplementary Data 3). An analogous experiment in murine NIH3T3 cells confirmed that depletion of both, Ncbp2 and Ncbp3 also inhibited cell growth in mouse cells, whereas individual knockdown of the two proteins did not (Fig. 2h). In summary, these experiments suggest that loss of NCBP2 and NCBP3 is synthetically lethal, and highlight NCBP3 as a prime candidate for formation of an alternative CBC in mammalian cells.

Mechanism of cap-RNA binding by NCBP3. NCBP1 lacks a canonical RNA or cap recognition domain. Hence, binding to capped RNA is mediated by its partner NCBP2. As suggested above, in an alternative scenario, cap binding could also be mediated by NCBP3. Indeed, homology-based structure prediction of NCBP3 suggested the presence of a canonical RNA recognition motif (RRM) fold (β1–α1–β2–β3–α2–β4)[25] in the N-terminal region of the protein (residues 126–187) (Fig. 3a). The C-terminal part of NCBP3 is unstructured, and a homology search did not reveal the presence of any other predicted domains. Two nuclear localization signals explain the predominant nuclear staining in confocal microscopy and subcellular fractionation analysis (Fig. 2c; Supplementary Fig. 3). Comparative modelling[26] using HHpred resulted in a structural model for the core RRM domain (Fig. 3b) based on the RRM of poly(A)-specific ribonuclease PARN, a protein known to bind the cap analogue m^7GpppG. The RRM of NCBP3 shares high sequence homology (identity of 21.4% for residues 121–191) and critical residues with the RRM of PARN (Fig. 3a–c; Supplementary Fig. 5a)[27–29] as well as with the RRM of NCBP2, although with lower sequence identity (16.4%)

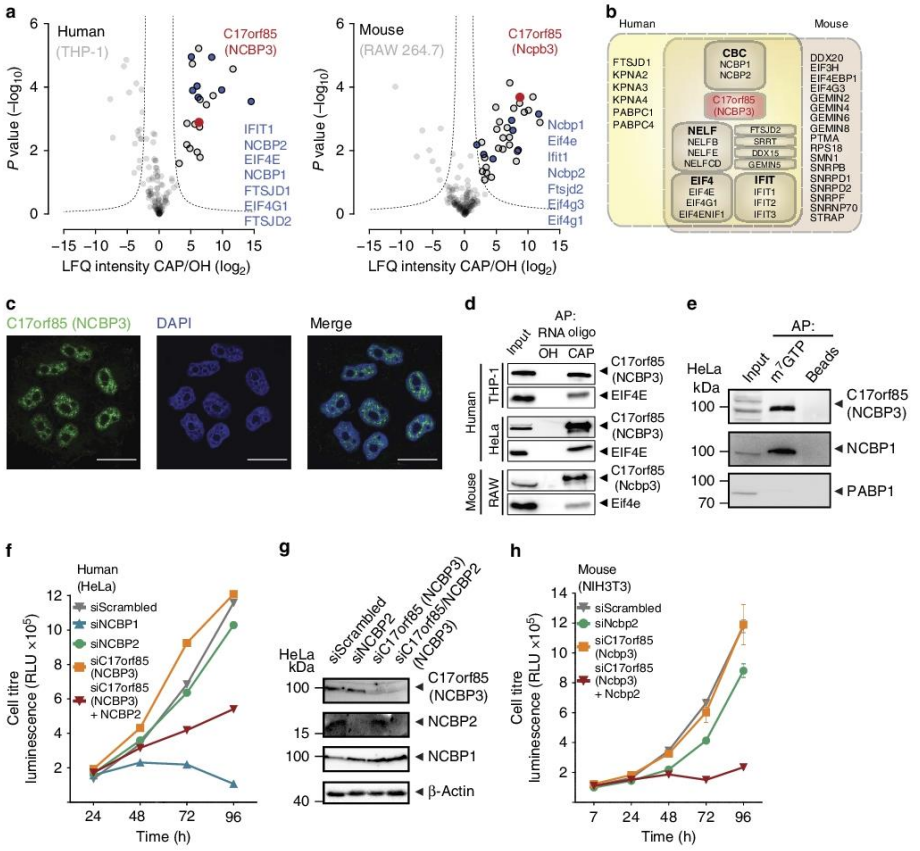


Figure 2 | C17orf85/NCBP3 binds capped RNA and compensates for loss of NCBP2. (**a**) Proteins enriched with 5′-capped (CAP) and uncapped (OH) RNA from THP-1 (human, left panel) and RAW 264.7 (mouse, right panel) macrophages were analysed by LC–MS/MS. Volcano plots show the average degrees of enrichment by CAP- over OH-RNA (ratio of label-free quantification (LFQ) protein intensities; *x* axis) and *P* values (*t*-test; *y* axis) for each protein. Significantly enriched proteins (circled in black) are separated from background proteins by a hyperbolic curve (dotted line). C17orf85/NCBP3 (red dot) and proteins known to bind capped RNA (blue dots) are highlighted. Three independent affinity purifications were performed for both baits. (**b**) Shared enrichment network. Proteins significantly enriched by CAP-RNA from either human or mouse cells (data from Fig. 1a). Proteins belonging to known protein complexes were grouped into boxes (CBC, cap-binding complex; NELF, negative elongation factor; IFIT, interferon-induced proteins with tetratricopeptide repeats; EIF4, eukaryotic translation initiation factor 4). (**c**) Subcellular localization of endogenous NCBP3 assessed by confocal microscopy. HeLa cells were grown on glass coverslips and stained for NCBP3 (green) and DAPI (blue). Scale bar, 20 μM. (**d**) Binding of endogenous C17orf85/NCBP3 to synthetic RNA oligomers (RNA oligo). Western blot analysis of precipitates after affinity purification (AP) with hydroxylated (OH) and m^7G-capped (CAP) RNA oligos from murine RAW 264.7, human THP-1 and HeLa cells using antibodies against C17orf85/NCBP3 and the cap-binding control protein EIF4E. (**e**) Western blot analysis of m^7GTP-precipitated proteins from HeLa cells detecting C17orf85/NCBP3, NCBP1 and PABP1 as control. (**f**,**h**) Growth of HeLa (**f**) and NIH3T3 (**h**) cells after RNAi-mediated knockdown. Cells were treated twice with siRNAs targeting NCBP2 or NCBP3 alone, NCBP3 together with NCBP2, or nonspecific siRNA (siScrambled) as control. The cell titre was determined by a luminescence-based cell viability assay at the indicated time points. The graph shows the mean ± s.d. of two individual treatments measured in triplicates. One representative experiment of three is shown. RLU, relative light units. (**g**) Knockdown efficiency and specificity of NCBP2 and NCBP3 in HeLa cells was confirmed by western blotting against indicated proteins.

(Supplementary Fig. 5b). To study whether NCBP3 has the ability to directly bind capped RNA, we generated recombinant NCBP3 in *Escherichia coli*. Both, the full-length protein and an N-terminal cleavage fragment of NCBP3 precipitated with RNA in a cap-dependent manner (Fig. 3d). The hallmark of the cap structure is a methyl group at the N7 position of the guanosine,

NATURE COMMUNICATIONS | DOI: 10.1038/ncomms9192 ARTICLE

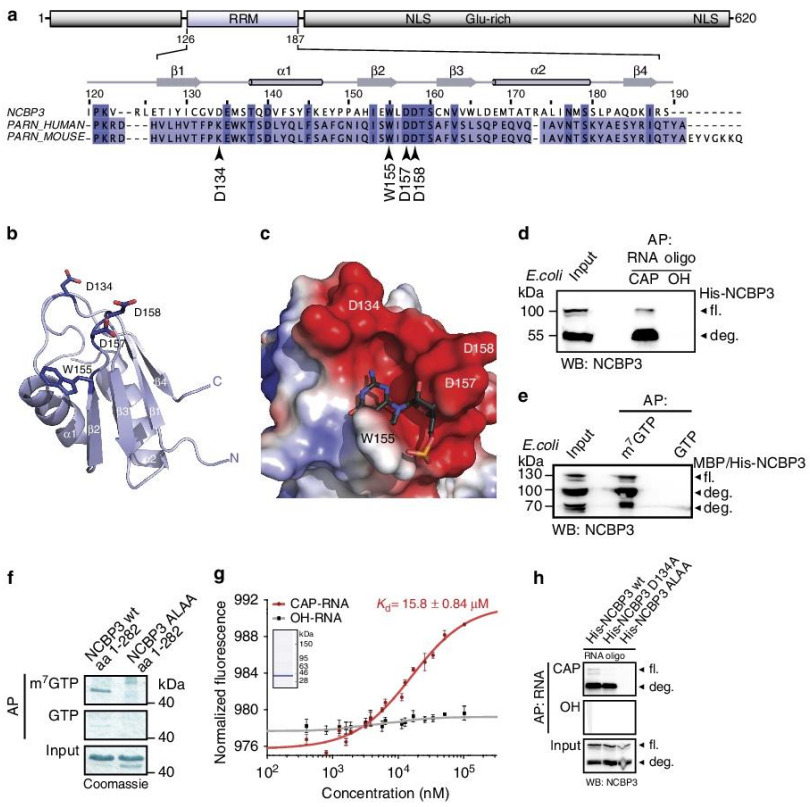


Figure 3 | Biochemical and biophysical properties of the cap-binding region of NCBP3. (**a**) Top, predicted features of NCBP3 and their positions in the full-length protein: RRM fold (aa 126–187, highlighted in blue), nuclear localization signal (NLS) and glutamine-rich region (Glu-rich). Bottom, secondary structure elements of NCBP3 (aa 120–189) and sequence alignment with human and mouse PARN RRMs. Arrowheads show residues chosen for a mutational analysis. (**b**) Homology model of the RRM of NCBP3. Secondary structure elements of the predicted RRM fold and residues chosen for mutational analysis are annotated. (**c**) Modelled RRM domain of NCBP3 superimposed with mouse PARN bound to m7GpppG (for clarity only the m^7Gp moiety is shown, PDB code 3d45) to indicate the putative binding site. The model shows the solvent-accessible surface coloured according to electrostatic charge distribution. (**d**) Binding of *E. coli*-expressed NCBP3 to RNA oligonucleotides. Western blot after affinity purification with 5′-CAP- or OH-RNA using lysate from bacteria expressing recombinant full-length NCBP3. Full-length (fl.) recombinant NCBP3 and N-terminal degradation products (deg.) are indicated. (**e**) m^7GTP versus GTP-affinity purification and western blotting of recombinant NCBP3 from *E. coli* lysates. (**f**) Coomassie-stained SDS-polyacrylamide gel electrophoresis gel after affinity purification of purified wild-type and mutant (ALAA) NCBP3 (aa 1–282) using m^7GTP and GTP as bait. (**g**) Binding affinity of NCBP3 (aa 1–282) to 5′-CAP- or 5′-OH-RNA as determined by microscale thermophoresis (MST). The graph shows normalized fluorescence versus concentration of NCBP3. Mean ± s.d. of three independent measurements are shown. The inset shows the quality of the recombinant NCBP3 (aa 1–282) as analysed by an Agilent protein chip. (**h**) Binding of recombinant wild-type and mutant NCBP3 to 5′-capped RNA. Western blot after affinity purification with biotinylated RNA oligos harbouring either a 5′-CAP or OH structure using lysate from *E. coli* expressing either recombinant full-length wild-type NCBP3 (wt), NCBP3 where aspartic acid at position 134 had been mutated to alanine (D134A), or where tryptophan at position 155 and two aspartic acids at position 157 and 158, respectively, had been mutated to alanines (ALAA).

and genuine cap-binding proteins such as EIF4E have the ability to distinguish between N7-methylated and unmethylated GTP. Consistent with this notion, recombinant NCBP3 selectively bound m^7GTP but not GTP (Fig. 3e). A highly purified N-terminal fragment (position 1–282 amino acids) consisting of the core RRM was sufficient for selective binding (Fig. 3f). Microscale thermophoresis (MST)-based affinity measurements showed that the RRM-containing fragment has an affinity to cap-RNA in the low micromolar range (15.8 ± 0.84 μM), which is comparable to the affinity of PARN to m^7GpppG[28,30] (Fig. 3g).

In case of PARN, a conserved WXDD motif located in loop β2–β3 is involved in m^7GpppG binding. N7-methylated GTP is bound to a surface pocket on the C-terminal RRM of PARN mainly through π-interaction with tryptophan W468 and polar

interactions with two aspartic acids, D471 and D470 (ref. 29). According to the predicted model, NCBP3 has a similar motif (WLDD) forming a negatively charged groove (D157, D158) with an aromatic platform for π-interaction (W155; Fig. 3b,c). To test whether NCBP3 uses these residues for cap-RNA binding, we generated point mutants in this putative RNA-binding motif. Whereas a single D134A mutation in close vicinity of the RNA-binding groove only marginally affected cap binding, mutation of the NCBP3 WLDD motif to ALAA resulted in loss of binding to capped RNA and m^7GTP (Fig. 3f,h; Supplementary Fig. 6a). To ensure that the introduced mutations did not influence the overall integrity and folding of the recombinant proteins, we used circular dichroism (CD) spectroscopy to analyse the secondary structure content. Both samples showed comparable CD spectra, irrespective of the introduced modification, suggesting that both proteins adopt a similar fold (Supplementary Fig. 5b). This shows that NCBP3 binds to the cap structure through a canonical binding mechanism using a core RRM and the conserved WLDD motif.

NCBP3 associates with NCBP1 and mRNA processing factors. To gain further insights into biological functions of NCBP3, we performed quantitative shotgun AP–LC–MS/MS (Supplementary Fig. 7a)[31]. To this aim, we used HeLa cell lines that stably express GFP-tagged NCBP3, NCBP2 and an unrelated control protein (RAB5C) from their endogenous promoters as baits[32,33]. NCBP3 significantly enriched for 88 proteins (Fig. 4a; Supplementary Fig. 7b; Supplementary Data 4) that are mostly related to RNA biogenesis, particularly mRNA transport[1,5–7,34,35] (Supplementary Fig. 7c). Most importantly, the CBC component NCBP1 was found to interact with NCBP3 (Fig. 4a,b; Supplementary Fig. 7b). Besides NCBP1, NCBP3 prominently interacted with components of the TRanscription EXport (TREX) complex (for example, THOC1, -2, -3, -5, -6, -7, DDX39B) that function in mRNA export from the nucleus. NCBP3 further associated with proteins belonging to the exon junction complex (EJC) (MAGOHB, SAP18, EIF4A3, PNN and ACIN1), which are deposited on spliced mRNAs and are involved in mRNA stability. Although NCBP2 precipitates also contained components of the TREX complex and EJC, their enrichment was not statistically significant as compared with control APs (Fig. 4a; Supplementary Fig. 7b). However, NCBP2 precipitates showed high enrichment for the phosphorylated adaptor for RNA export (PHAX) known for its function in U snRNA nuclear export[36] and the negative elongation factor (NELF) complex (consisting of NELFA, NELFB, NELFCD and NELFE), which participates in 3′-end processing of histone mRNAs[13] (Fig. 4a,b; Supplementary Fig. 7b). We independently confirmed the AP–LC–MS/MS data by co-precipitation experiments, followed by western blotting (Fig. 4c). We concluded that NCBP2 and NCBP3 can both bind NCBP1. However, directly comparing NCBP3 and NCBP2 suggests that they preferentially bind proteins involved in mRNA processing/export and snRNA export, respectively.

An alternative CBC. The association of NCBP3 to NCBP1 suggested formation of an alternative CBC. To formally test whether NCBP3 and NCBP1 directly bind to each other, we co-expressed both proteins in *E. coli* and performed co-precipitation experiments. In line with our hypothesis, NCBP1 co-precipitated with NCBP3 but not with a negative control, endorsing a direct interaction between NCBP3 and NCBP1 (Fig. 4d) and suggesting the formation of an alternative CBC that is reminiscent of the canonical CBC consisting of NCBP1 and -2. This result further suggested that NCBP2 and NCBP3 have the ability to individually bridge the association of NCBP1 to capped RNA. To test this in a mammalian system, we depleted NCBP2, NCBP3 or both proteins in HeLa cells and precipitated NCBP1 using m^7GTP beads. Consistent with the presence of two CBCs, NCBP1 could be recovered from m^7GTP beads in the absence of either NCBP2 or NCBP3 (Fig. 4e). However, in the absence of NCBP2 and -3, NCBP1 binding to m^7GTP beads was markedly impaired, indicating that in mammalian cells NCBP1 requires either NCBP2 or NCBP3 for cap-RNA association. This effect was specific for NCBP1, since the cap-binding protein EIF4E precipitated with m^7GTP beads with similar efficiency in all cases.

Binding of NCBP2 to NCBP1 requires residues located in the RRM. Sequence alignments of the NCBP2 and NCBP3 RRM domains showed that the residues critical for NCBP1 binding in NCBP2 are not conserved in NCBP3 (Supplementary Fig. 5b). Consistent with this finding, yeast two-hybrid data indicated that NCBP3 binds NCBP1 through its unstructured C-terminal region[37]. These findings provide evidence that NCBP2 and NCBP3 use different domains to bridge the association of NCBP1 to capped RNA.

Specificity of RNA types bound by NCBP2 and NCBP3. Collectively, our data show that NCBP3 has affinity for the RNA cap structure and bridges NCBP1 binding to m^7GTP in an NCBP2-independent manner. The binding of different RNA processing proteins highlights the possibility that distinct types of RNA associate with either NCBP2 or NCBP3. To test this, we precipitated GFP-tagged NCBP3, NCBP2 or control and sequenced bound RNA by deep sequencing (RNA immunoprecipitation followed by deep sequencing, RIP-Seq). In agreement with the ability of NCBP3 and NCBP2 to associate with the cap structure, both proteins enriched for capped RNAs as compared with control precipitates (Supplementary Data 5). Plotting the individual RNAs enriched by NCBP2 and NCBP3 and colour coding for the respective RNA species revealed an RNA-binding pattern for both proteins: NCBP2 precipitates showed particularly high enrichment for snRNA, lincRNA and asRNA (Fig. 5a,b). Among individual mRNAs, histone mRNAs were mainly apparent in NCBP2 precipitates (Supplementary Data 5). In contrast, NCBP3 associated with mRNA, but did not bind snRNA, and comparably less asRNA and lincRNA (Fig. 5b). The majority of mRNAs bound both NCBP2 and -3 with only a limited number of mRNAs showing more than twofold enrichment in either NCBP3 versus NCBP2 (Fig. 5b, orange) or NCBP2 versus NCBP3 precipitates (Fig. 5b, green). We validated the differences in binding to mRNA and snRNA using qRT–PCR (Fig. 5c). NCBP2 bound mRNAs (MYC, SLC43A3) and snRNAs (U1, U4), whereas NCBP3 selectively bound mRNAs (MYC, SLC43A3) but did not enrich for snRNAs. Collectively, this suggested that snRNA, lincRNA and asRNA preferentially bind NCBP2, while mRNAs can be bound by either NCBP2 or -3. The RIP-Seq data were in strong agreement with the protein–protein interaction data (Fig. 4) that showed selective binding of PHAX (transporting snRNA) and NELF (3′-end processing of replication-dependent histone mRNAs) to NCBP2 (refs 5,13).

NCBP2 and -3 individually contribute to mRNA export. Our data suggested that both, NCBP2 and -3 have the ability to bind mRNA and to recruit proteins involved in RNA processing. We therefore tested whether both proteins individually contribute to nuclear–cytoplasmic transport of mRNA. In RNA-FISH experiments, depletion of either NCBP2 or -3 did not considerably change poly(A) RNA distribution (Fig. 6a). Remarkably, simultaneous depletion of NCBP2 and -3 trapped poly(A) RNA in the nucleus. We quantified the ratio between the nuclear and cytoplasmic poly(A) RNA signal and found similar ratios for

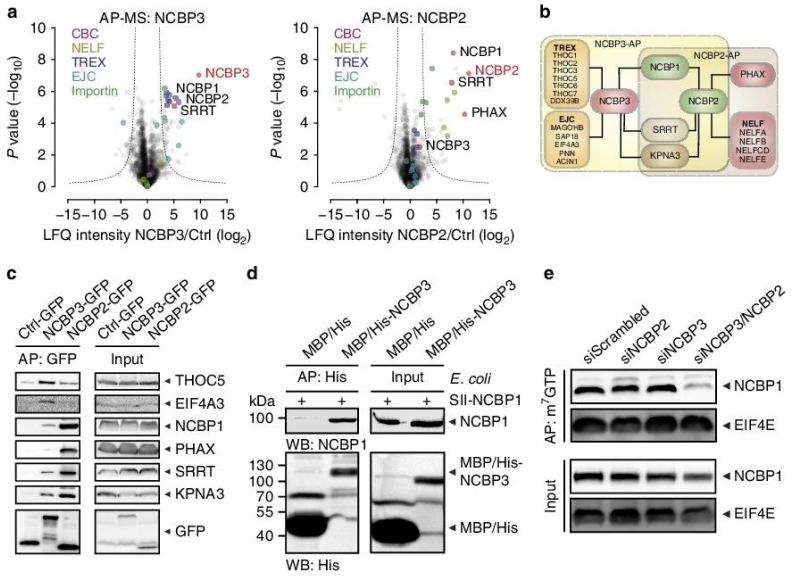


Figure 4 | Association of NCBP3 and NCBP2 with NCBP1 and proteins involved in RNA processing. (**a**) GFP-tagged NCBP3, NCBP2 or control protein (ctrl) stably expressed in HeLa cells were precipitated and associated proteins were analysed by LC–MS/MS. Volcano plots show the average degrees of enrichment by NCBP3 or NCBP2 over ctrl (ratio of LFQ protein intensities; *x* axis) and *P* values (*t*-test; *y* axis) for each protein. Bait proteins (red letters), proteins of the exon junction complex (EJC, turquoise dots) or the TREX complex (purple dots), CBC (magenta dots), NELF (olive dots), importin (green dots) and NCBP3, SRRT and PHAX (red dots) are highlighted. Significantly enriched proteins are separated from the background by a hyperbolic curve. Four independent affinity purifications were performed for each bait. (**b**) Schematic illustration of AP–MS analysis in **a** focusing on proteins involved in RNA biogenesis. (**c**) Western blot analysis of representative GFP-precipitated proteins identified by AP–MS analysis. (**d**) Binding of NCBP1 to NCBP3 in *E. coli*. Western blot after His-precipitation using RNAse-treated lysates from bacteria co-expressing recombinant SII-tagged NCBP1 and full-length MBP/His-tagged NCBP3 or MBP/His as control. (**e**) NCBP2 and NCBP3 are required for NCBP1 association to m^7GTP. Lysates of HeLa cells treated twice with siRNA against NCBP2, NCBP3 or both were used for m^7GTP-affinity purification (m^7GTP-AP), followed by western blot analysis. The m^7GTP-binding protein EIF4E served as control.

co-depletion of NCBP2 and -3 as compared with depletion of NCBP1 (Fig. 6b). To assess whether this is a direct effect of individual NCBP depletion or an indirect effect on expression of mRNA processing factors, we performed quantitative proteomics of siRNA-treated cells, which allows simultaneous quantification of over 5,000 proteins (Supplementary Data 3). As expected, NCBP1 was required for stability of NCBP2 (ref. 10); however, we could not observe an influence on NCBP1 and -2 levels after NCBP3 knockdown and vice versa (Supplementary Fig. 4c). The abundance of mRNA processing factors such as EJC, TREX and NELF components remained similarly unchanged (Supplementary Fig. 4c), collectively emphasizing a direct effect of NCBP depletion on poly(A) RNA export. Our data thus indicate that mRNA can be exported through the activity of either the canonical CBC consisting of NCBP1/-2 or the alternative CBC formed by NCBP1/-3. Furthermore, these data suggest that both CBCs can have functionally redundant activities under steady-state conditions.

NCBP3 activity is critical to inhibit virus growth. Perturbation of the RNA export machinery components often shows loss-of-function phenotypes under challenging conditions[38]. Infection with pathogens induces cellular stress and requires swift expression, processing and export of RNAs coding for antiviral proteins or factors required for virus growth[39,40]. We therefore tested the effect of NCBP3 knockdown on growth of viruses that replicate in the cytoplasm and are not dependent on the cellular mRNA export machinery. HeLa cells treated with siRNA against NCBP3 showed an almost 100-fold increase in growth of vesicular stomatitis virus variant M2 (VSV-M2) (Fig. 6c). Similarly, growth of two other cytoplasmic viruses, Semliki forest virus (SFV) and encephalomyocarditis virus (EMCV), was increased in the absence of NCBP3. These data suggest that NCBP3 becomes fundamental for cellular functions under stress conditions such as virus infection.

Discussion

The canonical CBC consisting of NCBP1 and -2 is considered to be one of the most fundamental parts of the biogenesis machinery of all capped RNAs. This notion is based on the involvement of the CBC in pre-mRNA splicing, 3′-end processing, intranuclear transport, nuclear–cytoplasmic transport of RNA and nonsense-mediated decay and translation[10,14,41]. Exogenously added cap analogues impair RNA metabolism, underlining the major importance of the cap structure in this process. However, *in vivo* antibody-mediated inhibition of NCBP2 only affected nuclear

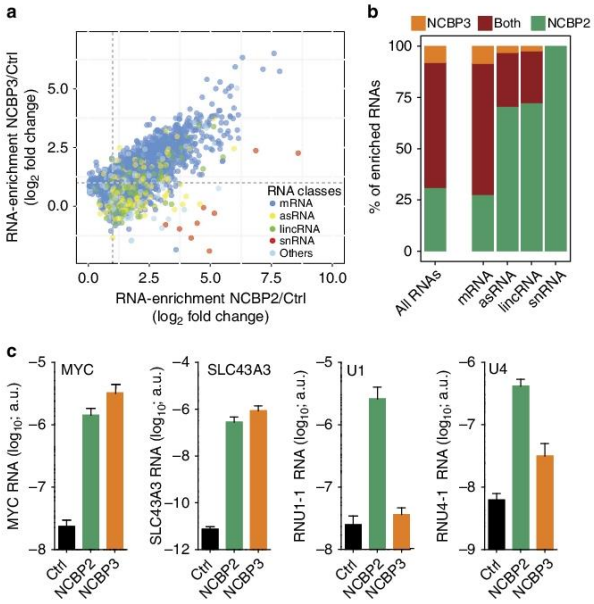


Figure 5 | Global analysis of NCBP3- and NCBP2-bound RNAs. Binding of endogenous RNA to NCBP3 and NCBP2. GFP-tagged NCBP3, NCBP2 or control protein (ctrl) stably expressed in HeLa cells were precipitated and associated RNAs were analysed by deep sequencing on the Illumina HiSeq platform (**a**,**b**) or qRT-PCR (**c**). All affinity purifications were performed in triplicates. (**a**) Scatter plot showing enrichment ($FPKM_{bait}+2/FPKM_{control}+2$) of transcripts binding to NCBP2 (*x* axis) and NCBP3 (*y* axis) quantified on the gene level. RNAs are colour-coded according to their annotated RNA types in Ensembl. (**b**) Percentage of enriched RNA types binding to NCBP3, NCBP2 or both proteins. RNAs have been considered enriched at a fold change bait over control >2, a false discovery rate (FDR) <0.01 and a minimal read count of 10. (**c**) Validation of sequencing data by qRT-PCR. RNA in NCBP2, NCBP3 or control (ctrl) precipitates was amplified by qRT-PCR using specific primers for two mRNAs (MYC and SLC43A3) and two snRNAs (U1 and U4). Data represent the mean ± s.d. a.u., arbitrary units.

export of U snRNAs but not general export of mRNA[8]. Using AP with synthetic capped RNA and unbiased protein–protein interaction studies, we identified the largely uncharacterized protein NCBP3 as a novel genuine cap-binding protein that links NCBP1 to capped RNA, and propose that NCBP1/-3 forms an alternative CBC that is involved in mRNA biogenesis.

Structure-guided binding studies using recombinant NCBP3 clearly established direct cap-binding through a canonical RRM bearing a WLDD motif at the centre of the RNA-binding grove. This motif is also conserved in the canonical cap-RNA-binding protein PARN and mediates affinity to cap-RNA in the low micromolar range. In addition, NCBP3 directly binds NCBP1 when co-expressed in bacteria and thus allows formation of a protein complex with affinity to capped RNA. A direct interaction between NCBP1 and -3 is further suggested from large-scale yeast two-hybrid interaction data[37]. Notably, NCBP2 and -3 were the only two proteins in this screen showing interaction with NCBP1 when being used either as bait or as prey. However, NCBP2 and -3 do not show an apparent homology that would suggest a shared binding mechanism to NCBP1. Both, NCBP2 and -3 have the ability to independently serve as adaptor proteins linking the RNA cap to NCBP1, and only loss of both proteins in HeLa cells reduced NCBP1 association with m^7GTP. Co-depletion of NCBP2 and -3 is synthetically lethal in human und murine fibroblasts, suggesting redundancy between both proteins. This redundancy can be explained by RIP-seq experiments, showing that NCBP2 and -3 share affinity for the majority of mRNA. In addition to mRNA binding, NCBP2 but not NCBP3 has the ability to bind snRNA. Consistent with the selective affinity for snRNA, only NCBP2 binds the snRNA transport protein PHAX. Our results consolidate reports that are incompatible with the current model of a single CBC for RNA export: while inhibition of NCBP2 selectively impairs U snRNA nuclear–cytoplasmic export, but not of mRNA[9], NCBP1 is strictly required as an adaptor for mRNA export[15].

It is commonly accepted that EJC components, the TREX complex, PHAX and the NELF complex are recruited to RNAs by virtue of the CBC. We here show that EJC and TREX precipitate superior with NCBP3 as compared with NCBP2, whereas PHAX and the NELF complex are exclusively enriched in NCBP2 precipitates. It is tempting to speculate that binding of either NCBP2 or -3 induces a conformational change of NCBP1 that in turn allows differential binding of adaptor proteins. Structural analysis of NCBP1 in the context of NCBP2 and -3 bound to capped RNA, as well as binding assays with purified proteins will be necessary to study this in detail.

Given the association with proteins of the EJC and TREX complex, NCBP3 may play a primary role in biogenesis of spliced mRNA. Such a role is supported by genetic correlation: although NCBP3 is remarkably conserved during evolution from fungi

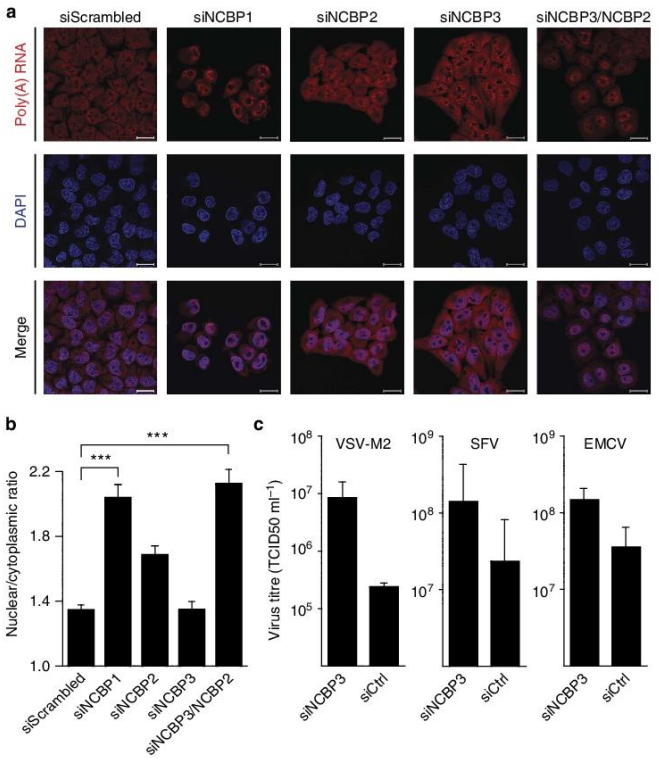


Figure 6 | Functional analysis of NCBP3 at normal and virus-challenged growth conditions. (**a**,**b**) Poly(A) RNA distribution in HeLa cells 72 h after RNAi-mediated knockdown. HeLa cells were transfected with the indicated siRNAs and localization of poly(A) RNA (red) was detected by fluorescence *in situ* hybridization (FISH) using fluorescently labelled oligo (dT) as probe. DAPI (blue) was used to visualize nuclei. (**a**) Representative images of poly(A) RNA distribution after RNAi-mediated knockdown. Scale bar, 20 μm. (**b**) The nuclear to cytoplasmic ratio of poly(A) RNA intensity was determined for a minimum of 125 cells per condition. Data represent the average ratio of the nuclear/cytoplasmic poly(A) RNA signal ± s.d. $^{***}P < 0.001$ as analysed by one-way analysis of variance statistics with Tukey's post-test. (**c**) Virus growth in HeLa cells after RNAi-mediated knockdown. HeLa cells were treated with siRNA against NCBP3 or nonspecific siRNA (siGFP) as control and infected with mutant vesicular stomatitis virus (VSV-M2; M51R substitution in the matrix protein), Semliki Forest virus (SFV) or encephalomyocarditis virus (EMCV). Virus titres in supernatants were determined by TCID50 24 h post infection.

to humans, we could not find homologues for NCBP3 in *Saccharomyces cervisiae*. Notably, in *S. cervisiae* only 3% of genes contain introns and only six genes contain two introns[42]. In contrast, the closely related *Schizosaccharomyces pombe* expresses an NCBP3 homologue and codes for introns in 43% of its genes[43]. Loss of NCBP3 in *S. cerevisiae* may therefore be a result of evolutionary adaptation.

The canonical CBC (consisting of NCBP1/-2) guides mRNA from the nucleus into the cytoplasm. Similarly, NCBP3 localizes to the nucleus and partially to the cytoplasm, suggesting that the alternative CBC drives mRNA export to the cytoplasm. RNA bound to the canonical CBC or EIF4E is sensitive to the nonsense-mediated mRNA decay (NMD) pathway[44–46]. Likewise, NCBP1/-3 associated with exported mRNA may be a target of NMD, particularly since the EJC, which co-purified with NCBP3, is known to promote NMD.

Given that RNA metabolism is of central importance for all physiological and pathophysiological processes, we envisage a central role of the alternative CBC in various diseases, as known for many proteins involved in RNA splicing, 3′-end processing and degradation[2,41,47]. A reason for metazoans to evolve an alternative CBC could be the availability of an additional control mechanism to respond to environmental cues and thus allow the cell to respond swiftly to appropriate changes in gene expression. This may be particularly important under environmental stress, such as occurring during virus infection. Loss of NCBP3 increased growth of viruses that replicate in the cytoplasm. The exact mechanisms in antiviral defence and the likely involvement in other diseases will be the focus of further studies. Our data show that RNA processing conveys additional complexity that may allow regulatory possibilities for both, cell intrinsic modulation as well as therapeutic intervention.

Methods

Cells and reagents. HeLa S3 (CCL-2.2) and Vero E6 cells (CRL-1586) were purchased from ATCC. Human THP-1, NIH3T3 and murine RAW 264.7 macrophages have been described previously[48]. HeLa Kyoto cells stably expressing GFP-tagged human C17orf85/NCBP3, NCBP2 and RAB5C from bacterial artificial chromosomes under control of their endogenous promoter were kindly provided by Ina Poser and Tony Hyman[33]. All cell lines were maintained in DMEM (PAA Laboratories) containing 10% fetal calf serum (PAA Laboratories) and antibiotics (100 U ml^{-1} penicillin and 100 µg ml^{-1} streptomycin). Streptavidin-agarose beads were obtained from Novagen, GFP-Trap-coupled agarose beads were from Chromotek, and γ-Aminohexyl m^{7}GTP- and GTP-agarose beads from Jena Bioscience and Biorbyt, respectively. Primary antibodies (1:1000 dilution for western blotting if not stated otherwise) used in this study were as follows: C17orf85 (Atlas Antibodies; HPA008959), NCBP1, SRRT, EIF4A3 and ALYREF (Thermo Scientific; PA 5–30098, PA5–31593, PA5–30329 and MA 1–26754), THOC5, PHAX and KPNA3 (Novus Biologicals; NBP1–19160, NBP2–22268 and NB100–81650), GFP (Invitrogen; A6455), PABP1 and β-tubulin (1:500 dilution for western blotting) (Santa Cruz; 10E10, sc-9104), histone H3 (Abcam; ab1791–100) and EIF4E (Cell Signaling; C46H6). Antibodies against NCBP2 (1:500 dilution for western blotting) were raised by immunizing rabbits with recombinant full-length protein purified from *E. coli*. Antibodies against β-actin (Santa Cruz; sc-47778), His-tag (Santa Cruz; sc-8036) and secondary antibodies detecting mouse or rabbit IgG (Jackson ImmunoResearch) were horseradish peroxidase-coupled. 4′,6-diamidino-2-phenylindole (DAPI) and secondary antibodies used for immunofluorescence were purchased from Invitrogen. Vesicular stomatitis virus M2 (mutant VSV with the M51R substitution in the matrix protein), SFV and EMCV have been described previously[49].

RNAi-mediated knockdown. Duplex siRNAs were transfected using either siPrime transfection reagent (GE Healthcare) or the Neon Transfection System (Invitrogen) for target gene knockdowns. Transfection was performed according to the manufacturer's instructions for HeLa or NIH3T3 cells. Briefly, we transfected 200 pmol of siRNA per 1×10^6 cells and repeated transfection 48 h later under the same conditions. For experiments where two genes were silenced simultaneously, we added scrambled-siRNA to single-gene knockdown controls. Cells were analysed at the indicated time points after the second transfection. Duplex siRNAs were either purchased from Qiagen or synthesized by the Core Facility at the MPI of Biochemistry. siRNA target sequences were as follows: human NCBP3 (#1: 5′-AAGAGCCGGTTAGATAACTTA-3′, #2: 5′-TCAGCGGGACGTGATCAA GAA-3′, #3: 5′-ATGACTATGTATGCTGACGAA-3′, #4: 5′-CAGATTGAAGTTA GTCGGGAA-3′), mouse NCBP3 (#1: 5′-TCAGATGTACATAGTAGGCTA-3′, #2: 5′-TCGCTTAGGATCTACACCCAA-3′, #3: 5′-TGGGTTGATGTTGAACAA TTT-3′, #4: 5′-CAGATTGAAGTCAGCCGGGAA-3′), human NCBP1 (#1: 5′-CCACAGATGATTGCTGTACTA-3′, #2: 5′-CAGGAACGGCACATCCTA AGA-3′, #3: 5′-CAGGTATGGACTGCTGATAAA-3′, #4: 5′-AGCCGTGTATTTG GTCCGTTT-3′), mouse NCBP1 (#1: 5′-AUGCAGAAAUGGACCGAAU-3′, #2: 5′-CGUCUGGACACGAUGAGUA-3′, #3: 5′-GGUACGAUGUGAAACGGAU-3′, #4: 5′-AGGCCUAACACUCGCGUUU-3′), human NCBP2 (#1: 5′-GCCAUGCG GUACAUAAAUG-3′, #2: 5′-UGGAUGAACUUAUCGUUAA-3′, #3: 5′-GCAU GAGAUAGCCUAAUAA-3′, #4: 5′-AUGAGUAUCGGCAGGACUA-3′), mouse NCBP2 (#1: 5′-CAGCAAAAGUGGUGAUAUA-3′, #2: 5′-GCAAUGCGGUACA UAAACG-3′, #3: 5′-GUAUGGACGUGGACGGUCU-3′, #4: 5′-ACGAGUAUCG GGAGGACUA-3′), huNXF1 (#1: 5′-GAACACGATGATGAACGCGTT-3′, #2: 5′-GATGACATGTCTAGCATTGTT-3′), GFP (5′-AAGCAGCACGACUUCUUC AAG-3′) and scrambled (5′-AAGGTAATTGCGCGTGCAACT-3′).

Cell growth assays. For cell growth assays, 1×10^4 cells were seeded in 24-well dishes and cell titres determined after the indicated time points using CellTiter-Glo (Promega) according to the manufacturer's instructions with the following modifications: cells were washed once with 1 × PBS and incubated with 100 µl CellTiter-Glo reagent (diluted 1:5 in 1 × PBS) for 10 min. Luminescence was measured using an Infinite 200 PRO series microplate reader (Tecan).

Immunofluorescence and RNA-FISH. For immunofluorescence, HeLa cells grown on coverslips were fixed with 4% (w/v) paraformaldehyde for 10 min, permeabilized with 0.1% (v/v) Triton X-100 for 10 min and washed three times with blocking buffer (1 × PBS containing 0.1% fetal calf serum (v/v)). Cells were then incubated with primary antibody (1:500 dilution), followed by incubation with secondary antibodies (1:200 dilution) covalently linked to fluorophores and DAPI. Coverslips were mounted on microscope slides using ProLong Gold Antifade Reagent (Molecular Probes). Confocal imaging was performed using a LSM780 confocal laser scanning microscope (Zeiss, Jena, Germany) equipped with a Plan-APO × 63/numerical aperture 1.46 oil immersion objective (Zeiss).

For RNA-FISH analysis, cells grown on coverslips were fixed with 4% (v/v) PFA for 10 min, followed by 100% (v/v) ethanol for 10 min and washed in 70% (v/v) ethanol. Subsequently, cells were incubated with 1 M Tris-HCl pH 8 for 10 min, followed by incubation with hybridization buffer (2 ng µl^{-1} 5′-Cy5-labelled oligo $(dT)_{45}$ (Sigma), 0.5 µg µl^{-1} transfer RNA (Ambion), 1% (v/v) bovine serum albumin, 10% (v/v) dextran sulphate, 20% (v/v) deionized formamide and 2 × sodium saline citrate (SSC)) for 3 h at 30 °C in a humid chamber. After washing once with 4 × SSC and twice with 2 × SSC, cells were incubated with DAPI diluted in 2 × SSC containing 0.1% (v/v) Triton X-100. Finally, cells were washed twice with 2 × SSC and coverslips placed on a microscope slide. Confocal analysis was performed as described above. Quantification of area and intensity of cytoplasm versus nuclei was performed with Volocity 6.3 analysis software (PerkinElmer, Waltham, MA, USA). With the protocol four populations of objects were defined: (1) Nuclei were masked using a DAPI channel with a threshold range of 13–100% and restriction to a minimum object size of 50 µm^2. 'Fill holes in objects' was applied to remove holes in the mask; (2) cytoplasm including nuclei was masked using a Cy5 channel with an intensity-based threshold of 6–255 and restriction to a minimum object size of 100 µm^2. (3) A mask for measurements in the cytoplasm was created using 'Exclusively combine' of the first two masks. (4) Finally, a mask for measurements in the Cy5-stained nuclei was created by subtracting (3) from (2). Within the mask (3) and (4), area and summarized intensity were measured and intensities were normalized to area. The nuclear to cytoplasmic ratio was calculated using the intensities from a minimum of 125 cells.

Generation of synthetic and *in vitro* transcribed RNA. Synthetic oligoribonucleotides (Chemgenes Corporation) with a 3′-terminal amino linker harbouring either an N7-methylated cap structure (CAP) or a 5′-hydroxyl group (OH) were generated as described previously[19]. RNA oligomers were modified at the 3′ end either with biotin using biotin-*N*-hydroxysuccinimide ester (Epicentre) or fluorophore using DyLight 488 NHS ester according to the manufacturer's instructions and purified by reverse-phase-HPLC. Biotinylated 7SK-as RNA bearing a 5′-CAP or –OH group, respectively, were synthesized by *in vitro* transcription and enzymatic modification of the 5′ termini[19] and purified using the NucleoSpin RNA II kit (Macherey-Nagel).

Affinity purifications. For APs with biotin-labelled RNA, streptavidin-affinity resin was first incubated with RNA in TAP buffer (50 mM Tris pH 7.5, 100 mM NaCl, 5% (v/v) glycerol, 0.2% (v/v) Nonidet-P40, 1.5 mM $MgCl_2$ and protease inhibitor cocktail (EDTA-free, cOmplete; Roche)) in the presence of 40 U RNase inhibitor (Fermentas) for 60 min at 4 °C on a rotary wheel and excess RNA removed by three washes with TAP buffer. Cell lysates were prepared by snap-freezing cells in liquid nitrogen, incubation in TAP buffer for 30 min on ice and clarification of the lysate by centrifugation at 16,000*g*. RNA-coated beads were incubated with 2 mg of clarified lysate for 60 min at 4 °C, washed three times with TAP buffer, boiled in Laemmli buffer for 10 min at 95 °C and subjected to SDS–polyacrylamide gel electrophoresis and western blot analysis. Uncropped western blots are provided in Supplementary Fig. 8.

To detect proteins binding to m^{7}GTP- and GTP-coupled beads, cell lysate was sonicated using a Branson Sonifier 250 with 15 pulses at a duty cycle of 50% and an output control of 1. The lysate was further incubated on ice for 15 min and clarified by centrifugation at 16,000*g* for 10 min at 4 °C. Similarly, we used GFP-Trap beads (Chromotek) for affinity purification of proteins binding to GFP-tagged baits. The sonicated lysate was further treated with 10 U of Benzonase (Core facility, MPI-B) for 30 min on ice before clarification of the lysate.

For co-precipitation of proteins expressed in *E. coli* expression of recombinant proteins was induced overnight at 18 °C in *E. coli* strain BL21-AI using 1 mM isopropyl-β-D-thiogalactoside (Thermo) and 0.2% L-(+)-Arabinose (Santa Cruz). Cells were lysed in HIS buffer (50 mM HEPES pH 7.5, 500 mM NaCl, 5% glycerol and 20 mM imidazole), including 50 U ml^{-1} of Benzonase (Core Facility, MPI-B) and protease inhibitor cocktail (EDTA-free, cOmplete; Roche) using sonication and lysates clarified by centrifugation at 16,000*g* for 10 min at 4 °C. Lysates were used for AP using Ni-NTA spin columns according to the manufacturer's instructions (Qiagen). Ni-NTA-bound proteins were eluted using HIS buffer containing 500 mM imidazole. Eluates were boiled in Laemmli buffer for 10 min at 95 °C and subjected to SDS–polyacrylamide gel electrophoresis and western blot analysis.

Quantitative LC–MS/MS-based proteomics and bioinformatics. To detect and quantify RNA-binding proteins and proteins bound to GFP-tagged baits by AP and MS, samples were prepared as described above. After the final three washes in TAP buffer, samples were, in addition, washed twice with TAP buffer lacking Nonidet-P40 to remove residual detergent. Three independent APs were performed for each bait. Sample preparations and LC–MS/MS analysis was performed as described previously[19]. Briefly, samples were sequentially digested with LysC (Wako Chemicals USA) and trypsin (Promega), acidified with 0.1% TFA, desalted with C18 stage tips and analysed by liquid chromatography coupled to MS either on Orbitrap XL or Q Exactive instruments (Thermo Fisher Scientific). For analysis of interaction proteomics data, MS raw files were processed with MaxQuant software versions 1.4.1.8 and 1.4.2.3 (ref. 50) using the built-in Andromeda engine to search against human and mouse proteomes (UniprotKB, release 2012_06) containing forward and reverse sequences. In MaxQuant, the label-free quantification (LFQ)[51] algorithm and Match Between Runs option were used. Only proteins identified on the basis of at least two peptides and a minimum of three quantification events in at least one experimental group were considered. LFQ protein intensity values were log-transformed and missing values filled by

imputation. Specific enrichment was determined by multiple equal variance *t*-tests with permutation-based false discovery rate (FDR) statistics, performing 250 permutations. FDR thresholds and S_0 parameters were empirically set to separate background from specifically enriched proteins.

For total proteome analysis, HeLa cells were electroporated twice with siScrambled or siRNAs against NCBP1, NCBP2, NCBP3 or NCBP2 and NCBP3 and lysed 5 days after knockdown. Three knockdowns were performed in parallel. Cells were pulsed with SILAC Arg10- and Lys8-containing medium for the last 6 h (ref. 49). Cells were lysed in SDS lysis buffer (50 mM Tris pH 7.5, 4% SDS (v/v)), boiled for 5 min at 95 °C and sonicated. Fifty-microgram aliquots were reduced with 10 mM dithiothreitol, alkylated with 55 mM iodoacetamide and precipitated with 80% acetone (v/v) at − 20 °C overnight. Pellets were dissolved in 6 M urea—2 M thiourea, digested with LysC and trypsin and desalted peptides analysed by LC–MS/MS on a Q Exactive instrument. Raw MS data were processed with MaxQuant 1.5.1.6 using the LFQ and iBAQ algorithms and the Match Between Runs option only considering light labelled amino acids. LFQ intensities were log-transformed and left non-imputed.

For functional annotation analysis, we used DAVID[52] (database for annotation, visualization and integrated discovery; http://david.abcc.ncifcrf.gov). Proteomics data were analysed using Perseus, results were plotted using R (www.R-project.org) and GraphPad Prism version 5.02 and visually adapted using Adobe Illustrator.

RNA immunoprecipitation sequencing. HeLa cells expressing GFP-tagged human NCBP3, NCBP2 or RAB5C were pelleted by centrifugation at 500*g* for 10 min at 4 °C and washed twice with ice-cold PBS. Cells were lysed in an equal volume of RIP lysis buffer (10 mM HEPES pH 7.0, 100 mM KCl, 5 mM $MgCl_2$, 25 mM EDTA, 0.5% (v/v) Nonidet-P40, 1 mM dithiothreitol, protease inhibitor cocktail (EDTA-free, cOmplete; Roche)) for 30 min on ice in the presence of 100 U ml^{-1} RNase inhibitor (Fermentas) and lysates clarified by centrifugation at 9,000*g* and 4 °C for 10 min. Clarified lysates were incubated with GFP-Trap beads (Chromotek) in RIP wash buffer (50 mM Tris pH 7.4, 150 mM NaCl, 1 mM $MgCl_2$, 0.05% (v/v) Nonidet-P40) containing 25 mM EDTA, protease inhibitors and 100 U ml^{-1} RNase inhibitor for 4 h at 4 °C on a rotary wheel, followed by five washes with RIP wash buffer. Beads were then resuspended in Trizol (Invitrogen) and RNA isolated using the Direct-zol RNA MiniPrep kit (Zymo Research) according to the manufacturer's instructions. Three RNA immunoprecipitations per bait were carried out in parallel. RNA quality was assessed on a Genetic Analyzer (Agilent) and TruSeq RNA library construction and next-generation sequencing were performed by the Max Planck-Genome-Center Cologne, Germany (http://mpgc.mpipz.mpg.de/home/). All samples were sequenced on an Illumina HiSeq2500 platform at 15 million 100-bp single reads per sample. After quality control of the sequencing libraries, reads were trimmed and mapped against the Ensembl genome annotation and the human genome assembly (hg19/GRCh37) using Tophat2 (ref. 53). Reads mapping to ribosomal RNAs or the mitochondrial genome were removed. RNAs binding to NCBP3 or NCBP2 were identified by differential quantification (bait over control) against the Ensembl genome annotation using cuffdiff from the cufflinks package[54]. RNAs with fold changes >2, FDR corrected *P* values <0.01 and minimal read counts of 10 were considered as enriched. To discover preferences of NCBP3 and NCBP2 for different RNA species, we extracted RNA types and gene-model-related features from the Ensembl annotations and plotted them using custom scripts.

Quantitative RT–PCR. RNA was reverse transcribed using the RevertAid H Minus First Strand cDNA Synthesis kit (Fermentas) and quantified by quantitative RT–PCR using the QuantiFast SYBR Green RT-PCR kit (Qiagen) and a CFX96 Touch Real-Time PCR Detection System (Bio-Rad). Each cycle consisted of 10 s at 95 °C and 30 s at 60 °C, followed by melting curve analysis. Primer sequences were as follows: huMYC (5′-CAGTGGGCTGTGAGGAGGTT-3′ and 5′-CAGGCTCC TGGCAAAAGGT-3′), huSLC43A3 (5′-CATGACATTCCCCACTGGCT-3′ and 5′-GATTCCCCCAATGGTGAGCA-3′), huRNU1-1 (5′-ACTTACCTGGCAGGG GAGATAC-3′ and 5′-ACATCCGGAGTGCAATGGATAA-3′), huRNU4-1 (5′-GCGCAGTGGCAGTATCGTAG-3′ and 5′-GGCGGGGTATTGGGAAA-3′), huNCBP1 (5′-GGCTGCAGCAGATCTTCCTA-3′ and 5′-TCTCCAGGGTCACC ATGTACT-3′), huNCBP2 (5′-TTTCCAACATATAACGTACAGCTTTT-3′ and 5′-CTACGTGGAGCTGAGCCAGT-3′), huNCBP3 (5′-GCAGGAAGACAGTTC AGATG-3′ and 5′-ACTTCTTCTGGCTGCTCCAA-3′).

Homology searches and homology modelling. Orthologous sequences were collected using the web-server morFeus[55] (http://chimborazo.biochem.mpg.de/morfeus). The multiple sequence alignment was generated using mafft[56] and submitted to the Gblocks web-server[57] to select conserved blocks for further phylogenetic analysis. The phylogenetic tree was calculated with Phyml[58] using standard parameters and bootstrapping with 100 iterations. The resulting tree was displayed using Dendroscope[59] and visually adapted using Adobe Illustrator.

A structural model of the RRM domain of C17orf85/NCBP3 (UniProt entry Q53F19) was obtained using the HHpred server[60] for the detection of homologous structures, and subsequent comparative modelling using Modeller[26] (as integrated in the Bioinformatics Toolkit (http://toolkit.tuebingen.mpg.de)) with the RRM of human poly(A) adenylate-binding protein 1 (PDB code 4f02) as a template model. The final best model was chosen according to DOPE (Descrete Optimized Protein Energy) scores[26]. Structure analysis and electrostatic surface potential calculations were done in Pymol (The PyMOL Molecular Graphics System, Version 1.5.0.4, Schrödinger, LLC).

Nuclear/cytoplasmic fractionation of mammalian cells. Cells grown in a 10-cm dish were fractionated as described previously[61] with the following modifications: RNLa buffer was replaced by C/N buffer (10 mM Tris-HCl pH 7.5, 140 mM NaCl, 1.5 mM $MgCl_2$, 10 mM EDTA and 0.5% (v/v) NP-40) containing protease inhibitor (EDTA-free, cOmplete, Roche) and Benzonase Nuclease (Novagen). For western blot analysis, we loaded 4% of the total cell lysate and 20% of the cytoplasmic and nuclear fractions.

Cloning and expression of recombinant proteins. The long isoform of C17orf85/NCBP3 (GenBank NM_001114118.2) was cloned from human complementary DNA (cDNA) into vector pDONR221 (Invitrogen). Mutations and truncations in C17orf85/NCBP3 were introduced by site-directed mutagenesis and PCR, respectively. Expression constructs containing N-terminal His-tagged full-length C17orf85/NCBP3 (pETG10A-NCBP3) or its mutant versions (pETG10A-NCBP3-D134A and -ALAA) were generated using the Gateway system (Invitrogen) as described earlier[49]. Expression constructs for N-terminal MBP/His-tagged full-length NCBP3 and NCBP3 expressing amino acids 1–282, as well as SII-tagged NCBP1 were generated by SLIC cloning as described elsewhere[62]. pETM44 expressing MBP/His was a kind gift from the Core Facility at the MPI of Biochemistry. Sequences of all cloning primers are available on request. Expression of recombinant proteins was induced overnight at 18 °C in *E. coli* strain BL21-AI using 1 mM isopropyl-β-D-thiogalactoside (Thermo) and 0.2% L-(+)-Arabinose (Santa Cruz). Cells were lysed in BL buffer (50 mM Tris-HCl pH 7.5, 500 mM NaCl, 5% glycerol, 5 mM β-mercaptoethanol, 20 mM imidazole and protease inhibitor cocktail (EDTA-free, cOmplete; Roche)) and the cleared lysate was used for further experiments or AP using a HisTrap HP column (GE Healthcare; 17-5247-01). Recombinant untagged wild-type and mutant (ALAA) NCBP3 1–282aa used for MST and CD measurements was purified by the Core Facility at the MPI of Biochemistry. Identity of recombinant NCBP3 was confirmed by MS.

Affinity measurements. For MST affinity measurements, labelled RNA oligos (CAP0- and OH-RNA-DyLight) were diluted in MST-1 buffer (50 mM Tris-HCL pH 7.4, 150 mM NaCl, 10 mM $MgCl_2$, 0.05% (v/v) Tween-20) to a final concentration of 50 nM. A twofold dilution series (16 samples) of NCBP3 1–282aa starting from 204 μM were performed in MST-2 buffer (20 mM Tris-HCL pH 7.4, 300 mM NaCl, 5% (v/v) glycerol). The two components were mixed in equal volumes and the mixture (~8 μl) was then measured in the Monolith NT.015T at room temperature using a blue light-emitting diode (LED). The MST settings were: LED power (20%), MST laser power (80%), fluorescence before (5 s), MST on (30 s) and fluorescence after (5 s). The normalized fluorescence signal ($F_{norm} = F_{hot}/F_{cold}$; hot = 34.5 s and cold = 5.5 s) was analysed and plotted by Monolith NT.015T analysis software. Three individual measurements were combined to calculate the K_d values with s.e.

CD spectroscopy. Far ultraviolet CD spectra of wild-type and mutant versions of NCBP3 were recorded on a Jasco J-810 CD-Photometer at room temperature in 20 mM sodium phosphate buffer pH 7.4 and 50 mM NaF. For each sample and the buffer (baseline), four scans were recorded and averaged. The averaged baseline spectrum was subtracted from the averaged sample spectra and the resulting spectra were smoothed using an FFT (Fast Fourier Transform) filter (as part of the software package). Measurements were only made down to wavelengths where the instrument dynode voltage indicated the detector was still in its linear range (190 nm). Spectra are shown as the mean residue ellipticity. Secondary structure compositions were estimated using the CONTINLL program[63].

Virus infections and determination of virus titres. To determine the impact of siRNA-mediated knockdown of NCBP3 on virus growth, aliquots of 1×10^5 HeLa cells that had been transfected for 48 h with siRNA targeting NCBP3 or GFP as control were infected with VSV-M2, EMCV or SFV at a multiplicity of infection of 0.01. At 24 h post infection, virus titres in supernatants were quantified by 50% tissue culture-infective dose (TCID50) assays on Vero E6 cells.

References

1. Topisirovic, I., Svitkin, Y. V., Sonenberg, N. & Shatkin, A. J. Cap and cap-binding proteins in the control of gene expression. *Wiley Interdiscip. Rev. RNA* **2,** 277–298 (2011).
2. Hallegger, M., Llorian, M. & Smith, C. W. Alternative splicing: global insights. *FEBS J.* **277,** 856–866 (2010).
3. Konarska, M. M., Padgett, R. A. & Sharp, P. A. Recognition of cap structure in splicing in vitro of mRNA precursors. *Cell* **38,** 731–736 (1984).
4. Edery, I. & Sonenberg, N. Cap-dependent RNA splicing in a HeLa nuclear extract. *Proc. Natl Acad. Sci. USA* **82,** 7590–7594 (1985).

5. Muller-McNicoll, M. & Neugebauer, K. M. How cells get the message: dynamic assembly and function of mRNA-protein complexes. *Nat. Rev. Genet.* **14,** 275–287 (2013).
6. Gonatopoulos-Pournatzis, T. & Cowling, V. H. Cap-binding complex (CBC). *Biochem. J.* **457,** 231–242 (2014).
7. Kohler, A. & Hurt, E. Exporting RNA from the nucleus to the cytoplasm. *Nat. Rev. Mol. Cell Biol.* **8,** 761–773 (2007).
8. Izaurralde, E. *et al.* A cap-binding protein complex mediating U snRNA export. *Nature* **376,** 709–712 (1995).
9. Izaurralde, E. *et al.* A nuclear cap binding protein complex involved in pre-mRNA splicing. *Cell* **78,** 657–668 (1994).
10. Mazza, C., Ohno, M., Segref, A., Mattaj, I. W. & Cusack, S. Crystal structure of the human nuclear cap binding complex. *Mol. Cell* **8,** 383–396 (2001).
11. Cook, A., Bono, F., Jinek, M. & Conti, E. Structural biology of nucleocytoplasmic transport. *Annu. Rev. Biochem.* **76,** 647–671 (2007).
12. Moore, M. J. & Proudfoot, N. J. Pre-mRNA processing reaches back to transcription and ahead to translation. *Cell* **136,** 688–700 (2009).
13. Narita, T. *et al.* NELF interacts with CBC and participates in 3' end processing of replication-dependent histone mRNAs. *Mol. Cell* **26,** 349–365 (2007).
14. Pabis, M. *et al.* The nuclear cap-binding complex interacts with the U4/U6.U5 tri-snRNP and promotes spliceosome assembly in mammalian cells. *RNA* **19,** 1054–1063 (2013).
15. Cheng, H. *et al.* Human mRNA export machinery recruited to the 5' end of mRNA. *Cell* **127,** 1389–1400 (2006).
16. Luo, B. *et al.* Highly parallel identification of essential genes in cancer cells. *Proc. Natl Acad. Sci. USA* **105,** 20380–20385 (2008).
17. Cheung, H. W. *et al.* Systematic investigation of genetic vulnerabilities across cancer cell lines reveals lineage-specific dependencies in ovarian cancer. *Proc. Natl Acad. Sci. USA* **108,** 12372–12377 (2011).
18. Kittler, R. *et al.* Genome-scale RNAi profiling of cell division in human tissue culture cells. *Nat. Cell Biol.* **9,** 1401–1412 (2007).
19. Habjan, M. *et al.* Sequestration by IFIT1 impairs translation of 2'O-unmethylated capped RNA. *PLoS Pathog.* **9,** e1003663 (2013).
20. Baltz, A. G. *et al.* The mRNA-bound proteome and its global occupancy profile on protein-coding transcripts. *Mol. Cell* **46,** 674–690 (2012).
21. Castello, A. *et al.* Insights into RNA biology from an atlas of mammalian mRNA-binding proteins. *Cell* **149,** 1393–1406 (2012).
22. Uhlen, M. *et al.* Tissue-based map of the human proteome. *Science* **347,** 1260419 (2015).
23. Robins, S. P., Duncan, A., Wilson, N. & Evans, B. J. Standardization of pyridinium crosslinks, pyridinoline and deoxypyridinoline, for use as biochemical markers of collagen degradation. *Clin. Chem.* **42,** 1621–1626 (1996).
24. Rodrigues, J. P. *et al.* REF proteins mediate the export of spliced and unspliced mRNAs from the nucleus. *Proc. Natl Acad. Sci. USA* **98,** 1030–1035 (2001).
25. Maris, C., Dominguez, C. & Allain, F. H. The RNA recognition motif, a plastic RNA-binding platform to regulate post-transcriptional gene expression. *FEBS J.* **272,** 2118–2131 (2005).
26. Eswar, N., Eramian, D., Webb, B., Shen, M. Y. & Sali, A. Protein structure modeling with MODELLER. *Methods Mol. Biol.* **426,** 145–159 (2008).
27. Wu, M. *et al.* Structural basis of m(7)GpppG binding to poly(A)-specific ribonuclease. *Structure* **17,** 276–286 (2009).
28. Monecke, T., Schell, S., Dickmanns, A. & Ficner, R. Crystal structure of the RRM domain of poly(A)-specific ribonuclease reveals a novel m(7)G-cap-binding mode. *J. Mol. Biol.* **382,** 827–834 (2008).
29. Nagata, T. *et al.* The RRM domain of poly(A)-specific ribonuclease has a noncanonical binding site for mRNA cap analog recognition. *Nucleic Acids Res.* **36,** 4754–4767 (2008).
30. Hu, G., Gershon, P. D., Hodel, A. E. & Quiocho, F. A. mRNA cap recognition: dominant role of enhanced stacking interactions between methylated bases and protein aromatic side chains. *Proc. Natl Acad. Sci. USA* **96,** 7149–7154 (1999).
31. Keilhauer, E. C., Hein, M. Y. & Mann, M. Accurate protein complex retrieval by affinity enrichment MS rather than affinity purification MS. *Mol. Cell. Proteomics* **14,** 120–135 (2014).
32. Hubner, N. C. *et al.* Quantitative proteomics combined with BAC TransgeneOmics reveals in vivo protein interactions. *J. Cell Biol.* **189,** 739–754 (2010).
33. Poser, I. *et al.* BAC TransgeneOmics: a high-throughput method for exploration of protein function in mammals. *Nat. Methods* **5,** 409–415 (2008).
34. Reed, R. & Cheng, H. TREX, SR proteins and export of mRNA. *Curr. Opin. Cell Biol.* **17,** 269–273 (2005).
35. Carmody, S. R. & Wente, S. R. mRNA nuclear export at a glance. *J. Cell Sci.* **122,** 1933–1937 (2009).
36. Ohno, M., Segref, A., Bachi, A., Wilm, M. & Mattaj, I. W. PHAX, a mediator of U snRNA nuclear export whose activity is regulated by phosphorylation. *Cell* **101,** 187–198 (2000).
37. Hegele, A. *et al.* Dynamic protein-protein interaction wiring of the human spliceosome. *Mol. Cell* **45,** 567–580 (2012).
38. Rehwinkel, J. *et al.* Genome-wide analysis of mRNAs regulated by the THO complex in Drosophila melanogaster. *Nat. Struct. Mol. Biol.* **11,** 558–566 (2004).
39. Pichlmair, A. *et al.* Viral immune modulators perturb the human molecular network by common and unique strategies. *Nature* **487,** 486–490 (2012).
40. Habjan, M. & Pichlmair, A. Cytoplasmic sensing of viral nucleic acids. *Curr. Opin. Virol.* **11,** 31–37 (2015).
41. Jenal, M. *et al.* The poly(A)-binding protein nuclear 1 suppresses alternative cleavage and polyadenylation sites. *Cell* **149,** 538–553 (2012).
42. Ast, G. How did alternative splicing evolve? *Nat. Rev. Genet.* **5,** 773–782 (2004).
43. Wood, V. *et al.* The genome sequence of Schizosaccharomyces pombe. *Nature* **415,** 871–880 (2002).
44. Rufener, S. C. & Muhlemann, O. eIF4E-bound mRNPs are substrates for nonsense-mediated mRNA decay in mammalian cells. *Nat. Struct. Mol. Biol.* **20,** 710–717 (2013).
45. Popp, M. W. & Maquat, L. E. Organizing principles of mammalian nonsense-mediated mRNA decay. *Annu. Rev. Genet.* **47,** 139–165 (2013).
46. Durand, S. & Lykke-Andersen, J. Nonsense-mediated mRNA decay occurs during eIF4F-dependent translation in human cells. *Nat. Struct. Mol. Biol.* **20,** 702–709 (2013).
47. Fackenthal, J. D. & Godley, L. A. Aberrant RNA splicing and its functional consequences in cancer cells. *Dis. Model. Mech.* **1,** 37–42 (2008).
48. Burckstummer, T. *et al.* An orthogonal proteomic-genomic screen identifies AIM2 as a cytoplasmic DNA sensor for the inflammasome. *Nat. Immunol.* **10,** 266–272 (2009).
49. Pichlmair, A. *et al.* IFIT1 is an antiviral protein that recognizes 5'-triphosphate RNA. *Nat. Immunol.* **12,** 624–630 (2011).
50. Cox, J. & Mann, M. MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat. Biotechnol.* **26,** 1367–1372 (2008).
51. Cox, J. *et al.* MaxLFQ allows accurate proteome-wide label-free quantification by delayed normalization and maximal peptide ratio extraction. *Mol. Cell. Proteomics* **13,** 2513–2526 (2014).
52. Huang, da, W., Sherman, B. T. & Lempicki, R. A. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat. Protoc.* **4,** 44–57 (2009).
53. Kim, D. *et al.* TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome. Biol.* **14,** R36 (2013).
54. Trapnell, C. *et al.* Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nat. Biotechnol.* **28,** 511–515 (2010).
55. Wagner, I. *et al.* morFeus: a web-based program to detect remotely conserved orthologs using symmetrical best hits and orthology network scoring. *BMC Bioinformatics* **15,** 263 (2014).
56. Katoh, K. & Standley, D. M. MAFFT: iterative refinement and additional methods. *Methods Mol. Biol.* **1079,** 131–146 (2014).
57. Castresana, J. Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. *Mol. Biol. Evol.* **17,** 540–552 (2000).
58. Guindon, S., Delsuc, F., Dufayard, J. F. & Gascuel, O. Estimating maximum likelihood phylogenies with PhyML. *Methods Mol. Biol.* **537,** 113–137 (2009).
59. Huson, D. H. *et al.* Dendroscope: an interactive viewer for large phylogenetic trees. *BMC Bioinformatics* **8,** 460 (2007).
60. Soding, J. Protein homology detection by HMM-HMM comparison. *Bioinformatics* **21,** 951–960 (2005).
61. Barbier, J. *et al.* Regulation of H-ras splice variant expression by cross talk between the p53 and nonsense-mediated mRNA decay pathways. *Mol. Cell. Biol.* **27,** 7315–7333 (2007).
62. Scholz, J., Besir, H., Strasser, C. & Suppmann, S. A new method to customize protein expression vectors for fast, efficient and background free parallel cloning. *BMC Biotechnol.* **13,** 12 (2013).
63. Provencher, S. W. & Glockner, J. Estimation of globular protein secondary structure from circular dichroism. *Biochemistry* **20,** 33–37 (1981).

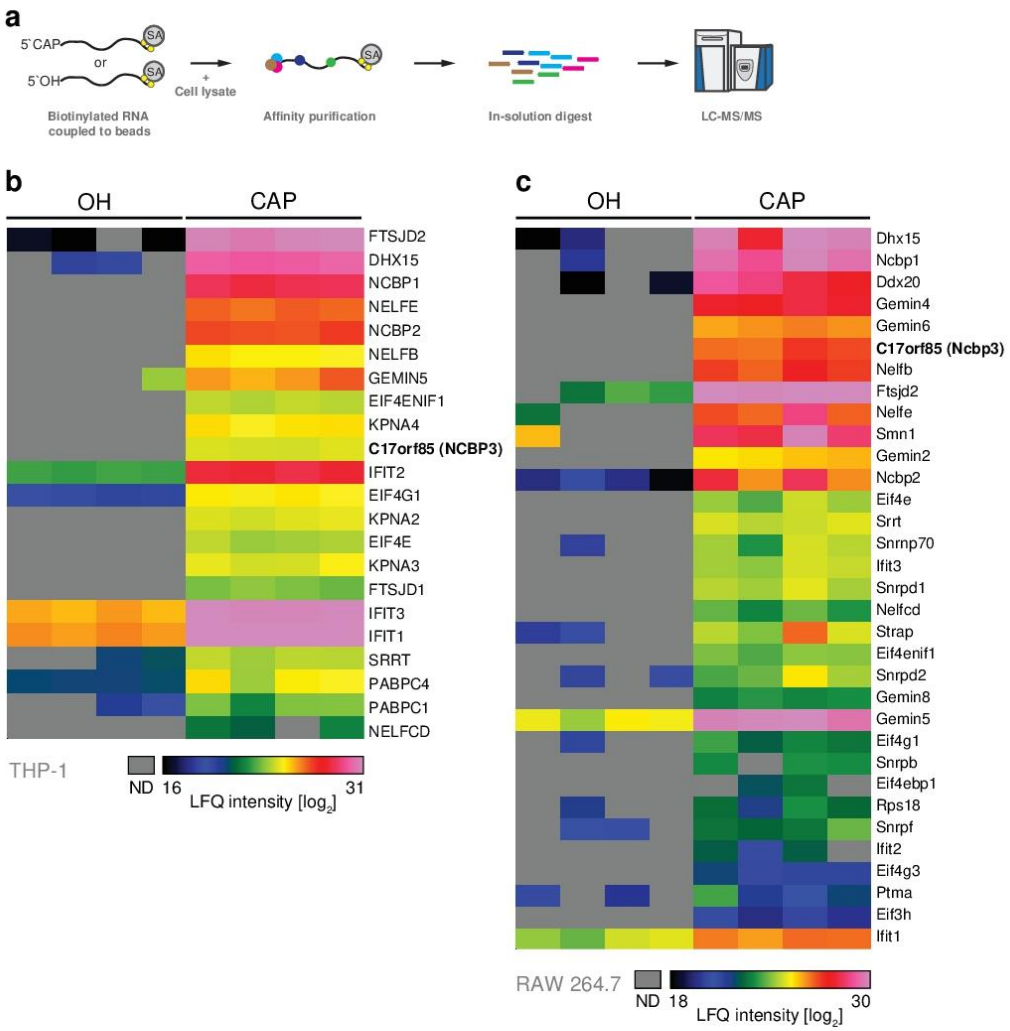
Acknowledgements

We acknowledge the innate immunity laboratory for critical discussions and suggestions, Korbinian Mayr, Igor Paron and Gaby Sowa for maintaining mass spectrometers, Bernd Haas for purifying RNAs, Ralf Zenke for support with image analysis, Ina Poser and Tony Hyman for GFP-expressing cell lines, Ulrich Stelzl for cDNA constructs, the MPI-B core facility for technical assistance and the Max Planck-Genome-Center Cologne for RNA sequencing. We particularly thank Elena Conti, Jernej Ule, Stefan Jentsch and Fulvia Bono for helpful suggestions. This work was supported by the Max-Planck Free Floater program, an ERC starting grant (311339; iVIP) and the german research foundation (P1084/2) to A.P. and an Alexander von Humboldt Fellowship to M.H.

Author contributions

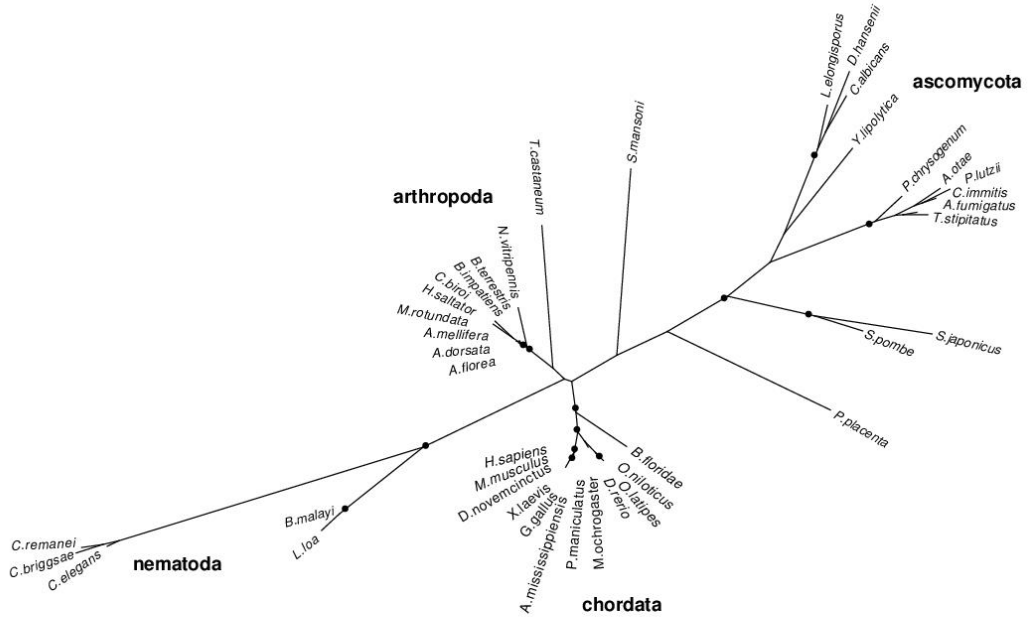
A.G., M.H. and A.P. designed the experiments. A.G., M.H., C.B., D.A.H., An.M. and A.P. performed the experiments. M.H., D.A.H., C.B., Ar.M., M.Y.H. and B.H. performed bioinformatics analysis. M.M. provided critical input. A.G., M.H. and A.P. wrote the manuscript.

Supplementary Figure 1



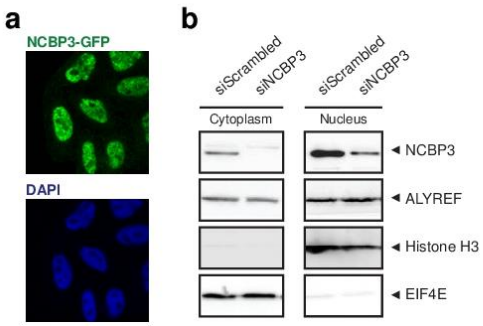
Supplementary Figure 1. Heat map of proteins enriched by capped RNA.
(a) Schematic of quantitative interaction proteomics approach to identify proteins associated with capped RNA. Synthetic biotinylated RNA oligos harbouring either a 5'm^7 guanine cap structure (CAP) or hydroxyl group (OH) were coupled to streptavidin beads, incubated with cell lysates, and bound proteins identified by shotgun liquid chromatography-tandem MS (LC-MS/MS). **(b, c)** Heat maps showing hierarchical clustering (Euclidean distances) of proteins significantly enriched by synthetic cap-RNA oligos (compared to OH-oligos) in murine RAW 264.7 **(b)** and human THP-1 macrophages **(c)**. The plots show non-imputed log2 transformed LFQ intensities for each individual replicate in rainbow colours (see colour scale). Grey colour denotes missing values (ND).

Supplementary Figure 2



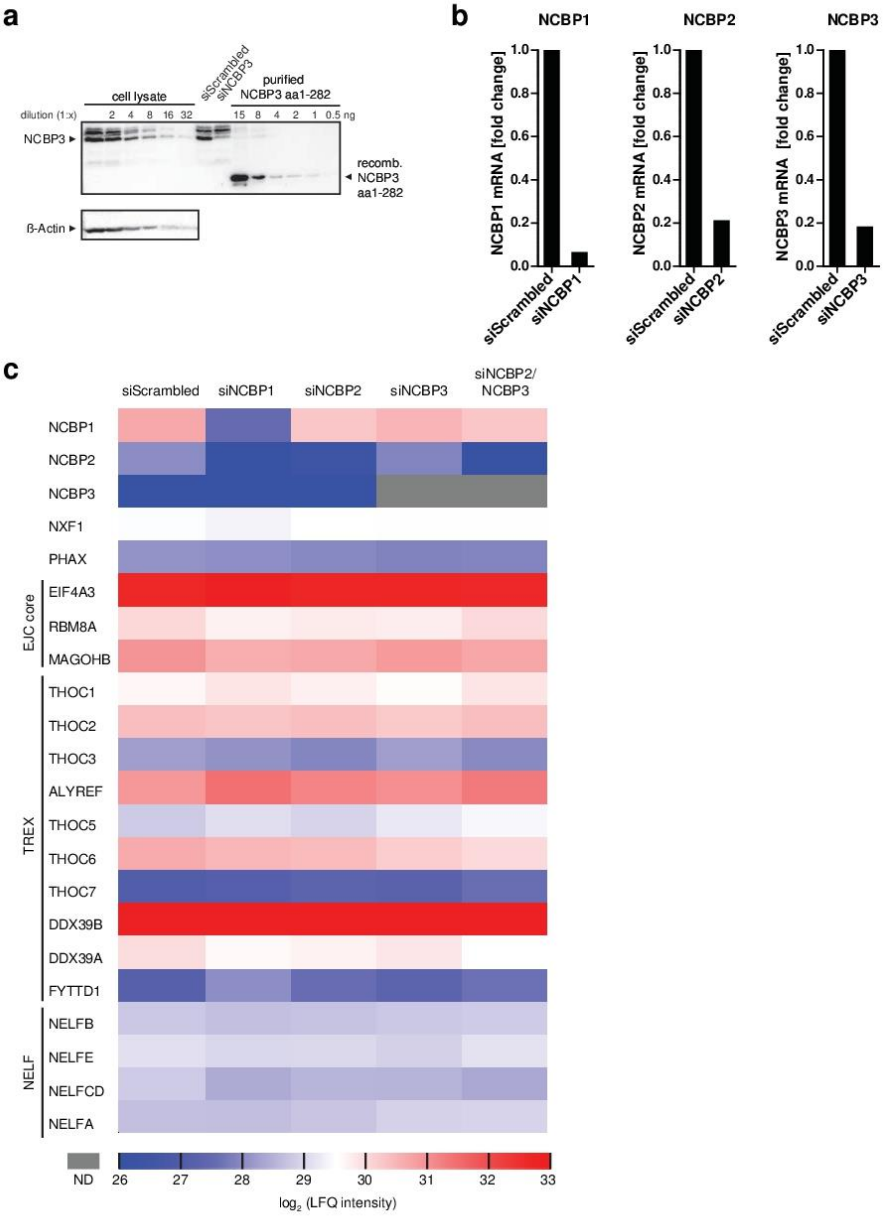
Supplementary Figure 2. NCBP3 is an evolutionary highly conserved protein. Phylogenetic tree of the NCBP3 family of proteins. NCBP3 can be found in Chordata, Arthropoda, Nematoda and Fungi. It is, however, not found in the model organisms Drosophila melanogaster (or any sequenced Muscomorpha) and Saccharomyces cerevisiae (or any sequenced Saccharomycetaceae). Note that we also detected similarity to proteins from plants (green and red algae), which were omitted from further analysis. Dots indicate stable tree branches according to Phyml bootstrapping. Species abbreviations and accession numbers are shown in Table S2.

Supplementary Figure 3



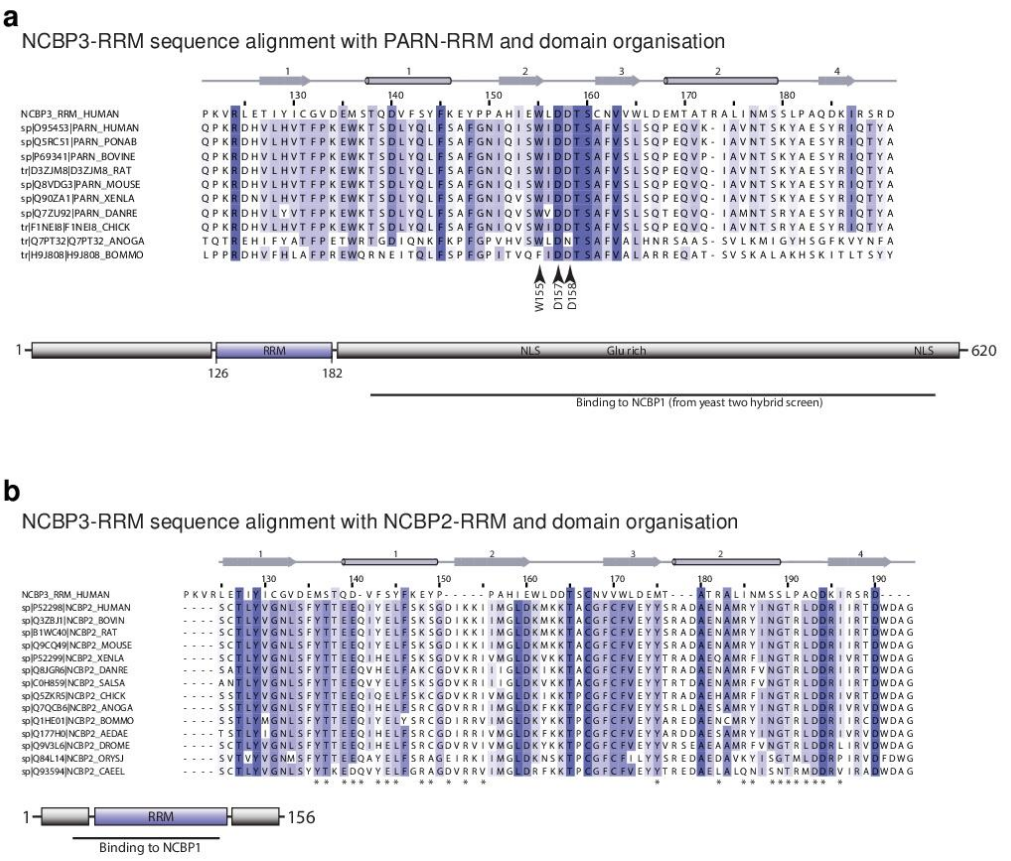
Supplementary Figure 3. Subcellular distribution of NCBP3.
(a) Immunofluorescence analysis to determine the cellular localization of NCBP3. HeLa cells stably expressing NCBP3-GFP were stained with antibodies against GFP (green) and DAPI (blue) and analysed by confocal microscopy. **(b)** Subcellular fractionation of RNAi-treated HeLa cells. Cells were treated with siRNA against NCBP3 or siScrambled as control, and cytoplasmic (C) and nuclear (N) fractions were analysed by western blotting using antibodies against endogenous NCBP3 and ALYREF. Histone H3 and EIF4E were used as controls.

Supplementary Figure 4



Supplementary Figure 4. Knockdown efficiency of NCBP1, -2 and -3.
(a) Western blot for knockdown efficiency of NCBP3. HeLa cells were treated with siRNAs targeting NCBP3 or control for a total of five days and abundance of NCBP3 tested by western blotting. Serial 2-fold dilutions of HeLa cell lysates and dilutions of purified NCBP3-RRM (1-282) protein were used to assess the sensitivity of the antibody. **(b)** HeLa cells were treated with siRNA against endogenous NCBP1, NCBP2 or NCBP3. To validate knockdown efficiencies. NCBP1, NCBP2 and NCBP3 mRNA levels were determined by qRT-PCR two (siNCBP1) or five days (siNCBP2, siNCBP3) after knockdown. Values were normalized to GAPDH mRNA levels and are expressed as fold change to siScrambled control. **(c)** Proteome analysis of HeLa cells after NCBP knockdowns. HeLa cells were treated with siScrambled control, NCBP1, NCBP2, NCBP3 or NCBP2 and NCBP3 targeting siRNAs for 5 days and analysed by LC-MS/MS. Heat map showing non-imputed log2 transformed averaged LFQ intensities from three replicates for the indicated proteins. Red colour (see scale) represents maximum intensity, blue – minimum. Grey colour indicates

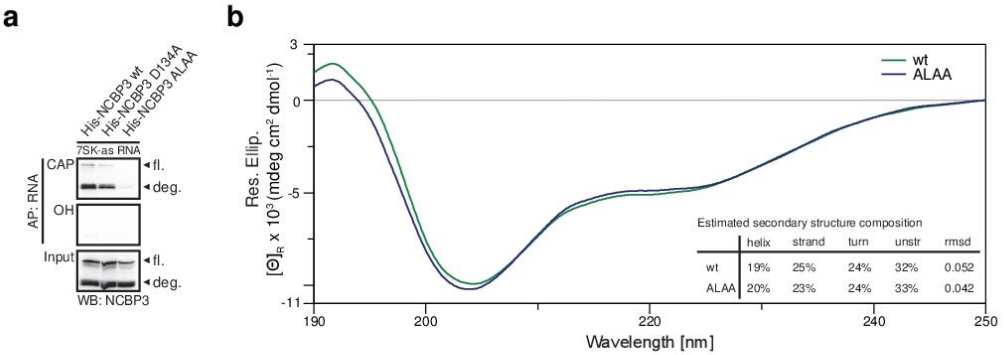
Supplementary Figure 5



Supplementary Figure 5. NCBP3-RRM amino acid sequence alignment and domain organisation.

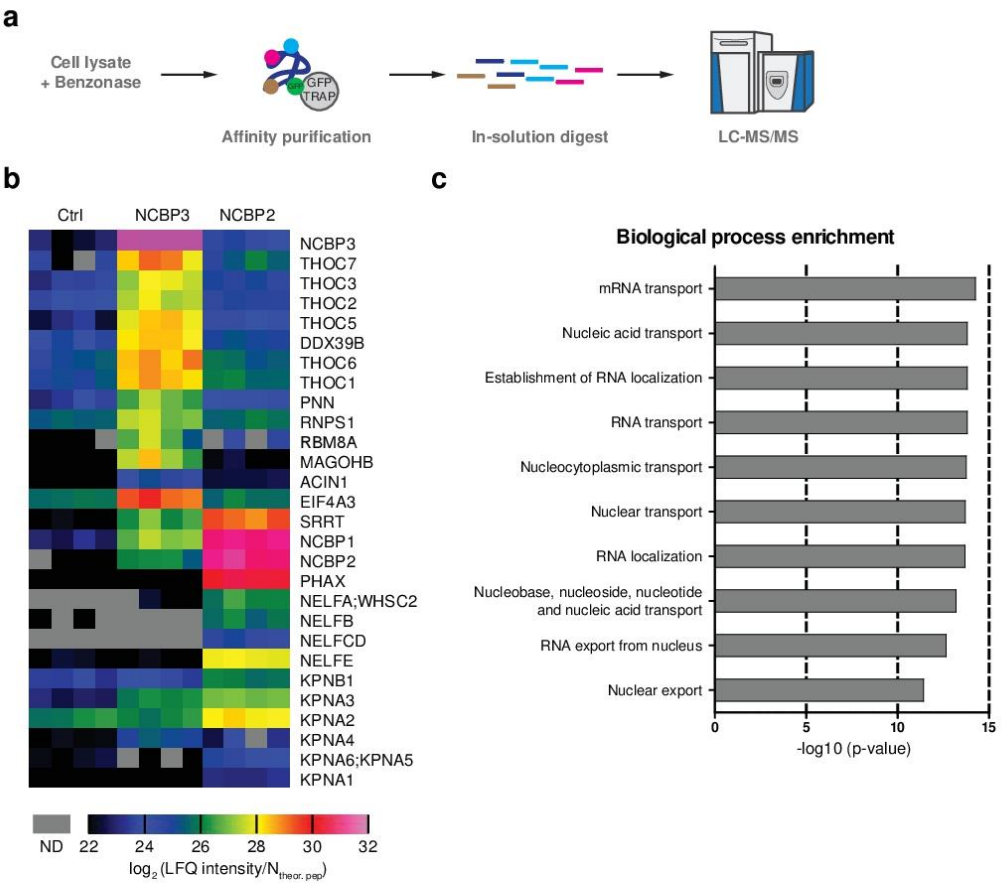
Shown are representative sequence alignments (Clustal Omega)[1] between the RRMs of human NCBP3 and Poly(A)-specific ribonuclease (PARN) **(a)**, and between the RRMs of NCBP3 and NCBP2 **(b)**, as well as the respective domain organisations (bar diagrams). Regions that were shown to be necessary for binding to NCBP1 are marked[1, 2]. Sequence numbering refers to full-length human NCBP3. Secondary structure elements are shown above (NCBP3, predicted) and below (NCBP2, PDB code 1h6k10) the alignments. Asterisks in **(b)** mark residues of NCBP2 engaged in the interface formation with NCBP1.

Supplementary Figure 6



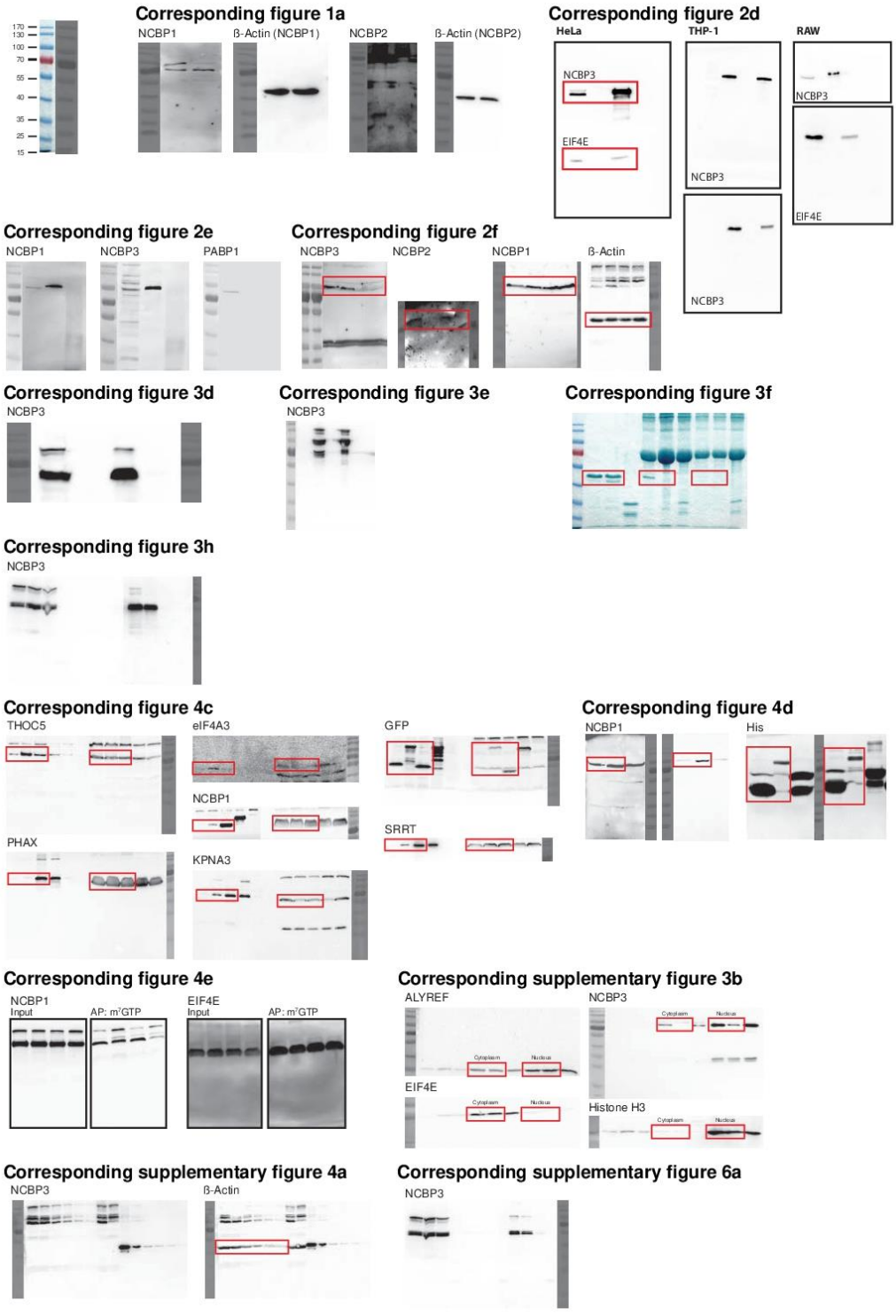
Supplementary Figure 6. Characterization of NCBP3 cap-binding using full-length protein and the RRM-domain.
(a) Binding of recombinant wild-type and mutant NCBP3 to 5'capped RNA. Western blot after affinity purification with biotinylated in-vitro transcribed 7SK-as RNA harbouring either a 5'CAP or OH structure using lysate from E.coli expressing either recombinant full-length wild-type NCBP3 (wt), NCBP3 where aspartic acid at position 134 had been mutated to alanine (D134A), or where tryptophan at position 155 and two aspartic acids at position 157 and 158 had been mutated to alanines (ALAA). Full-length (fl.) recombinant NCBP3 and N-terminal degradation products (deg.) are indicated. **(b)** Far UV circular dichroism (CD) spectra for wild type (wt; green) and mutant (ALAA; blue) NCBP3 (aa 1-282). The overall similarity of the spectra as well as the estimated secondary structure compositions (insert table) are shown.

Supplementary Figure 7



Supplementary Figure 7. Heat map and functional annotation analysis of proteins enriched by NCBP3 and NCBP2.
(a) Schematic of quantitative interaction proteomics approach to identify proteins associated with NCBP3 or NCBP2. GFP-Trap beads were incubated with benzonase treated cell lysates and bound proteins identified by shotgun liquid chromatography-tandem MS (LC-MS/MS). **(b)** Heat map showing hierarchical clustering (Euclidean distances) of proteins significantly enriched by precipitation of GFP-NCBP3 and GFP-NCBP2 as compared to GFP-RAB5C (ctrl) belonging to the TREX complex, NELF complex, EJC, Importins or CBC. The plot shows non-imputed log2 transformed LFQ intensities for each individual replicate in rainbow colours (see colour scale). Grey colour denotes missing values (ND). **(c)** Functional annotation analysis of enriched gene ontology biological processes (GOBP). The gene list of enriched NCBP3 interactors (p-value < 0.001 and LFQ intensity fold change ≥ 8 as compared to GFP-RAB5C) has been analysed for GOBP enrichment using DAVID. Plotted are the top ten GOBP with the lowest p-values.

Supplementary Figure 8



Supplementary Figure 8. Original western blots.

Supplementary References

1. McWilliam, H. *et al.* Analysis Tool Web Services from the EMBL-EBI. *Nucleic acids research* **41**, W597-600 (2013).

2. Mazza, C., Ohno, M., Segref, A., Mattaj, I. W. & Cusack, S. Crystal Structure of the Human Nuclear Cap Binding Complex. *Mol. Cell* **8**, 383–396 (2001).

3. Hegele, A. *et al.* Dynamic protein-protein interaction wiring of the human spliceosome. *Mol. Cell* **45**, 567–80 (2012).

2.2 PUBLICATION 2: The alternative cap-binding complex is required for antiviral response *in vivo*

Gebhardt, A., Schnepf, D., Moser, M., Meiler, A., Michaudel, C., Mackowiak, C., Sedda, D., Stukalov, A., Reinert, L., Paludan, S.R., Ryffel, B., Stäheli, P., and Pichlmair, A. (2018). The alternative cap-binding complex is required for antiviral response *in vivo*. (prepared for submission)

The manuscript "The alternative cap-binding complex is required for antiviral response *in vivo*." investigates the requirement of the alternative CBC during viral infections. *In vitro* experiments showed that the alternative CBC subunit NCBP3 is required to regulate appropriate expression of innate immune response genes and is involved in restriction of viral infections. I established a NCBP3 knockout mouse model and could show that NCBP3 is required to pertain influenza A virus mediated pathology *in vivo*. Combining *in vitro* and *in vivo* data highlighted that the alternative CBC is fundamental to combat viral infections. NCBP3 appears to be involved in the precise regulation of antiviral gene expression.

The manuscript prepared for submission is presented below and includes results, discussion, material and methods and supplementary information. Supplementary table can be provided on request.

The alternative cap-binding complex is required for antiviral response *in vivo*

Anna Gebhardt[1], Daniel Schnepf[2,3], Markus Moser[4], Arno Meiler[1], Chloé Michaudel[5], Claire Mackowiak[5], Delphine Sedda[5], Alexey Stukalov[1], Line Reinert[6], Søren R Paludan[6], Bernhard Ryffel[5,7], Peter Stäheli[2,8] and Andreas Pichlmair[1,9,10]

[1] Innate Immunity Laboratory, Max-Planck Institute of Biochemistry, Martinsried, Munich, Germany. [2] Institute of Virology, University of Freiburg, Freiburg, Germany. [3] Spemann Graduate School of Biology and Medicine, Albert Ludwigs University Freiburg, Freiburg, Germany. [4] Department of Molecular Medicine, Max-Planck Institute of Biochemistry, Martinsried, Munich, Germany. [5] INEM, Experimental Molecular Immunology, UMR7355 CNRS and University, Orleans, France. [6] Department of Biomedicine, University of Aarhus, Aarhus, Denmark. [7] Institute of Infectious Diseases and Molecular Medicine, University of Cape Town, Cape Town, South Africa. [8] Faculty of Medicine, University of Freiburg, Freiburg, Germany. [9] School of Medicine, Institute of Virology, Technical University of Munich, Munich, Germany. [10] German Center for Infection Research (DZIF), Munich partner site, Munich, Germany.

Corresponding author:
Prof. Dr. Andreas Pichlmair, PhD
Innate Immunity laboratory
Max-Planck Institute for Biochemistry
Am Klopferspitz 18
D-82152 Martinsried, Germany
Email: apichl@biochem.mpg.de
Telephone: +49 89 8578 2220

Abstract

Cellular adaption to environmental changes requires immediate and precise regulation of transcriptional programs. During viral infections, this includes the expression of innate immune regulated genes like type I interferons and pro-inflammatory cytokines as well as interferon stimulated genes (ISGs) that are needed to successfully control viruses. The translation into functional proteins depends on the transcriptional activation of specific genes and precise export of respective mRNAs to the cytoplasm. Until recently, it was believed that all polymerase II transcribed mRNAs are protected and guarded to the cytoplasm bound to the canonical cap-binding complex (CBC). We recently identified an alternative CBC consisting of Ncbp1 and -3 with redundant function under physiological condition. Here, we provide evidence that the alternative CBC is essential to mount precise and appropriate antiviral programs. In Ncbp3 deficient cells, infection with Influenza A virus and Vesicular stomatitis virus resulted in higher levels of viral RNA and protein as well as viral titers. Notably, regardless of heightened viral stimuli, the expression of ISGs was comparable between Ncbp3 deficient and wild-type cells. Stimulation with low concentration of synthetic polyinosinic-polycytidylic acid or interferon led to reduced IL6 secretion or Ifit1 expression in Ncbp3 deficient cells, respectively, suggesting a defect in innate immune responses. In support, proteomics analysis of Influenza A virus infected cells revealed impaired expression of a subset of ISGs in Ncbp3 deficient cells. Importantly, during Influenza A virus infection *in vivo* loss of Ncbp3 resulted in severe lung pathology leading to heightened morbidity. In sum, we show that the alternative CBC is essential for viral defense by regulating the expression of antiviral genes in a fine-tuned and precise manner.

Introduction

Successful control of virus infections requires immediate and appropriate responses to pathogens, a process centrally mastered by the innate immune system. The activation of this system is predominantly initiated by viral nucleic acids that are delivered during the infection process and bear virus-specific chemical or structural modifications [1]. Pathogen-associated molecular patterns (PAMPs) are sensed by specific pattern recognition receptors (PRRs) which activate signaling cascades leading to transcription of antiviral defense genes. This includes antiviral and inflammatory cytokines such as type-I interferons (IFN-α/β), Tumour necrosis factor alpha (TNF-α) and interleukins (IL) including IL6, IL8 and IL12 [2]. Cytokines act in a paracrine and autocrine manner and bind to specific receptors, which trigger signaling cascades leading to transcriptional activation of antiviral and inflammatory effector proteins. The effector proteins include several hundred interferon-stimulated genes (ISGs) such as Interferon-induced proteins with tetratricopeptide repeats (IFITs), orthomyxovirus resistance genes (MX) or 2'-5'-oligoadenylate synthases (OAS) [3–5]. Transcripts are exported from the nucleus and translated into bioactive proteins by the cytoplasmic translation machinery. The cellular apparatus of transcription, mRNA processing/export and translation is required to effectively control virus infection and to mount an appropriate immune response.

Newly transcribed RNA is one of the key components of the innate immune response. To protect cellular RNA transcripts and to coordinate the steps from transcription to translation, RNA polymerase II transcripts like messenger RNAs (mRNAs) are co-transcriptionally capped on the RNA 5`end by a mono-methylated guanosine which is linked through a 5`5`-phosphodiester bond (cap) [6]. The RNA cap is bound by the cap-binding complex (CBC), which coordinates RNA processing, export and initiation of translation. The CBC is a heterodimeric complex consisting of

nuclear cap binding protein (NCBP) 1 and -2. While NCBP2 is a cap-binding protein, NCBP1 allows recruitment of cellular factors that are required for RNA processing [6–8]. All processing steps such as splicing and 3`end polyadenylation are highly controlled by multiple protein complexes and it was shown that accurate processing advantages export of mRNAs to the cytoplasm [9]. One of these steps is the excision of intronic regions by the spliceosome and loading of the exon-junction-complex (EJC) on spliced RNA. This allows association to the export receptor NXF1-NXT1 (also known as TAP-p15) and facilitates the translocation of the mature mRNA through the nuclear pore complex (NPC) [10].

We recently identified an alternative CBC consisting of the cap-binding protein NCBP3 and NCBP1 [11]. Under physiological conditions the canonical and alternative CBCs carry out similar functions and are to large extent functionally redundant to process and export mRNAs. Only co-depletion of the two cap binding proteins NCBP2 and -3 retains polyadenylated RNA in the nucleus, while depletion of the common adapter NCBP1 has the same effect. However, why eukaryotes evolved two CBCs remained unclear. We speculated that the alternative CBC may be particularly important in situations of cellular stress such as during virus infections.

Viruses have identified mRNA processing and splicing as a primary battleground to down-regulate immune responses and to support viral replication and spread. Influenza A virus (IAV), for instance, encodes for the nonstructural protein 1 (NS1) to counteract and downregulate host innate immune responses, in part through targeting multiple steps in the mRNA processing and nuclear export process [12]: IAV-NS1 (1) blocks RNA 3`end processing through inhibiting components of the cellular and polyadenylation specificity factor (CPSF) complex, (2) binds directly to PABPN and thereby pertains proper mRNA polyadenylation and transport to the cytoplasm and (3) associates to the export receptor NXF1-NXT1 and the nuclear pore protein Nup98. Another

example for viral manipulation of the mRNA export machinery is the ability of the Vesicular stomatitis virus (VSV) matrix protein (M), which interacts with the RNA binding protein Rae1 and prevents its interaction with nucleoporin Nup98, leading to accumulation of mRNA and U snRNAs in the nucleus [13]. Impaired mRNA export and processing inhibit appropriate immune responses and therefore paralyze cells in a vulnerable state. Conversely, mutations in IAV-NS1 or VSV-M proteins lead to attenuation of these virus strains and even allows their usage as safe therapeutic agents [14–19].

Although the recently identified alternative CBC consisting of NCBP1/-3 functioned in a redundant-manner under physiological conditions, we could show that NCBP3 was of critical importance to impair virus growth *in vitro*. However, the functional basis for this phenomenon has remained unclear. Furthermore, it was not clear to what extent this finding was relevant for the immune response in a multi-cellular organism. Here, we show that Ncbp3 is important to mount a proper immune response during virus infections *in vitro*. Genetic depletion of Ncbp3 *in vivo* was rendering mice highly susceptible to IAV infection, indicating a critical role of Ncbp3 during antiviral responses *in vivo*.

Results

Characterization and generation of Ncbp3 knockout mice

To test the function of Ncbp3 *in vivo* we generated Ncbp3 knockout (ko) mice by homologous recombination (Fig Supplementary Fig1a). Correct insertion of the knockout cassette ($Ncbp3^{tm1a}$) was confirmed by specific PCR assays and resulted in 49% of heterozygous F0 mice (Supplementary Fig 1b, c). Heterozygous Ncbp3 mice ($Ncbp3^{+/-}$) showed no phenotype as compared to littermate controls. However, when breeding Ncbp3 heterozygous mice we obtained reduced numbers of Ncbp3 deficient mice (Fig 1a, Supplementary Fig 1d). Similar reduced numbers of homozygous offsprings were evident when breeding Ncbp3 knockout to heterozygous animals (Fig 1a). Born Ncbp3 ko mice showed reduced body weight compared to littermate controls but no other obvious phenotype during an observation period of 14 month (Supplementary Fig 1e, f). Tissues of 12-13 weeks old mice were analyzed for Ncbp3 expression investigating the LacZ expression which is expressed under the control of the endogenous Ncbp3 promoter (Supplementary Fig. 1g). Ncbp3 expression could be classified in three expression patterns: low (bone marrow, pancreas, spleen, thymus, kidney, sexual organs and liver), intermediate (lung and brain) and high (heart and muscle) expression with slight differences between male and female animals.

Ncbp3 wild-type (wt) and ko mouse embryonic fibroblasts (MEF) isolated from heterozygous crossings were morphologically similar. Loss of Ncbp3 was confirmed by western blot analysis employing Ncbp3 specific antibody (Fig 1b). Heterozygous MEFs showed intermediate expression of Ncbp3. We proceeded to use these cells to test the function of the alternative CBC *in vitro*. To test whether the canonical CBC is operative in murine cells, we employed a transient knockdown experiment depleting Ncbp1 and Ncbp2 by siRNA. Knockdown efficiency was

confirmed by western blot analysis (Fig 1c). In line with results in human cells, Ncbp2 depletion did not affect growth of wt MEFs (Fig 1d). In contrast, knockdown of Ncbp2 in Ncbp3 ko MEFs significantly reduced cell growth. This effect was comparable to depletion of Ncbp1. In summary, these experiments showed that the canonical CBC (consisting of NCBP1/2) and the alternative CBC (consisting of NCBP1/3) are evolutionary conserved in murine and human cells, that both CBCs are operative in MEFs and that they serve redundant functions under physiological conditions.

Loss of Ncbp3 supports RNA virus replication *in vitro*

A phenotype of Ncbp3 depletion was previously shown to be predominantly apparent under challenging environmental conditions [11]. We used different viruses to assess the function of Ncbp3 during virus infections. To this aim we selected viruses that encode their genome as RNA (influenza A virus: IAV-wt, vesicular stomatitis virus: VSV-wt) and replicate in the nucleus (IAV) or cytoplasm (VSV). In addition, we used viral mutants that lost their ability to perturb the innate immune system (IAV-ΔNS1, VSV-M2). Infection of Ncbp3 wt and ko MEFs with different doses of an IAV reporter virus that expresses renilla luciferase showed that Ncbp3 ko MEFs allowed higher expression of IAV driven renilla signals as compared to wt MEFs (Fig 2a). These results were confirmed by testing IAV accumulation in the supernatant of MEFs. IAV accumulation was 30-fold higher in Ncbp3 ko as compared to wt MEFs (Fig 2b). Likewise, the interferon inducing variant IAV-ΔNS1 accumulated to more than 10-fold higher virus titers in Ncbp3 deficient as compared to control cells (Fig 2b). Similar results were obtained when assessing mRNA accumulation of virus transcripts in infected cells. IAV RNA accumulated to 10-fold higher levels in NCBP3 ko cells as compared to wt controls (Fig 2c). Higher accumulation of virus mRNA was corroborated with higher mRNA levels for IFN-β in Ncbp3 ko cells as compared to wt controls.

Surprisingly, despite higher viral and IFN-β mRNA expression in Ncbp3 deficient MEFs, accumulation of interferon responsive Ifit3 mRNA was undistinguishable in both cell types (Fig 2c). RNA of the housekeeping gene Hmbs were similar expressed upon infection and in Ncbp3 ko compared to wt cells.

We set out to confirm these data by western blotting. In line with a role in virus growth regulation, accumulation of the viral nucleoprotein (NP) and non-structural protein 1 (NS1) were increased in Ncbp3 ko MEFs as compared to wt counterparts (Fig 2d). Again, despite the higher accumulation of viral proteins, expression of Ifit1 in response to virus infection was similar in both cell types (Fig 2d). In sum, these experiments suggested that Ncbp3 is required to control IAV infection *in vitro*. However, despite significantly higher accumulation of viral nucleic acids in Ncbp3 ko cells, the induction of the interferon response was equal in both, Ncbp3 wt and ko MEFs.

Since IAV replicates in the nucleus of infected cells, the increased virus replication could also be interpreted by an effect of Ncbp3 on viral RNA export. We therefore assessed growth of VSV, a virus that replicates in the cytoplasm. In line with data obtained for IAV, VSV-driven luciferase was increased in Ncbp3 ko MEFs as compared to wt controls (Fig 3a) and accumulation of viruses in supernatants of wt and Ncbp3 deficient MEFs was 100-fold increased for VSV-wt and the interferon inducing variant VSV-M2 (Fig 3b). Notably, despite this significantly increased virus growth, expression of Ifit1 protein was again similar in Ncbp3 wt and ko MEFs (Fig 3c). These results indicate that Ncbp3 deficient MEFs do not mount an innate immune response that is quantitatively equalling the amount of viruses present in infected cells. Altogether, these data show that infection with RNA viruses, regardless of their replication site, results in higher accumulation of viral particles in supernatant of Ncbp3 depleted cells. Activation of the innate immune responses

is similar in Ncbp3 wt and ko MEFs but execution of the innate immune response was severely impaired in Ncbp3 depleted cells.

Global proteomic analysis suggests altered innate immune response of a subset of ISGs in Ncbp3 deficient MEFs

Our data indicate that Ncbp3 wt and ko MEFs differ in their ability to translate innate immune signals into functional proteins. To investigate this on a proteome-wide scale we performed unbiased proteomic analysis of IAV infected MEFs. Briefly, MEFs were infected with IAV-wt and proteomic analysis were performed by liquid chromatography tandem mass spectrometry (LC-MS/MS) 24 hours after infection (Fig 4a). This analysis allowed quantification of 5422 proteins in Ncbp3 wt and ko MEFs (Supplement Table 1). In line with increased virus growth, we could observe increased expression of viral proteins in Ncbp3 deficient as compared to wt MEFs (Fig 4b). Investigating the differences in protein expression between Ncbp3 ko and wt MEFs upon virus infection, we analyzed the fold change difference of Ncbp3 IAV infection versus mock in the two cell lines (Fig 4c, Supplementary Table 1). In this analysis, 282 proteins were differentially expressed in Ncbp3 ko as compared to wt cells, including 19 differentially expressed ISGs. Among those ISGs, four differentially expressed classes could be observed (Fig 4c, green dots, Fig 4d): (1) ISGs that were up-regulated in wt cells after virus infection and less up-regulated in Ncbp3 ko: Bst2, Irgm2 and Rtp4; (2) ISGs that were up-regulated in wt MEFs upon virus infection and not regulated in Ncbp3 ko: Ctnna2, H2-K1, Ifi44, Irgm1, Lgals3bp, Timp2, Ube2l6 and Usp18; (3) ISGs that were up-regulated in wt cells and slightly down regulated in Ncbp3 ko cells: Helz2 and Tspo; and (4) ISGs that were not regulated in wt cells and slightly down-regulated in Ncbp3 ko cells: Lipa, Gla, Crabp2 and Tor1aip2. In addition, Noc4l was found to be significantly regulated between Ncbp3 ko and wt being marginally upregulated in Ncbp3 ko upon IAV infection. These

analyses show that a subset of ISGs is differentially expressed in the absence of Ncbp3 potentially required to control virus infection in a fine-tuned manner.

Ncbp3 is required for cytokine induction and response *in vitro*

Viral infection experiments suggested regulation of innate response genes in an Ncbp3-dependent manner. These results prompted us to test MEFs for their ability to respond to the synthetic dsRNA analogue and innate immune stimulus polyinosinic-polycytidylic acid (PIC). Intriguingly, compared to wt MEFs, PIC-treated Ncbp3 deficient cells showed significantly reduced accumulation of IL6 in their supernatants (Fig 5a). This difference was particularly apparent when low amounts of PIC were used (Fig 5b), clearly suggesting that proper antiviral responses require a functional alternative CBC. To test whether Ncbp3 is also required to properly execute an antiviral response we stimulated MEFs with recombinant IFN-α/β and tested for the ability to synthesize the antiviral protein Ifit1 by western blot analyses. In the absence of Ncbp3, interferon elicited Ifit1 levels were clearly reduced as compared to Ifit1 expression in wt MEFs (Fig 5c). This difference was again particularly evident at low cytokine concentrations and less prominent when high amounts of IFN-α/β were used for stimulation of cells. In sum, these data indicate that Ncbp3 is required for fully functional antiviral immune response *in vitro*.

Ncbp3 depletion results in increased IAV induced mortality *in vivo*

Lack of Ncbp3 leads to increased virus growth, altered cytokine response and reduced induction of a subset of ISGs *in vitro*. However, the role of the Ncbp3 *in vivo* is not known. To test the relevance of Ncbp3 *in vivo* we infected Ncbp3 ko and littermate control mice intranasally with a sublethal dose of IAV. As expected, control mice survived virus challenge. However, IAV infection resulted in severe disease of Ncbp3 deficient mice of which 66.7% died within 13 days of infection (Fig 6a). In this infection regime, mice died 7-8 days after infection suggesting

224 inability to properly control virus infection and/or triggering of detrimental pathological effects.
225 In order to test the latter possibility we performed histological analyses of lungs from infected
226 mice (Fig 6b, c). Compared to lungs isolated from infected wt mice, lungs from Ncbp3 ko mice
227 augmented lung inflammatory cell recruitment and pathology characterized by severe epithelial
228 damage and infiltration of lymphocytes, hallmarks of necrotic peribronchial pneumonia caused by
229 influenza virus infection. Collectively, these experiments show that Ncbp3 is required to control
230 virus growth and adequate innate responses *in vitro* and that loss of Ncbp3 results in severe
231 virus-induced inflammation and pathology *in vivo*.

Discussion

Regulation of gene expression is one of the most fundamental process in a cellular organism. To obtain functional proteins, the genetic information need to be transcribed into RNA and shuttle between the two main cellular compartments, the nucleus and the cytoplasm. Coordination of this event is fundamental for precise gene expression in a timely manner. Under challenging conditions such as virus infections, proper functionality of mRNA processing machineries becomes even more important since an immediate and adequate response needs to be mounted. We previously identified human NCBP3 as a cap-binding protein assembling the alternative CBC together with NCBP1 and being involved in the export of polyadenylated RNA from the nucleus in a similar manner as the canonical CBC under physiological conditions [11]. Here, we show that Ncbp3 deficient mice are viable, which supports the *in vitro* finding that Ncbp3 is dispensable under physiological conditions since its exhibits redundant functions with Ncbp2. Surprisingly, breeding of Ncbp3 heterozygous or knockout mice revealed an atypical Mendelian ratio of born knockout animals suggesting that Ncbp3 is required for proper embryonic development, which may reflect highly complex and regulated gene transcription programs that are particularly active during embryonal development [20].

The alternative CBC may be predominantly required to respond to environmental cues and to allow appropriate gene expression [11]. Such responses are particularly required to coordinate antiviral responses triggered by virus infections [21,22]. *In vitro*, infections with RNA viruses resulted in elevated viral titers, increased viral protein production as well as in higher viral transcript expression in Ncbp3 deficient cells. Since this was observed for both, nuclear and cytoplasmic replicating viruses, the effect of Ncbp3 appears to rather be of regulatory nature of antiviral immune responses than a direct effect of Ncbp3 on viral RNA export and growth. Given the high

abundance of viral transcripts in Ncbp3 deficient cells, expression of ISGs monitored by Ifit3 mRNA and Ifit1 protein did not proportionally follow the high amount of stimulus. This indicates dysregulated immune responses in Ncbp3 deficient cells and a general inability of Ncbp3 ko cells to translate a stimulatory trigger into protective antiviral immunity. Proteomic analyses of virally infected Ncbp3 ko and wt MEFs showed reduced expression of a subset of ISGs in Ncbp3 deficient cells, including antiviral restriction factors as well as negative regulators of type I interferon, additionally supporting the requirement of Ncbp3 for proper antiviral responses. Consolidating our data, Ncbp3 was necessary for proper expression of antiviral proteins such as IFIT proteins and Bst2 (also known as Tetherin), which are antiviral factors active against a wide variety of viruses [3,23–25]. While IFIT proteins are nucleic acid binding proteins that restrict virus replication, Bst2 impairs the release of enveloped viruses including IAV and VSV [3,25–27]. In line with elevated accumulation of viral particles in the supernatants of Ncbp3 ko compared to wt cells, the missing induction of IFIT proteins and Bst2 and their functions in restricting virus replication and virion release could lead to the higher accumulation of viral particles. Moreover, Bst2 and Timp2 can inhibit metalloproteinases leading to reduced cell growth and migration [28–31] and their reduced expression in Ncbp3 ko cells could result in enhanced cell growth and migration, which is beneficial for viral replication. Similarly, Ncbp3 ko MEFs showed reduced expression of Usp18 and Ube2l6, which were described to be involved in ISGylation induced by innate responses. During IAV infection, NS1 is ISGylated which leads to the inhibiting of viral replication [3,32,33]. Alteration in ISGylation in the absence of Ncbp3 could therefore further contribute to enhanced viral replication as seen by *in vitro* experiments. In sum, our results indicate defects in the induction of an appropriate antiviral immune response in Ncbp3 ko cells *in vitro*, which may accumulate and result in the observed higher viral titers.

An involvement of Ncbp3 in induction of antiviral immune responses was further supported by stimulation experiments using the synthetic dsRNA analogue PIC and recombinant IFN-α. Ncbp3 deficient cells showed reduced cytokine and ISG expression, particularly when stimulated with low concentrations of ligands. However, at higher concentrations of stimuli, the dependency on Ncbp3 gradually decreased. This indicates that high amounts of mRNA may overcome dependency on the alternative CBC and that regulation of mRNA export depends on the stoichiometry of individual mRNAs. In case of viral infections, Ncbp3 may therefore be predominantly important early after infection when little virus stimulus initiates an antiviral response, which requires full activity of antiviral programs in order to control initial infections. Similarly, expression of negative regulators of the innate immune system may be delayed in case of NCBP3 deficient systems, thereby leading to potential overshooting of immune responses at later stages of infection. However, the exact timing of antiviral responses is of central importance to control virus infections and can dramatically affect the outcome of disease [34]. A fully functional mRNA processing system may be required to allow such a timed response.

In an organism, the antiviral program needs to be strictly coordinated to effectively prevent viral spread and associated pathologies [35,36]. While timely expression of cytokines and antiviral proteins is crucial at early stages of infection, expression of negative regulators is essential to prevent overshooting immune reactions after successful virus clearance [21,37–39]. This highly dynamic and regulated expression may require the alternative CBC. Indeed, mice lacking Ncbp3 showed increased virus induced mortality and lung pathology. Ncbp3 deficiency may results in the inability of the organism to negatively regulate the immune response, which has been demonstrated to be important to prevent overshooting and prolonged immune reactions in response to viruses and cytokines *in vivo* [21,37–39]. Prolonged and/or stronger IFN-α/β, IL-6 and TNF-α

expression results in the up-regulation of apoptosis-inducing proteins contributing to immunopathology and increased disease severity with greater morbidity in mice infected with IAV [35,36,40,41]. Furthermore, elevated cytokine levels correlate with more serious illness during influenza infection and duration of hospitalization in patients. Thus, increased pathology in IAV infected Ncbp3 deficient mice may be due to disease promoting effects of the innate immune system, which results in virus-induced tissue damage.

In sum, our study shows that the alternative CBC is required to mount appropriate immune responses and underlines the requirement of fine-tuned gene expression and export programs during virus infection processes.

Acknowledgements

We want to acknowledge the innate immunity laboratory for critical discussions and suggestions, A. Mann for technical assistance. K. Mayr, I. Paron and G. Sowa for maintaining mass spectrometers, Soo-Jin Weißenhorn and the MPI-B transgenic and animal facility for generating and housing mice. This work was supported by the Max-Planck Free Floater program to A.P., the German research foundation to A.P. (PI1084/3 and TRR179 TP11) and P.S. (SFB 1160, project 13), the Lunbeck Foundation to L.R. (R198-2015-171), an ERC starting grant to A.P (ERC-StG iVIP, 311339) and the Infect-Era and the German Federal Ministry of Education and Research (ERASE) to A.P.

Material and Methods

Reagents

Hybrid human IFN-α (IFN B/D) [42], which is highly active on mouse cells, was a kind gift from Peter Stäheli. Polyinosinic-polycytidylic acid (PIC; Sigma Aldrich; P9582) and siRNAs (Dharmacon) were transfected into cells using Neon Transfection System (Invitrogen). Primary antibodies were used according to manufacturer recommendation and were as follows: Ncbp3 (Atlas Antibodies; HPA008959), Histone H3 (Abcam; ab1791−100). For viral protein detection, we used antibodies against IAV (Millipore; AB1074) and VSV-G (Santa Cruz; sc-66180). Antibodies against NCBP2 were purified from serum of rabbits immunized with recombinant full-length protein purified from E. coli. Serum isolated from Ifit1 deficient mice immunized with full-length murine protein purified from E.coli was used to detect Ifit1. Antibodies against β-Actin (Santa Cruz; sc-47778) and secondary antibodies detecting mouse (Dako), rabbit (Sigma-Aldrich), goat (Santa Cruz) IgG were horseradish peroxidase-coupled. CellTiter-Glo® Luminescent Cell Viability Assay kit was purchased from Promega. RT-PCR reagents were from Takara/Clontech and Qiagen. Interleukin 6 was measured by enzyme-linked immunosorbent assay obtained from R&D Systems. Mass spectrometry grade trypsin and LysC was obtained from Wako Chemicals USA and Sigma-Aldrich, respectively.

Viruses

All used viruses are classified as BSL2 pathogens in Germany and experiments were carried out according official regulations. Wild-type Influenza A virus strain SC35M was used in mice experiments. In cell culture experiments, we used wild-type as well as NS1-deleted Influenza A virus strain SC35M [43] and wild-type and M2-mutated (M51R substitution in M2 protein) Vesicular stomatitis virus [44]. IAV-ΔNS1 and VSV-M2 are deficient in their ability to block innate immune

responses due to the lack of NS1 and mutation in M2 protein, respectively. In addition, we used an IAV-wt reporter virus [45] that expresses renilla luciferase and a VSV-wt reporter virus expressing firefly luciferase, which were a kind gift from Peter Reuther and Gert Zimmer, respectively.

Generation of Ncbp3 knockout mice

Genetically modified ES cell clones (JM8.F6; C57BL/6N background) carrying a promotor-less targeting cassette to generate "knockout-first" allelic mutation were obtained from the European Conditional Mouse Mutagenesis Program (EUCOMM) [46,47]. The cassette encodes for a neomycin resistance and a LacZ gene with a splice acceptor and a polyA site flanked by FRT sites. Additionally, the second exon of Ncbp3 is flanked by loxP sites, which results in a frame-shift mutation (Suppl. Fig 1a). Through the insertion of this cassette, a so-called tm1a mutation is generated resulting in the disruption of the targeted Ncbp3 gene.

Chimeras were generated by ES cell injection into C57BL/6 albino (B6(Cg)-Tyr<c-2J>/J) blastocyst donors, which were implanted into pseudo-pregnant mice. Germ-line male chimeras generated from one ES clone (Ncbp3 clone A06) were bred with C57BL/6 albino females to produce heterozygous, tm1a-carrying *F1* mice (C57BL/6N-Ncbp3^{tm1a}/Mpi). Heterozygous *F1* mice carrying the tm1a mutation in one Ncbp3 allele were further bred and maintained in the animal facility of the MPI of Biochemistry at SPF (specific pathogen-free) conditions. Body weight was monitored at weaning (age 19-23 days) and for a group of mice for a period of 13 month. Mouse husbandry were carried out in accordance with animal welfare regulations and have been approved by the responsible authorities (TVA 55.2-1-54-2532-116-2015).

Genotyping of Ncbp3 tm1a mice

Ncbp3^{tm1a} mice were genotyped using the following primers: Ncbp3_tm1a-fw 5`-CTGTATGTCCGGTCGTCATC-`3, Ncbp3_tm1a-rev 5´-GCCTGCATGTACCATGCATT-`3.

PCR conditions were as followed: (1) 94 °C for 1 minute (1×); (2) 94 °C for 30 sec, 57°C for 20 sec, 72°C for 30 sec (35×). PCR products were visualized using a 1% agarose gel electrophoresis.

LacZ/ß-Galactosidase assay

Ncbp3 expression in tissues was determined by ß-Galactosidase expression driven under endogenous Ncbp3 promotor from the LacZ gene of the tm1a cassette. Tissues were isolated from heterozygous ($Ncbp3^{tm1a+/-}$) and wild-type (background control) mice and homogenated in 2 µL reaction buffer per 1 mg tissue using SS matrix beads (MP Biomedicals) and the FastPrep-24 machine with the following setting: 4x 20 seconds, MP, 4 m/s. Lysates were cleared by centrifugation at 18000x g at 4°C for 30 minutes. Cleared lysated were used to determine ß-Galactosidase expression using the FluoReporter lacZ/Galactosidase Quantification kit (ThermoFischer Scientific; F-1905) according manufacturer`s guidelines. The assay was done in a 384-well format, wherefore, half of the manufacturer`s volumes were used. ß-Galactosidase signals were first divided by total protein content in the lysate measured using Pierce™ 660nm Protein Assay Reagent (ThermoFischer Scientific; 22660) according manufacturer`s protocol and second normalized to mean ß-Galactosidase signals in the respective tissue of the same gender.

Isolation and immortalization of mouse embryonic fibroblasts

Mouse embryonic fibroblasts were generated from day 13.5 embryos from heterozygous intercrosses. All following steps were carried out in a tissue culture hood under sterile conditions. First, uterine horns were dissected out and washed several times in 1x PBS. Embryos were isolated separately from the placenta and its embryonic sac and placed in sterile 1x PBS. Heads, red organs and extremities were removed from the embryos and the remaining body placed in fresh 1x PBS. Subsequently, the remaining body was minced and incubated for 10 minutes at 37 °C in 3 ml 0.05% trypsin/EDTA (ThermoFischer Scientific). Supernatant containing single cell suspension

was cleaned through a cell strainer (Greiner Bio-One International; 542070) adding 30 ml pre-warmed complete DMEM (ThermoFischer Scientific; containing 10% (v/v) fetal bovine serum and antibiotics (100 U/ml penicillin and 100 mg/ml streptomycin)) and subsequently centrifuged at 1000x g for 5 min. Trypsin treatment was repeated several times with remaining tissue. Cells were resuspended in 3 ml complete DMEM and seeded in a cell culture dish. After three passages, the cells were immortalized with SV40-LT by retroviral infection. Immortalized cells were selected with 3 µg/ml puromycin (Sigma-Aldrich; P8833) for eight passages and used for further experiments.

RNAi-mediated knockdown and cell growth assay

MEFs were electroporated with duplex siRNAs using the Neon Transfection System (Invitrogen) for target gene knockdowns. Duplex siRNAs were obtained from Dharmacon and had the following target sequences: Ncbp1 (#1: 5´-AUGCAGAAAUGGACCGAAU-3`, #2: 5`-CGUCUGGACACGAUGAGUA-3`, #3: 5`-GGUACGAUGUGAAACGGAU-3`, #4: 5`-AGGCCUAACACUCGCGUUU-3`), Ncbp2 (#1: 5`-CAGCAAAAGUGGUGAUAUA-3`, #2: 5`-GCAAUGCGGUACAUAAACG-3`, #3: 5` GUAUGGACGUGGACGGUCU-3`, #4: 5`-ACGAGUAUCGGGAGGACUA-3`) and scrambled (5`-AAGGTAATTGCGCGTGCAACT-3`). Transfection of siRNA and cell growth assay was described elsewhere [11]. Cell growth were analyzed at the indicated time points after the second/repeated siRNA transfection. For western blot analysis, cells were lysed in SSB buffer (62.5 mM Tris-HCl pH 6.8, 2 % sodium dodecyl sulfate, 10 % glycerol, 50 mM dithiothreitol, 0.01 % bromophenol blue) and boiled for 10 minutes at 95 °C and subjected to SDS–polyacrylamide gel electrophoresis and western blot analysis.

Cell growth assays

Cell growth was determined after RNAi-mediated knockdown of indicated time points using CellTiter-Glo (Promega) according to the manufacturer's instructions with the modifications decribed elsewhere [11].

***In vitro* virus infection**

MEFs were seeded on the day before infection and infected with IAV-wt, -ΔNS1 and VSV-wt, -M2. Duration and multiplicity of infection (MOI) is indicated in the figure legends. Supernatants of virus-infected cells were collected and virus titers were quantified by 50% tissue culture-infective dose (TCID50) assays on Vero E6 cell. For western blot analysis, cells were lysed in RIPA buffer (50 mM Tris-HCl pH 7.5, 150 mM NaCl, 0.25 % sodium deoxycholate, 1 % NP-40, 1 mM EDTA) and boiled in Laemmli buffer for 10 minutes at 95 °C and subjected to SDS–polyacrylamide gel electrophoresis and western blot analysis. For proteomic analysis, cell pellets were snap-frozen in liquid nitrogen before further processing. For RT-PCR analysis, RNA was isolated using NucleoSpin® RNA Plus kit (Macherey Nagel) or Direct-zol™ RNA MiniPrep Plus (Zymo Research) according manufacturer`s protocol.

For luciferase reporter viruses, MEFs were infected with IAV-wt and VSV-wt with different MOIs (indicated in figures). After 21 hours, cells were lysed in 1x passive lysis buffer (Promega) and lysates were mixed with equal volume of 2x renilla reagent solution (100 mM K_3PO_4, 500 mM NaCl, 1 mM EDTA, 25 mM Thiourea, 30 µM Coelenterazine) or 2x firefly reagent solution (20 mM Tris-HCl pH 7.8, 0.1 mM EDTA, 3.74 mM magnesium sulfate, 33.3 mM dithiothreitol, 0.27 mM coenzyme A, 0.47 mM D-luciferin, 0.53 mM adenosine triphosphate) in technical duplicates. Subsequently, luminescence was measured using an Infinite 200 PRO series micro plate reader

(Tecan). Luminescence counts were normalized to total protein concentration measured using Pierce™ 660nm Protein Assay Reagent (Thermo Fischer Scientific; 22660).

IFN and PIC treatment

All used concentrations are indicated in figures and figure legends. For IFN treatment, IFN B/D was added to the medium of homozygously attached MEFs in a 24-well format and cells were further incubated at 37°C with 5% CO_2. Cells were lysed in RIPA buffer (50 mM Tris-HCl pH 7.5, 150 mM NaCl, 0.25 % sodium deoxycholate, 1 % NP-40, 1 mM EDTA) and boiled in Laemmli buffer for 10 minutes at 95 °C and subjected to SDS–polyacrylamide gel electrophoresis and western blot analysis. For PIC stimulation, MEFs were transfected using the Neon Transfection System (Invitrogen) and seeded in a 24-well format. Supernatants were collected after indicated time points and IL6 concentration was determined using the mouse IL6 DuoSet Elisa (R&D systems) according manufacturer`s instructions.

Quantitative RT-PCR analysis

RNA was reverse transcribed using the PrimeScript™ RT Reagent Kit with or without gDNA Eraser (Takara/Clontech) and quantified by quantitative RT-PCR using the QuantiFast SYBR Green RT-PCR kit (Qiagen) and a CFX96 Touch Real-Time PCR Detection System (Bio-Rad). Each cycle consisted of 10 seconds at 95 °C and 30 seconds at 60 °C, followed by melting curve analysis. Primer sequences were as follows: mTbp (5`-CCTTCACCAATGACTCCTATGAC-3` and 5`-CAAGTTTACAGCCAAGATTCA-3`), mHmbs (5`-GAGTCTAGATGGCTCAGATAG CATGC-3` and 5`-CCTACAGACCAGTTAGCGCACATC-3`), mIfit3 (5`-TGGTCATGTGCC GTTACAGG-3` and 5`-GCTGCGAGGTCTTCAGACTT-3`), mIFN-β (5`-CGGAGAAGATGC AGAAGAGT-3` and 5`-TCAAGTGGAGAGCAGTTGAG-3`) and IAV-M [48] (5`-AGATGAGY CTTCTAACCGA-3` and 5`-GCAAAGACATCTTCAAGTYTC-3`).

Quantitative LC-MS/MS-based proteomics

For proteome analysis, four virus infections were performed in parallel. Frozen MEF cell pellets were lysed in U/A buffer (8 M Urea, 100 mM Tris-HCl pH 8.5) shaking for 10 minutes at RT followed by sonication for 15 minutes using a biorupter with 30 sec on/off cycles. Lysates were reduced with 10 mM dithiothreitol for 30 minutes at RT and incubated for 20 minutes with 55 mM iodacetamid in the dark to alkylate proteins. Subsequently, thirty microgram protein were digested with LysC (Wako Chemicals USA) and trypsin (Sigma-Aldrich), acified with TFA and desalted with C18 stage tips. Desalted peptides were analyzed by liquid chromatography coupled to mass spectrometry on a Q Exactive™ HF MS system (Thermo Fischer Scientific). Raw MS data were processed with MaxQuant version 1.5.5.1 [49] using the built-in Andromeda engine to search against mouse (UniprotKB, mus musculus; Proteome ID UP000000589; release 29/08/2016) and influenza A virus (UniprotKB, strain A/Seal/Massachusetts/1/1980 H7N7; Proteome ID UP000008576; release 27/07/2017) proteomes containing forward and reverse sequences. Label-free quantification (LFQ) algorithm [50] and Match between Runs option were used with standard settings.

LFQ intensities were log2-transformed and missing values imputed by normal distribution of values around the detection limit. Differential expression analysis was performed using the limma package in R [51]. Proteins showing differential response in infection between Ncbp3 ko and wt were calculated using the interaction term of the linear model. Significantly changing proteins must have met the criterion [abs((log2 fold change)x(-log10(interaction p-value)))]>3. ISGs among identified proteins were annotated using the INTERFEROME v2.0 database [52] filtered for mouse fibroblast and a minimal fold change of 2. Graphs were plotted in R and adapted using Adobe Illustrator.

478 ***In vivo* virus infection**

479 6-8 weeks old mice were anesthetized by intra-peritoneal injection of ketamine (100 µg per gram
480 body weight) and xylazine (5 µg per gram body weight). The infection was administered intranasal
481 with 1500 pfu of influenza A virus diluted in 40 µl sterile 1x PBS. Body weight (weight loss) and
482 survival was monitored for 13 days. In addition to mice that were found dead, mice with a body
483 weight loss greater than 25% of the initial weight were sacrificed and recorded as dead. All animal
484 experiments have been performed according to animal welfare regulations and have been approved
485 by the responsible authorities (Freiburg, G-12/46). For histology of lung sections, left lobes were
486 isolated from mice, fixed in 4% PFA and embedded in paraffin. Paraffin-embeded sections were
487 stained with hematoxylin and eosin and cell infiltration scores (0-5) were defined for severity of
488 tissue inflammation as described before [53].

Figure legends

Figure 1: Characterization of Ncbp3 knockout mice.

(a) Mendelian frequency of Ncbp3 tm1a crossings. Expected and obtained Mendelian ratios of *F1* Ncbp3 tm1a crosses are presented. **(b)** Western blot analysis of Ncbp3 tm1a MEFs. MEFs were generated from Ncbp3 wild-type (wt; $Ncbp3^{+/+}$, $tm1a^{-/-}$), heterozygous (het; $Ncbp3^{+/-}$, $tm1a^{+/-}$) and knockout (ko; $Ncbp3^{-/-}$, $tm1a^{+/+}$) embryos and western blot analysis was performed using antibodies against indicated proteins. **(c)** Knockdown efficiency of Ncbp2 was confirmed by western blotting staining with antibodies against indicated proteins. **(d)** Cell growth of Ncbp3 wt and ko MEFs after RNAi-mediated knockdown. MEFs were electroporated twice with siRNAs against Ncbp1, Ncbp2 and a control siRNA (siScrambled). After repeated knokdown, cell viability was measured using a luminescence-based assay at indicated time points. RNAi treatment was performed in triplicates and the graph displays the mean ± SD. Ncbp1/2/3, Nuclear cap-binding protein 1/2/3; wt, wild-type; het, heterozygous; ko, knockout; RLU, relative light units.

Figure 2: Loss of Ncbp3 supports IAV replication *in vitro*.

(a) Virus load in Ncbp3 wt and ko MEFs using different MOIs. MEFs were infected with an IAV-wt reporter virus (strain SC35M) expressing renilla luciferase using indicated MOIs for 21 hours. Luminescence counts were normalized to total protein concentration. Graphs display the mean ± SD of virus infections performed in triplicates. **(b)** Accumulation of infectious viral particles in supernatants of IAV-infected MEFs quantified by TCID50. MEFs were infected with IAV-wt and -ΔNS1 (strain SC35M) using a MOI of 0.01 and 0.1, respectively, for 24 hours. *** $P<0.001$ as analyzed by two-way analysis of variance (ANOVA) statistics with Bonferroni`s post-test. **(c)** RNA expression in IAV-infected Ncbp3 wt and ko MEFs. MEFs were infected with indicated MOIs of IAV-wt and -ΔNS1 for 24 hours and RNA was amplified by qRT-PCR using

specific primers for murine Ifit3, IFN-β, Hmbs and IAV-M. Data were normalized to murine Tbp RNA and the mean of two technical replicates are represented. **(d)** Western blot analysis of IAV-infected Ncbp3 wt and ko MEFs using antibodies against indicated proteins. MOI, multiplicity of infection; IAV, Influenza A virus; Tbp, TATA-box-binding protein; Ifit1, Interferon-induced protein with tetratricopeptide repeats 1; IFN-β, Interferon beta; IAV-M, Influenza A virus matrix protein; IAV-NP, Influenza A virus nucleoprotein; IAV-NS1, Influenza A virus non-structural protein 1;AU, arbitrary units.

Figure 3: Ncbp3 depletion increases VSV infection *in vitro.*

(a) Virus expression in Ncbp3 wt and ko MEFs using different MOIs. MEFs were infected with a VSV-wt reporter virus expressing firefly luciferase using indicated MOIs for 21 hours. Luminescence counts were normalized to total protein concentration. Graphs display the mean ± SD of virus infections performed in triplicates. **(b)** Viral particle accumulation in supernatants of VSV-infected MEFs quantified by TCID50. MEFs were infected with VSV-wt and -M2 (M51R substitution in the matrix protein) using a MOI of 0.001 for 24 hours. *** $P<0.001$ as analyzed by two-way analysis of variance (ANOVA) statistics with Bonferroni`s post-test. **(c)** Western blot analysis of VSV-infected Ncbp3 wt and ko MEFs using antibodies against indicated proteins. MOI, multiplicity of infection; VSV, Vesicular stomatitis virus; VSV-G, Vesicular stomatitis virus glycoprotein; Ifit1, Interferon-induced protein with tetratricopeptide repeats 1; Ncbp3, Nuclar cap-binding protein 3.

Figure 4: Global proteomic analyses suggest differential expression of a subset of ISGs.

(a) Experimental schematic of proteomic analyses. Ncbp3 wt and ko MEFs were infected for 24 hours with IAV-wt (MOI 0.1). Cells were lysed and proteomic changes were analyzed by shotgun liquid chromatography-tandem MS (LC-MS/MS). **(b)** Heat map showing hierarchical clustering

(Euclidean distances) of viral protein expression in infected MEFs. Log2-transformed non-imputed LFQ intensities are shown for each individual replicate. Grey color denotes missing values (ND). **(c)** Scatter plot showing differential expression of host proteins in Ncbp3 ko and wt MEFs of of SC35M infection versus uninfected. Proteins showing a significant differential response between Ncbp3 ko and wt upon infection are highlighted in blue or green if they were classified as ISGs. ISGs that were significantly changing either in Ncbp3 ko or wt upon infection but show no differential response between Ncbp3 ko and wt are highlighted in pink. **(c)** Z-scored replicate expression levels of significantly changing ISGs between NCBP3 ko and wt. Grey color represents missing values (ND).

Figure 5: Ncbp3 depletion alters innate immune responses.

(a-b) IL6 secretion in Ncbp3 wt and ko MEFs. MEFs were electroporated with 0.15 **(a)**, 0.5 and 0.05 **(b)** µg/mL PIC and IL6 concentration in supernatants was determined 3, 8 and 24 hours **(a)** or 24 hours **(b)** after stimulation by enzyme-linked immunosorbent assay (Elisa). Data represent the individual and the mean value ± SD of PIC treatments performed in triplicates. *** $P<0.001$, ns $P>0.05$ as analyzed by two-way analysis of variance (ANOVA) statistics with Bonferroni`s post-test. **(c)** Ifit1 expression in Ncbp3 wt and ko MEFs stimulated with IFN. Ncbp3 wt and ko MEFs were treated with IFN B/D for 18 hours. Ifit1 expression was determined by western blot analysis. Depletion of Ncbp3 and equal loading was confirmed by western blotting against indicated proteins. IL6, interleukin 6; PIC, Polyinosinic:polycytidylic acid; IFN, interferon; Ncbp3, Nuclear cap-binding protein 3; Ifit1, Interferon-induced protein with tetratricopeptide repeats 1.

Figure 6: Ncbp3 is required to defend IAV infection *in vivo*.

(a) Survival of Ncbp3 wt and ko mice after infection with IAV-wt (strain SC35M). Mice (n=9, C57BL/6 background) were infected intranasal with 1500 pfu of IAV and survival was monitored over 13 days. **(b)** Histological sections of lungs from IAV-wt infected Ncbp3 wt and ko mice. Mice (n=5, C57BL/6 background) were infected with 1500 pfu of IAV for 7 days and paraffin embedded sections were stained with hematoxylin and eosin. Representative images are shown in 5x (left panel) and 20x (right panel) magnification. **(c)** Cell infiltration in Ncbp3 wt and ko lung tissue. Lung sections of 7 days IAV-infected mice (n=5, C57BL/6 background) were scored for severity of cell infiltration (0-5). Data represent the individual and the mean value ± SD of arbitrary infiltration scores. * $P<0.05$ as analyzed by one-way analysis of variance (ANOVA) with non-parametric Kruskal-Wallis statistics and Dunn`s post-test.

Supplementary figure 1

(a) Schematic overview of the promoterless tm1a cassette inserted in intronic Ncbp3 region. The cassette was inserted into the first intron of Ncbp3 gene locus (NM_025818.3) flanked by FRT sites and encodes for a neomycin resistance and a LacZ gene with splice acceptor and a polyA site. Exon 2 is flanked by loxP sites which after recombination would result in a frame-shift mutation. **(b)** Genotyping PCR of Ncbp3 tm1a mice. PCR amplification results in a 324 bp construct for Ncbp3 wt mice and a 204 bp construct for Ncbp3 ko (tm1a promoterless cassette insertion) mice. **(c)** Genotypes of Ncbp3 chimera crossings. Male Ncbp3 chimeras were crossed with C57BL/6 albino and obtained genotypes are represented. **(d)** Genotypes of Ncbp3 tm1a crossings. Ncbp3 tm1a mice were bred to homogeneity and genotypes obtained for the indicated breeding combination are shown. **(e)** Body weight of Ncbp3 wt and ko mice at the age of weaning. Body weight of 52 animals per genotype in the age of 19-23 days were monitored. *** $P<0.001$ as

analyzed by one-way analysis of variance (ANOVA) statistics with Bonferroni`s post-test. **(f)** Body weight development of Ncbp3 wt and ko mice over 13 month. Body weight of 6 animals per genotype and gender were monitored for 13 month. **(g)** Expression of Ncbp3 in different tissues. LacZ/β-galactosidase expression is driven under the endogenous Ncbp3 promotor. Ncbp3 expression was determined in lysed tissues using a fluorometric β-galactosidase detection assay. Expression from heterozygous Ncbp3 tm1a mice was normalized to background β-galactosidase signal using tissues from Ncbp3 wt mice. Tissues from 3 animals per genotype and gender were isolated and individual and mean values ± SD normalized to background signal are presented. FRT, Flippase Recognition Target; EnS 2A, splice acceptor site; T2A, peptide sequence with self-cleaving function; lacZ, lacZ gene encoding for ß-galactosidase; neo, neomycine resistance gene; pA, simian virus 40 polyadenylation signal; loxP, locus of X-over P1; bp, base pair; wt, wild-type; ko, knockout; Ncbp3, Nuclear cap-binding protein 3.

References

1. Gebhardt, A., Laudenbach, B. T. & Pichlmair, A. Discrimination of Self and Non-Self Ribonucleic Acids. *J. Interf. Cytokine Res.* **37,** 184–197 (2017).

2. Zhang, J.-M. & An, J. Cytokines, inflammation, and pain. *Int. Anesthesiol. Clin.* **45,** 27–37 (2007).

3. Schneider, W. M., Chevillotte, D. & Rice, C. M. Interferon-Stimulated Genes: A Complex Web of Host Defenses. *Annu. Rev. Immunol* **32,** 513–45 (2014).

4. Schoggins, J. W. & Rice, C. M. Interferon-stimulated genes and their antiviral effector functions. *Curr. Opin. Virol.* **1,** 519–25 (2011).

5. Schoggins, J. W. Interferon-stimulated genes: roles in viral pathogenesis. *Curr. Opin. Virol.* **6,** 40–46 (2014).

6. Gonatopoulos-Pournatzis, T. & Cowling, V. H. Cap-binding complex (CBC). *Biochem. J.* **457,** 231–42 (2014).

7. Topisirovic, I., Svitkin, Y. V, Sonenberg, N. & Shatkin, A. J. Cap and cap-binding proteins in the control of gene expression. *Wiley Interdiscip. Rev. RNA* **2,** 277–98 (2011).

8. Mazza, C., Ohno, M., Segref, A., Mattaj, I. W. & Cusack, S. Crystal Structure of the Human Nuclear Cap Binding Complex. *Mol. Cell* **8,** 383–396 (2001).

9. Jalkanen, A. L., Coleman, S. J. & Wilusz, J. Determinants and implications of mRNA poly(A) tail size--does this protein make my tail look big? *Semin. Cell Dev. Biol.* **34,** 24–32 (2014).

10. Köhler, A. & Hurt, E. Exporting RNA from the nucleus to the cytoplasm. *Nat. Rev. Mol.*

Cell Biol. **8,** 761–73 (2007).

11. Gebhardt, A. *et al*. mRNA export through an additional cap-binding complex consisting of NCBP1 and NCBP3. *Nat. Commun.* **6,** 8192 (2015).

12. Rivas, H. G., Schmaling, S. K. & Gaglia, M. M. Shutoff of host gene expression in influenza A virus and herpesviruses: Similar mechanisms and common themes. *Viruses* **8,** 1–26 (2016).

13. Kuss, S. K., Mata, M. a, Zhang, L. & Fontoura, B. M. a. Nuclear imprisonment: viral strategies to arrest host mRNA nuclear export. *Viruses* **5,** 1824–49 (2013).

14. Jhawar, S. R. *et al*. Oncolytic Viruses—Natural and Genetically Engineered Cancer Immunotherapies. *Front. Oncol.* **7,** 202 (2017).

15. Bartlett, D. L. *et al.* Oncolytic viruses as therapeutic cancer vaccines. *Mol. Cancer* **12,** 103 (2013).

16. Chiocca, E. A. & Rabkin, S. D. Oncolytic viruses and their application to cancer immunotherapy. *Cancer Immunol. Res.* **2,** 295–300 (2014).

17. Hale, B. G., Randall, R. E., Ortin, J. & Jackson, D. The multifunctional NS1 protein of influenza A viruses. *J. Gen. Virol.* **89,** 2359–2376 (2008).

18. García-Sastre, A. Induction and evasion of type I interferon responses by influenza viruses. *Virus Res.* **162,** 12–18 (2011).

19. Lichty, B. D., Power, A. T., Stojdl, D. F. & Bell, J. C. Vesicular stomatitis virus: re-inventing the bullet. *Trends Mol. Med.* **10,** 210–216 (2004).

20. Zeitlinger, J. & Stark, A. Developmental gene regulation in the era of genomics. *Dev.*

Biol. **339,** 230–239 (2010).

21. Ivashkiv, L. B. & Donlin, L. T. Regulation of type I interferon responses. *Nat. Rev. Immunol.* **14,** 36–49 (2013).

22. Carpenter, S., Ricci, E. P., Mercier, B. C., Moore, M. J. & Fitzgerald, K. A. Post-transcriptional regulation of gene expression in innate immunity. *Nat. Publ. Gr.* **14,** (2014).

23. Bego, M. G. *et al.* Vpu Exploits the Cross-Talk between BST2 and the ILT7 Receptor to Suppress Anti-HIV-1 Responses by Plasmacytoid Dendritic Cells. *PLOS Pathog.* **11,** e1005024 (2015).

24. Cao, W. *et al.* Regulation of TLR7/9 responses in plasmacytoid dendritic cells by BST2 and ILT7 receptor interaction. *J. Exp. Med.* **206,** 1603–1614 (2009).

25. Watanabe, R., Leser, G. P. & Lamb, R. A. Influenza virus is not restricted by tetherin whereas influenza VLP production is restricted by tetherin. *Virology* **417,** 50–56 (2011).

26. Weidner, J. M. *et al.* Interferon-Induced Cell Membrane Proteins, IFITM3 and Tetherin, Inhibit Vesicular Stomatitis Virus Infection via Distinct Mechanisms. *J. Virol.* **84,** 12646–12657 (2010).

27. Yang, H. *et al.* Structural insight into the mechanisms of enveloped virus tethering by tetherin. *Proc. Natl. Acad. Sci.* **107,** 18428–18432 (2010).

28. Gu, G., Zhao, D., Yin, Z. & Liu, P. BST-2 binding with cellular MT1-MMP blocks cell growth and migration via decreasing MMP2 activity. *J. Cell. Biochem.* **113,** 1013–1021 (2012).

29. Stetler-Stevenson, W. G., Krutzsch, H. C. & Liotta, L. A. Tissue inhibitor of metalloproteinase (TIMP-2). A new member of the metalloproteinase inhibitor family. *J. Biol. Chem.* **264,** 17374–8 (1989).

30. Goldberg, G. I. *et al.* Human 72-kilodalton type IV collagenase forms a complex with a tissue inhibitor of metalloproteases designated TIMP-2. *Proc. Natl. Acad. Sci. U. S. A.* **86,** 8207–11 (1989).

31. Chattopadhyay, N., Mitra, A., Frei, E. & Chatterjee, A. Human cervical tumor cell (SiHa) surface alphavbeta3 integrin receptor has associated matrix metalloproteinase (MMP-2) activity. *J. Cancer Res. Clin. Oncol.* **127,** 653–8 (2001).

32. Zhao, C., Hsiang, T.-Y., Kuo, R.-L. & Krug, R. M. ISG15 conjugation system targets the viral NS1 protein in influenza A virus-infected cells. *Proc. Natl. Acad. Sci. U. S. A.* **107,** 2253–8 (2010).

33. Zhang, D. & Zhang, D.-E. Interferon-stimulated gene 15 and the protein ISGylation system. *J. Interferon Cytokine Res.* **31,** 119–30 (2011).

34. Sandler, N. G. *et al.* Type I interferon responses in rhesus macaques prevent SIV infection and slow disease progression. *Nature* **511,** 601–605 (2014).

35. Peiris, J. S. M., Cheung, C. Y., Leung, C. Y. H. & Nicholls, J. M. Innate immune responses to influenza A H5N1: friend or foe? *Trends Immunol.* **30,** 574–584 (2009).

36. Newton, A. H., Cardani, A. & Braciale, T. J. The host immune response in respiratory virus infection: balancing virus clearance and immunopathology. *Semin. Immunopathol.* **38,** 471–482 (2016).

37. Quicke, K. M., Diamond, M. S. & Suthar, M. S. Negative regulators of the RIG-I-like receptor signaling pathway. *Eur. J. Immunol.* **47,** 615–628 (2017).

38. Mogensen, T. H. Pathogen recognition and inflammatory signaling in innate immune defenses. *Clinical Microbiology Reviews* **22,** 240–273 (2009).

39. Cao, X. Self-regulation and cross-regulation of pattern-recognition receptor signalling in health and disease. *Nature Reviews Immunology* **16,** 35–50 (2016).

40. McNab, F., Mayer-Barber, K., Sher, A. & Wack, A. Type I interferons in infectious disease. *Nat. Publ. Gr.* **15,** (2015).

41. Peiris, J. S. M., Hui, K. P. Y. & Yen, H.-L. Host response to influenza virus: protection versus immunopathology. *Curr. Opin. Immunol.* **22,** 475–81 (2010).

42. Horisberger, M. A. & de Staritzky, K. A Recombinant Human Interferon- B/D Hybrid with a Broad Host-range. *J. Gen. Virol.* **68,** 945–948 (1987).

43. Kochs, G. *et al.* Properties of H7N7 influenza A virus strain SC35M lacking interferon antagonist NS1 in mice and chickens. *J. Gen. Virol.* **88,** 1403–1409 (2007).

44. Pichlmair, A. *et al.* IFIT1 is an antiviral protein that recognizes 5′-triphosphate RNA. *Nat. Immunol.* **12,** 624–630 (2011).

45. Reuther, P. *et al*. Generation of a variety of stable Influenza A reporter viruses by genetic engineering of the NS gene segment. *Sci. Rep.* **5,** 1–17 (2015).

46. Skarnes, W. C. *et al*. A conditional knockout resource for the genome-wide study of mouse gene function. *Nature* **474,** (2011).

47. Pettitt, S. J. *et al*. Agouti C57BL/6N embryonic stem cells for mouse genetic resources.

Nat. Methods **6,** 493–495 (2009).

48. Penski, N. *et al.* Highly pathogenic avian influenza viruses do not inhibit interferon synthesis in infected chickens but can override the interferon-induced antiviral state. *J. Virol.* **85,** 7730–41 (2011).

49. Cox, J. & Mann, M. MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat. Biotechnol.* **26,** 1367–72 (2008).

50. Cox, J. *et al.* Accurate Proteome-wide Label-free Quantification by Delayed Normalization and Maximal Peptide Ratio Extraction, Termed MaxLFQ. *Mol. Cell. Proteomics* **13,** 2513–2526 (2014).

51. Ritchie, M. E. *et al.* limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* **43,** e47–e47 (2015).

52. Rusinova, I. *et al.* Interferome v2.0: an updated database of annotated interferon-regulated genes. *Nucleic Acids Res.* **41,** D1040-6 (2013).

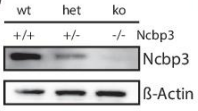
53. Michaudel, C. *et al.* Ozone exposure induces respiratory barrier biphasic injury and inflammation controlled by Interleukin-33. *J. Allergy Clin. Immunol.* (2018). doi:10.1016/j.jaci.2017.11.044

Figure 1

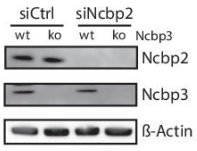
a

Offspring \ Crossing type	♀Ncbp3 het × ♂Ncbp3 het		♀Ncbp3 ko × ♂Ncbp3 het		♀Ncbp3 het × ♂Ncbp3 ko	
	expected	obtained	expected	obtained	expected	obtained
Ncbp3 ko	1	0.7	1	0.7	1	0.5
Ncbp3 het	2	2.2	1	1.3	1	1.5
Ncbp3 wt	1	1.1				

b



c



d

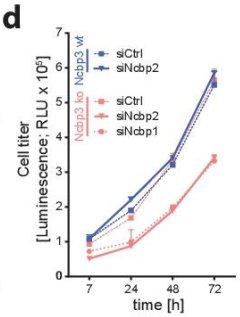


Figure 2

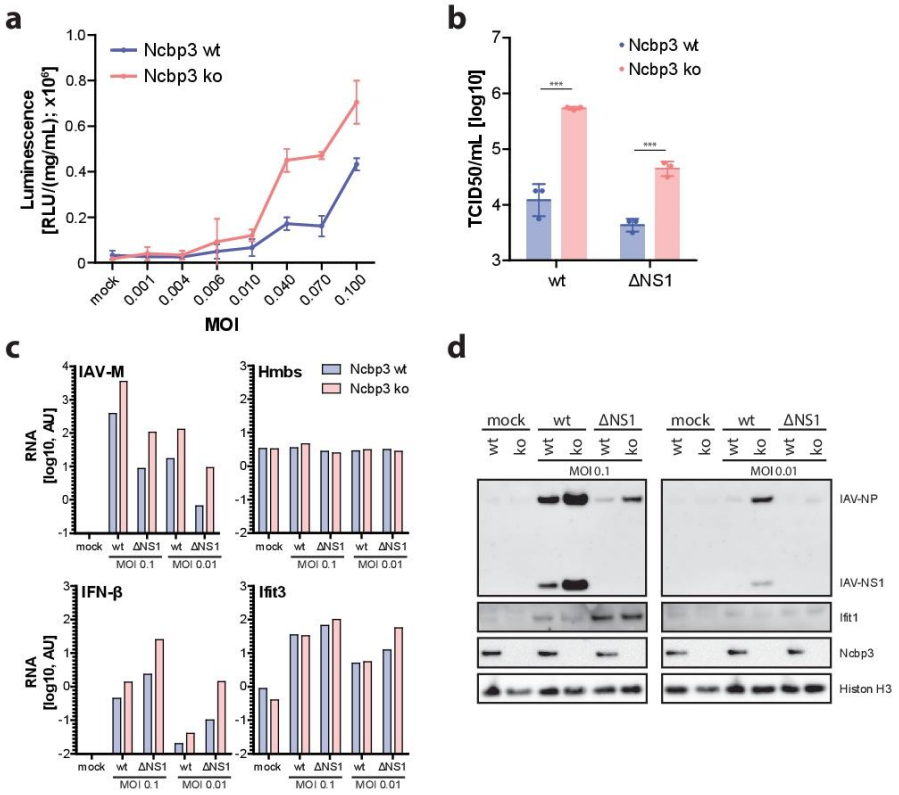

a
Ncbp3 wt
Ncbp3 ko
Luminescence [RLU/(mg/mL); x10⁶]
MOI
mock
0.001
0.004
0.006
0.010
0.040
0.070
0.100
b
Ncbp3 wt
Ncbp3 ko
TCID50/mL [log10]
wt
ΔNS1
c
IAV-M
Hmbs
IFN-β
Ifit3
RNA [log10, AU]
MOI 0.1
MOI 0.01
d
mock
wt
ΔNS1
ko
IAV-NP
IAV-NS1
Ifit1
Ncbp3
Histon H3

Figure 3

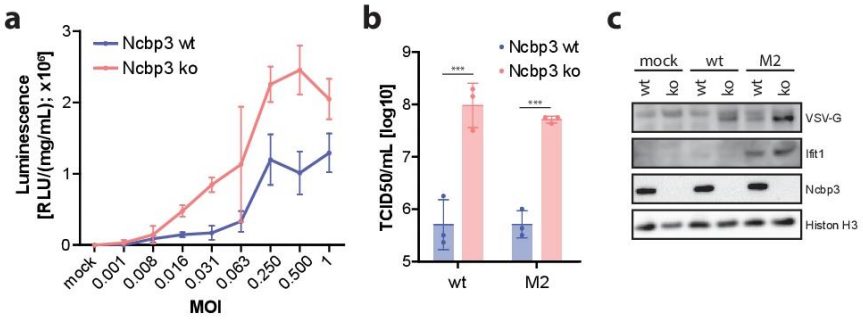

a
Ncbp3 wt
Ncbp3 ko
Luminescence [RLU/(mg/mL); x10⁶]
MOI
mock
0.001
0.008
0.016
0.031
0.063
0.250
0.500
1
b
Ncbp3 wt
Ncbp3 ko
TCID50/mL [log10]
wt
M2
c
mock
wt
M2
VSV-G
Ifit1
Ncbp3
Histon H3

Figure 4

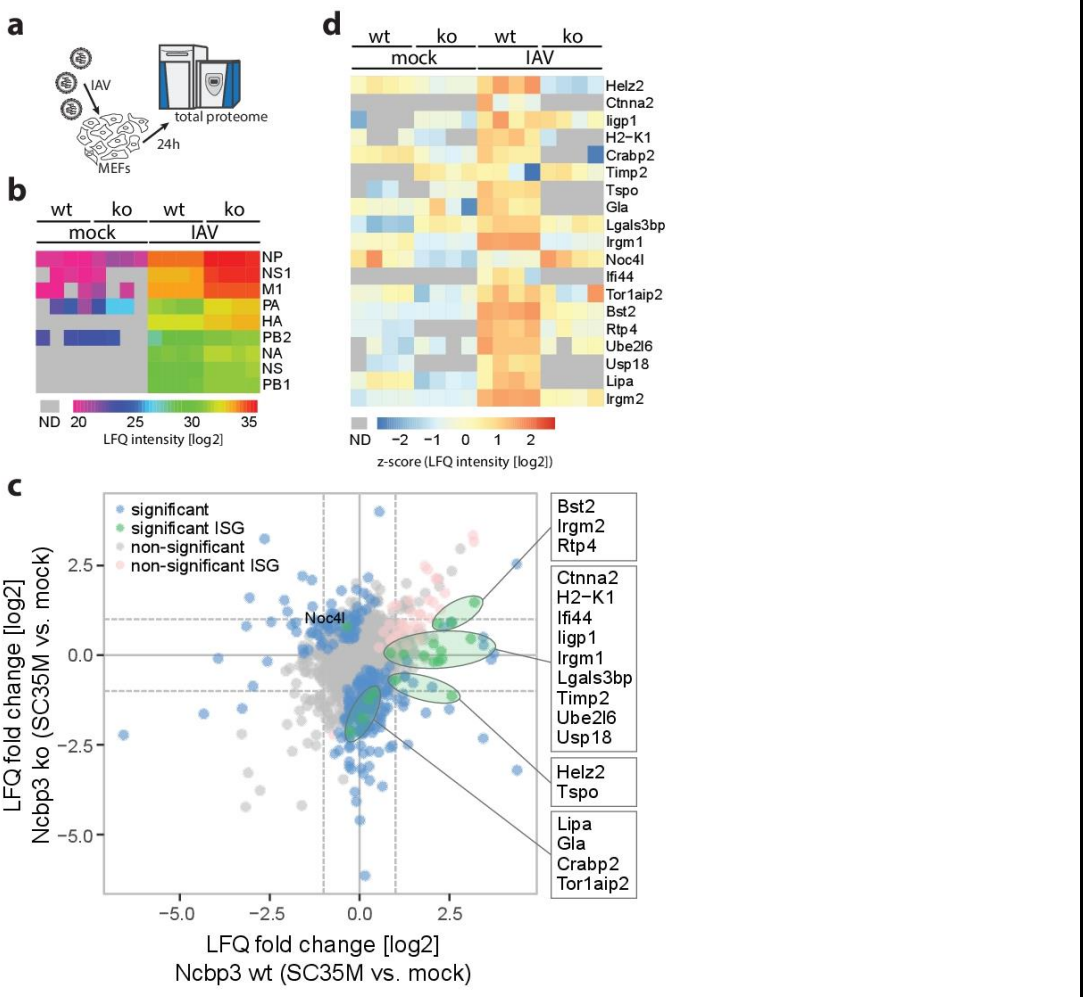
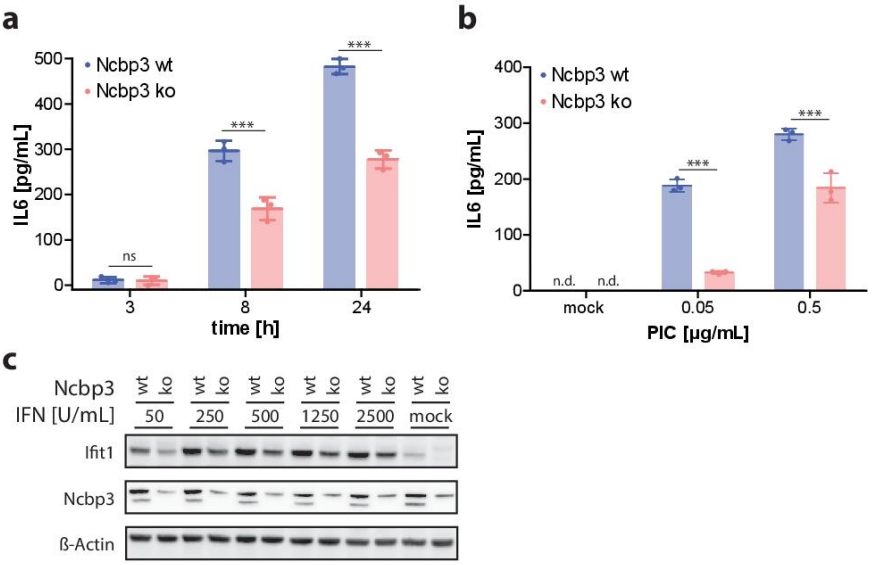
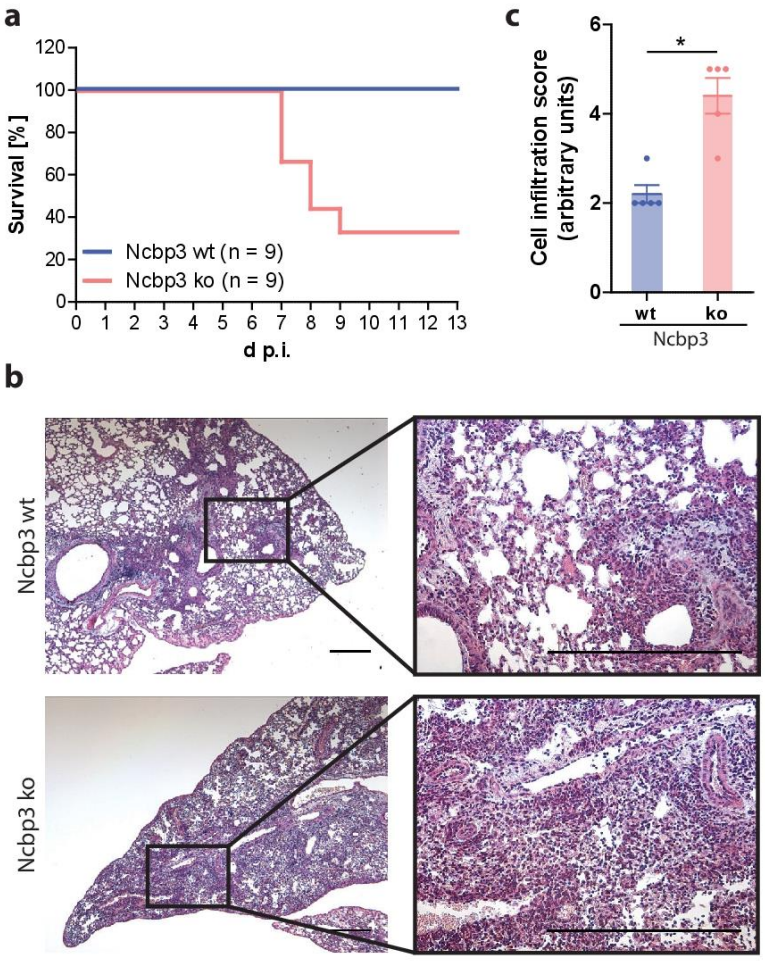

a
IAV
total proteome
24h
MEFs
b
wt ko wt ko
mock IAV
NP
NS1
M1
PA
HA
PB2
NA
NS
PB1
ND 20 25 30 35
LFQ intensity [log2]
d
wt ko wt ko
mock IAV
Helz2
Ctnna2
Iigp1
H2−K1
Crabp2
Timp2
Tspo
Gla
Lgals3bp
Irgm1
Noc4l
Ifi44
Tor1aip2
Bst2
Rtp4
Ube2l6
Usp18
Lipa
Irgm2
ND −2 −1 0 1 2
z-score (LFQ intensity [log2])
c
significant
significant ISG
non-significant
non-significant ISG
Noc4l
Bst2
Irgm2
Rtp4
Ctnna2
H2−K1
Ifi44
Iigp1
Irgm1
Lgals3bp
Timp2
Ube2l6
Usp18
Helz2
Tspo
Lipa
Gla
Crabp2
Tor1aip2
2.5
0.0
−2.5
−5.0
LFQ fold change [log2]
Ncbp3 ko (SC35M vs. mock)
−5.0 −2.5 0.0 2.5
LFQ fold change [log2]
Ncbp3 wt (SC35M vs. mock)

Figure 5

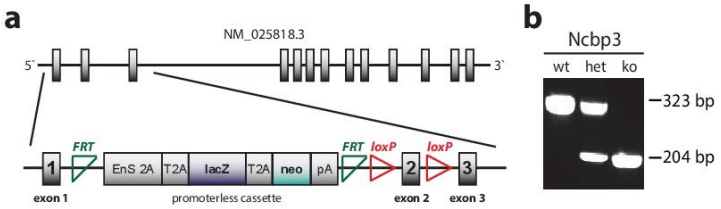

a
Ncbp3 wt
Ncbp3 ko
IL6 [pg/mL]
500
400
300
200
100
0
ns
***
***
3
8
24
time [h]
b
Ncbp3 wt
Ncbp3 ko
IL6 [pg/mL]
400
300
200
100
0
n.d.
n.d.
***
***
mock
0.05
0.5
PIC [µg/mL]

c
Ncbp3
wt
ko
IFN [U/mL]
50
250
500
1250
2500
mock
Ifit1
Ncbp3
ß-Actin

Figure 6


a
Survival [%]
120
100
80
60
40
20
0
0 1 2 3 4 5 6 7 8 9 10 11 12 13
d p.i.
Ncbp3 wt (n = 9)
Ncbp3 ko (n = 9)
c
Cell infiltration score (arbitrary units)
6
4
2
0
*
wt
ko
Ncbp3
b
Ncbp3 wt
Ncbp3 ko

Supplementary figure 1



c

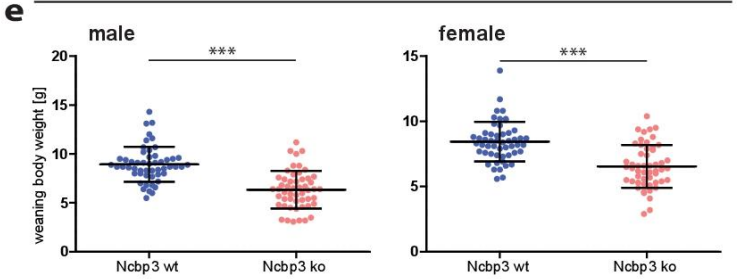
Genotypes of Ncbp3 ^{chimeras} crossings

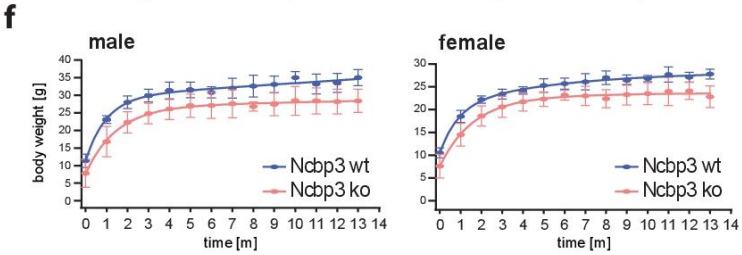
Crosing type		Ncbp3 wt (% of total)	Ncbp3 het (% of total)	Total
♀ B6 albino	x ♂ Ncbp3 ^{chimera}	19 (51)	18 (49)	37

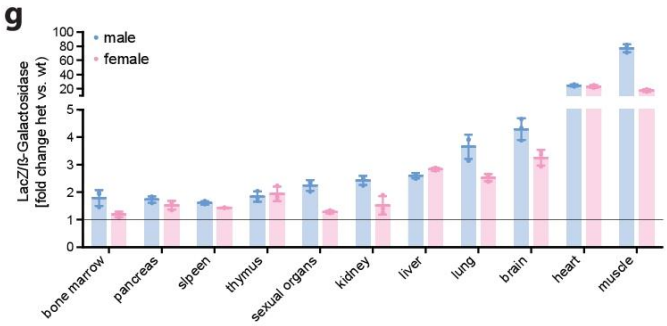
d

Genotypes of Ncbp3^{tm1a} crossings

Crosing type		Ncbp3 wt (% of total)	Ncbp3 het (% of total)	Ncbp3 ko (% of total)	Total
♀ Ncbp3 het	x ♂ Ncbp3 het	93 (28)	185 (55)	57 (17)	335
♀ Ncbp3 ko	x ♂ Ncbp3 het		26 (63)	15 (37)	41
♀ Ncbp3 het	x ♂ Ncbp3 ko		29 (76)	9 (24)	38







2.3 PUBLICATION 3: Phosphorylation of serine 225 in hepatitis C virus NS5A regulates protein-protein interactions

Goonawardane, N., **Gebhardt, A.**, Bartlett, C., Pichlmair, A., and Harris, M. (2017). Phosphorylation of serine 225 in hepatitis C virus NS5A regulates protein-protein interactions. *J. Virol.* JVI.00805-17.

The manuscript "Phosphorylation of serine 225 in hepatitis C virus NS5A regulates protein-protein interactions" describes the molecular mechanism of NS5A serine 225 (S225) in regulating hepatitis C virus (JFH-1, genotype 2a) genome replication. Mass spectrometry based analysis revealed that the phosphoablatant mutant (S225A) lost the ability to interact with a number of cellular proteins including the nucleosome assembly protein 1-like protein 1 (NAP1L1), bridging integrator 1 (Bin1, also known as amphiphysin II), and vesicle-associated membrane protein-associated protein A (VAP-A). Previous studies demonstrated that infection with NS5A phosphoablatant mutant (S225A) HCV results in reduced virus replication and perinuclear distribution of NS5A. Similar, knockdown of the three cellular NS5A interacting proteins, namely NAP1L1, Bin1 and VAP-A, impaired viral genome replication and led to perinuclear distribution of NS5A demonstrating the importance of NS5A S225 phosphorylation in regulating cellular interactions which are required for successful viral genome replication.

The main text including material and methods is represented in the following. Supplementary tables can be downloaded from *Journal of Virology*`s publishing website (http://jvi.asm.org/content/early/2017/06/08/JVI.00805-17).

AMERICAN SOCIETY FOR MICROBIOLOGY | Journal of Virology®

VIRUS-CELL INTERACTIONS

Phosphorylation of Serine 225 in Hepatitis C Virus NS5A Regulates Protein-Protein Interactions

Niluka Goonawardane,[a] Anna Gebhardt,[b] Christopher Bartlett,[a] Andreas Pichlmair,[b] Mark Harris[a]

School of Molecular and Cellular Biology, Faculty of Biological Sciences and Astbury Centre for Structural Molecular Biology, University of Leeds, Leeds, United Kingdom[a]; Max Planck Institute of Biochemistry, Martinsried, Germany[b]

ABSTRACT Hepatitis C virus (HCV) nonstructural protein 5A (NS5A) is a phosphoprotein that plays key, yet poorly defined, roles in both virus genome replication and virion assembly/release. It has been proposed that differential phosphorylation could act as a switch to regulate the various functions of NS5A; however, the mechanistic details of the role of this posttranslational modification in the virus life cycle remain obscure. We previously reported (D. Ross-Thriepland, J. Mankouri, and M. Harris, J Virol 89:3123–3135, 2015, doi:10.1128/JVI.02995-14) a role for phosphorylation at serine 225 (S225) of NS5A in the regulation of JFH-1 (genotype 2a) genome replication. A phosphoablatant (S225A) mutation resulted in a 10-fold reduction in replication and a perinuclear restricted distribution of NS5A, whereas the corresponding phosphomimetic mutation (S225D) had no phenotype. To determine the molecular mechanisms underpinning this phenotype we conducted a label-free proteomics approach to identify cellular NS5A interaction partners. This analysis revealed that the S225A mutation disrupted the interactions of NS5A with a number of cellular proteins, in particular the nucleosome assembly protein 1-like protein 1 (NAP1L1), bridging integrator 1 (Bin1, also known as amphiphysin II), and vesicle-associated membrane protein-associated protein A (VAP-A). These interactions were validated by immunoprecipitation/Western blotting, immunofluorescence, and proximity ligation assay. Importantly, small interfering RNA (siRNA)-mediated knockdown of NAP1L1, Bin1 or VAP-A impaired viral genome replication and recapitulated the perinuclear redistribution of NS5A seen in the S225A mutant. These results demonstrate that S225 phosphorylation regulates the interactions of NS5A with a defined subset of cellular proteins. Furthermore, these interactions regulate both HCV genome replication and the subcellular localization of replication complexes.

IMPORTANCE Hepatitis C virus is an important human pathogen. The viral nonstructural 5A protein (NS5A) is the target for new antiviral drugs. NS5A has multiple functions during the virus life cycle, but the biochemical details of these roles remain obscure. NS5A is known to be phosphorylated by cellular protein kinases, and in this study, we set out to determine whether this modification is required for the binding of NS5A to other cellular proteins. We identified 3 such proteins and show that they interacted only with NS5A that was phosphorylated on a specific residue. Furthermore, these proteins were required for efficient virus replication and the ability of NS5A to spread throughout the cytoplasm of the cell. Our results help to define the function of NS5A and may contribute to an understanding of the mode of action of the highly potent antiviral drugs that are targeted to NS5A.

KEYWORDS hepatitis C virus, NS5A, phosphorylation, RNA replication, protein phosphorylation, subgenomic replicon

Received 15 May 2017 **Accepted** 3 June 2017

Accepted manuscript posted online 14 June 2017

Citation Goonawardane N, Gebhardt A, Bartlett C, Pichlmair A, Harris M. 2017. Phosphorylation of serine 225 in hepatitis C virus NS5A regulates protein-protein interactions. J Virol 91:e00805-17. https://doi.org/10.1128/JVI.00805-17.

Editor Michael S. Diamond, Washington University School of Medicine

Address correspondence to Mark Harris, m.harris@leeds.ac.uk.

Hepatitis C virus (HCV) infects approximately 130 to 170 million individuals worldwide and is a leading cause of liver disease (1). There is no vaccine available, and current antiviral treatments are less effective against some viral strains (2). HCV belongs to the *Flaviviridae* family (genus *Hepacivirus*) of enveloped viruses with a positive-sense RNA genome (9.6 kb) coding for a single polyprotein that is processed co- and posttranslationally by viral and host proteases, yielding four structural proteins (core, E1, E2, and p7) and six nonstructural proteins (NS2, NS3, NS4A, NS4B, NS5A, and NS5B) (3). NS3 to NS5B are necessary and sufficient for viral genome replication (4) and thus constitute the essential components of the genome replication complex. Further to its requirement in genome replication, NS5A has been shown to play a critical role in virion assembly, as discussed below.

Early in infection, HCV remodels endoplasmic reticulum (ER)-derived membranes to form a "membranous web" (MW) comprised of single, double, and multimembrane vesicles (SMVs, DMVs, and MMVs) that are enriched in viral (e.g., NS3, NS5A, and NS5B) and host cell proteins (5, 6). The MW is proposed as the site of viral genome replication, and in Huh7 cells, the MW is extensively distributed throughout the cytoplasm, correlating with the observed subcellular distribution of NS5A as discrete punctae throughout the cytoplasm. The cellular lipid kinase, phosphatidylinositol kinase type III alpha isoform (PI4KIIIα), is activated by NS5A (7), and the subsequent increase in abundance of phosphatidylinositol-4-phosphate (PI4P) is critical for establishment and maintenance of the MW (8–10). NS5A is also thought to be involved in delivery of nascent virus genomes from the MW to sites of assembly. While the latter are yet to be unambiguously defined, it is accepted that an association of both NS5A and the HCV capsid (core) protein with lipid droplets (LDs; a host organelle responsible for storage of neutral lipids) is required during this process. It has been hypothesized that NS5A switches from a role in replication to an alternative function in assembly which might involve transporting nascent genomes via LDs to assembly sites.

In this regard, NS5A is highly phosphorylated, and it is possible that this reversible posttranslational modification could mediate a switch in NS5A function, by altering protein conformation and/or protein-protein interactions. NS5A comprises three domains and is tethered to membranes by an N-terminal amphipathic helix (Fig. 1b). Domain I is highly structured (11–13), while domains II and III are intrinsically disordered, with elements of transient secondary structure (14). The domains are linked by low-complexity sequences (LCS); LCSI is serine rich, and LCSII is proline rich. To address the potential functional role of NS5A phosphorylation, we and others have used mass spectrometry (MS) to identify phosphorylation sites (15–20). These studies have identified multiple phosphorylation sites, but in particular they show that LCSI is highly phosphorylated. Subsequent mutagenesis of these phosphorylation sites revealed that a subset of them is required for efficient genome replication. In particular, a number of groups have presented evidence that phosphorylation of serine 235 is critical, as replacement of this residue with alanine (S235A) resulted in a 100-fold reduction in genome replication (15–17, 21). It should be noted that other approaches, such as *in vitro* phosphorylation assays, genetic approaches, and use of selective inhibitors, have identified additional sites toward the C terminus of NS5A, such as T360 (22) and S457 (23).

In this study, we focused on another phosphorylated residue within LCSI: serine 225. We demonstrated previously that phosphorylation of serine 225 was required for efficient genome replication and contributed to the subcellular localization of viral proteins during infection (19, 24). Alanine substitution (S225A) resulted in a 10-fold reduction in genome replication and was concomitant with a restricted distribution of NS5A, and other factors known to participate in genome replication (NS3 and PI4P lipids), to a perinuclear region (24). This restriction was dramatic compared to the extensive distribution of these components throughout the cytoplasm in wild-type (WT)-infected cells.

To understand the molecular mechanism underpinning the phenotype of the S225A mutation, we hypothesized that it might be explained by a role of S225 phosphoryla-

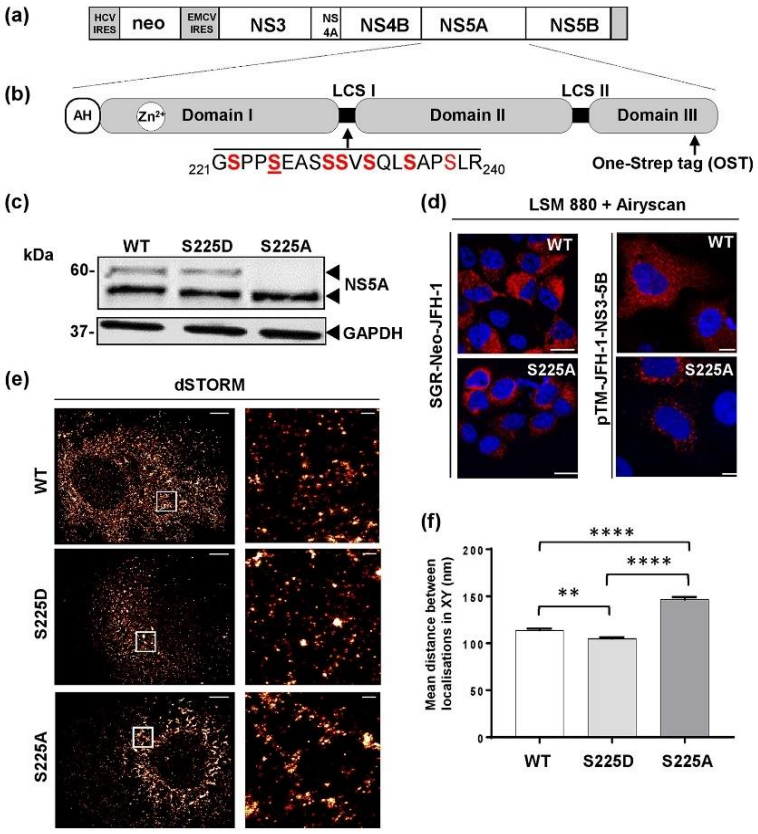


FIG 1 Construction and validation of One-Strep-tagged (OST) NS5A neomycin reporter SGR JFH-1. (a) Schematic of SGR-Neo-JFH-1-5A-OST. (b) NS5A structure showing amino acid residues for LCSI and location of the OST. The OST was introduced between residues 418 and 419, a site previously shown to tolerate insertions (25). AH, amphipathic helix; neo, neomycin phosphotransferase. (c) G418-resistant Huh7 cells stably harboring the SGRs were isolated by selective culture. Cells were lysed and analyzed by Western blotting for NS5A or glyceraldehyde-3-phosphate dehydrogenase (GAPDH). (d) SGR-harboring Huh7 cells (left) or Huh7-Lunet T7 cells transfected with pTM-NS3-5B plasmids (right) were seeded onto coverslips, fixed at 48 h posttransfection (hpt), and permeabilized prior to staining with a sheep polyclonal anti-NS5A serum and Alexa Fluor 594 secondary antibody. Scale bars, 30 μm and 20 μm. (e) Huh7 cells stably harboring wild-type, S225D, or S225A SGRs were fixed, immunostained for NS5A (Alexa Fluor 647-labeled 9E10), and imaged by dSTORM. Images were reconstructed from 873185, 354952, and 641609 localizations for wild-type, S225D, and S225A SGRs, respectively. Histogram bins of 100 nm and 10 nm (with image smoothing) were used for the left and right images, respectively. Scale bars, 5 μm (left) and 500 nm (right). (f) Size of NS5A protein clusters identified by DBSCAN analysis. The mean Euclidean distance between each localization to every other localization in an identified cluster was measured. Data represent the means ± SEM from three independent cells containing 2,062, 1,561, and 1,515 clusters from wild-type, S225D, and S225A SGRs, respectively. Significant differences are indicated as follows: **, $P < 0.005$, and ****, $P < 0.0001$.

tion in regulating interactions between NS5A and cellular proteins. To address this, we used affinity purification of One-Strep-tagged NS5A in conjunction with label-free quantitative proteomics analysis to compare the interactome of wild-type and S225 mutants of NS5A. In this study, we focused on 3 cellular proteins—nucleosome assembly protein 1-like protein 1 (NAP1L1), bridging integrator 1 (Bin1, also known as amphiphysin II), and vesicle-associated membrane protein-associated protein A

(VAP-A)—that exhibited a loss of interaction with S225A mutant NS5A compared to the wild type and the S225D mutant. In contrast, the binding of VAP-B to NS5A was not dependent on S225 phosphorylation and acted as a control. These interactions were validated by immunoprecipitation, immunofluorescence, and proximity ligation assay (PLA). Furthermore, small interfering RNA (siRNA) ablation of endogenous NAP1L1, Bin1, and VAP-A reduced both HCV RNA replication and NS5A expression significantly; however, consistent with lack of dependence on S225 phosphorylation, VAP-B ablation had a modest effect. Importantly, ablation of NAP1L1, Bin1, and VAP-A recapitulated the restricted distribution of the NS5A protein. We propose that S225 phosphorylation is required for the interaction of NS5A with these cellular proteins and enables the formation and distribution of replication complexes throughout the cytoplasm, thereby enhancing the efficiency of genome replication.

RESULTS

Identification of serine 225 phosphorylation-dependent NS5A-interacting proteins. Previously, we reported that phosphorylation of serine 225 within LCSI of NS5A played a role in the regulation of JFH-1 genome replication (24). Mutation of this residue to alanine (S225A; phosphoablatant) resulted in a 10-fold reduction in genome replication and altered subcellular distribution of NS5A, whereas the phosphomimetic mutation (S225D) had no phenotype. To understand the mechanism behind this phenotype, we sought to identify cellular proteins that interacted with NS5A in an S225 phosphorylation-dependent fashion. For this, we exploited the One-Strep tag (OST) affinity purification strategy that we had previously used to identify sites of phosphorylation within NS5A (19). The OST is a peptide that structurally resembles biotin and binds to recombinant streptavidin (Strep-Tactin). S225A and S225D mutations were cloned into pSGR-Neo-JFH-1-5A-OST, which contained the OST cloned into a well-tolerated insertion site near the C terminus of NS5A domain III (25) (Fig. 1a and b). These subgenomic replicons (SGRs) were used to establish stable Huh7 cell lines expressing either the wild-type or the two mutant SGRs. The phenotype of these mutants was confirmed by Western blotting and fluorescence microscopy (Fig. 1c and d). As expected (24), the S225A mutation resulted in a reduction in hyperphosphorylation (Fig. 1c) and in a distribution of the protein that was restricted to the perinuclear region (Fig. 1d). We had previously demonstrated that this phenotype was not a consequence of the reduced level of RNA replication exhibited by the S225A mutant (24). We further confirmed this by expressing either wild-type or S225A mutant NS5A in the context of the NS3-NS5B polyprotein from a T7 RNA polymerase-driven construct (pTM, a kind gift from Volker Lohmann). Following transfection of these plasmids into Huh7-Lunet T7 cells (stably expressing T7 RNA polymerase, also a kind gift from Volker Lohmann), the restricted distribution of NS5A S225A was recapitulated (Fig. 1d), confirming that the phenotype was not dependent on genome replication or the level of NS5A expression. We also applied a superresolution approach (Fig. 1e and f), and this revealed an additional S225A phenotype: discrete clusters of NS5A localizations were equivalent to the diffraction limited puncta observed by wide-field microscopy. Clusters were observed distributed throughout the cytoplasm apart from the S225A mutant, which was more condensed and perinuclear. Consistent with our previous findings (24), larger NS5A clusters were observed for the S225A mutant than for the wild type and the S225D mutant. Taken together, these data demonstrate that S225 phosphorylation regulated not only the distribution of the replication complexes but also their architecture. These observations gave further impetus for the need to understand the molecular mechanism behind this phenotype.

To identify cellular candidates that are potentially involved in the S225A phenotype, we next performed affinity purification from cytoplasmic lysates of Strep-tagged wild-type, S225A, and S225D NS5A and analyzed the bound fractions by mass spectrometry. A large number of known NS5A binding proteins were identified by this procedure. Wild-type NS5A and S225D mutant NS5A showed no difference in their binding to associated proteins (see Data Set S1 in the supplemental material). As we were

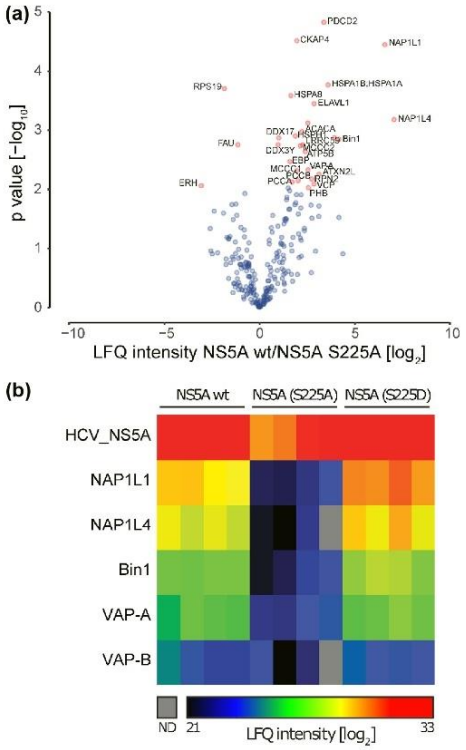


FIG 2 Proteomic analysis of NS5A-interacting proteins. OST-tagged wild-type (wt), S225A, and S225D NS5A proteins were affinity purified from Huh7 cells stably harboring SGRs, and associated proteins were identified by LC-MS/MS. Four independent affinity purifications were performed for each bait. (a) Volcano plot displaying the average degree of enrichment by wt NS5A over S225A NS5A (ratio of label-free quantification [LFQ] protein intensities) and the *P* value (Welch's *t* test) for each protein. Significantly enriched cellular proteins are highlighted in red. (b) Heat map showing nonimputed $\log_2$ transformed LFQ intensities for each individual replicate in rainbow colors (see color scale). Only the bait protein and selected cellular interaction partners are depicted in the plot. Gray color represents missing values (not determined [ND]).

interested in the role of S225 phosphorylation, we mined the data set for those cellular proteins that bound well to wild-type or S225D NS5A but exhibited a reduced level of interaction with S225A mutant NS5A (Fig. 2). We focused on three of these differentially enriched proteins. Bin1 (also known as amphiphysin II) is a BAR domain-containing protein involved in generating membrane curvature and was previously characterized as an NS5A interactor (26–28). VAP-A (also known as hVAP-33) is a vesicle-associated protein previously shown to participate in the formation of HCV genome replication complexes and to bind to NS5A (29–32). As well as these previously characterized NS5A partners, we also focused on the nucleosome assembly protein 1-like proteins 1 and 4 (NAP1L1 and NAP1L4). NAP1L1 has been shown to interact with HCV core and NS5A (33, 34) as well as Kaposi's sarcoma herpesvirus (KSHV) LANA (35) and HIV-1 Tat (36). Although its roles in the nucleus associated with chromatin structure and gene expression are well defined, NAP1L1 is predominantly located in the cytoplasm, where its function(s) remains to be elucidated. We chose these proteins for further analysis

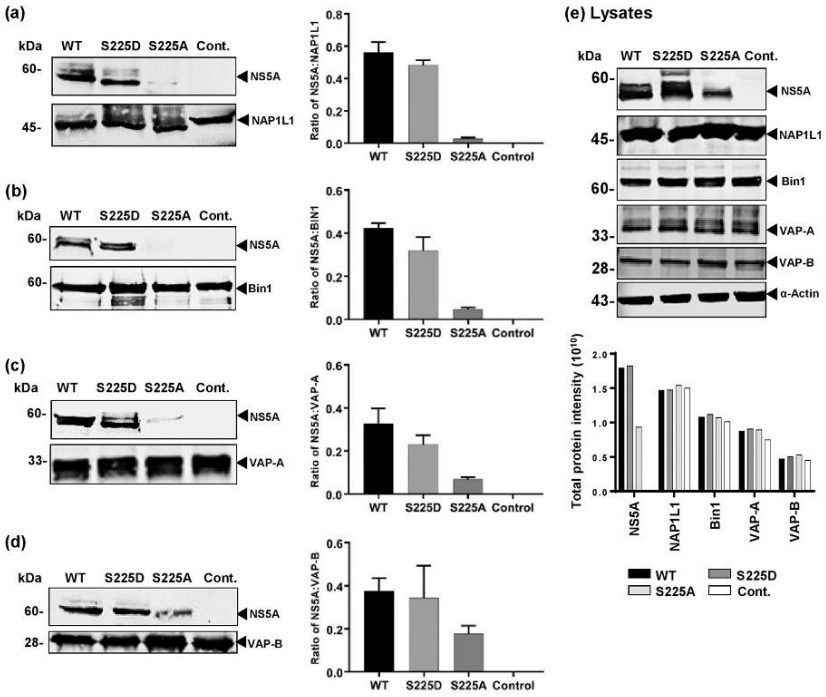


FIG 3 Validation of NS5A-interacting proteins NAP1L1, Bin1, VAP-A, and VAP-B. Huh7 cells stably harboring wild-type (WT), S225D, or S225A SGRs were lysed and immunoprecipitated with antibodies to NAP1L1 (a), Bin1 (b), VAP-A (c), and VAP-B (d) prior to analysis by Western blotting with mouse monoclonal antiserum to NS5A or antibodies to the indicated cellular proteins. The ratios of NS5A to the cellular proteins were quantified from fluorescent Western blots (n = 3). Western blots are representative of three independent experiments. (e) Levels of NS5A and cellular proteins in cytoplasmic lysates were confirmed by Western blotting and quantified (graph). Cont., control.

because they exhibited the greatest difference between wild-type and S225A NS5A (Fig. 2b) and thus represented the best candidates to explain the mechanism underpinning the S225A phenotype. As a control, we also examined VAP-B, which was enriched as an NS5A interactor but did not exhibit S225 phosphorylation dependence.

Validation of S225 phosphorylation-dependent interacting proteins. We next sought to validate the proteomic analysis/mass spectrometry results by a specific immunoprecipitation approach. To this end, lysates from stable Huh7 cell lines harboring either the wild-type or S225A/S225D mutant SGRs, or control parental cells, were immunoprecipitated with antibodies to NAP1L1, Bin1, VAP-A, or VAP-B. The immunoprecipitates were probed with an anti-NS5A antiserum by Western blotting (Fig. 3a to d). The adjacent graphs show quantification of 3 independent assays. Consistent with the proteomic analysis, wild-type and S225D mutant NS5A bound equally to all four cellular proteins; in contrast, S225A NS5A bound weakly, or not at all, to NAP1L1, Bin1, and VAP-A. In addition, binding to VAP-B was confirmed to be less dependent on S225 phosphorylation, as there was only a 50% reduction in the level of S225A NS5A bound to VAP-B. Input levels of NS5A and the four cellular proteins were confirmed by Western blotting of lysates and quantification (Fig. 3e, graph). As seen previously, levels of S225A NS5A were reduced by approximately 50%; however, this did not account for the

reduction in levels of NS5A bound to NAP1L1, Bin1, or VAP-A. As the reduction in expression of S225A correlated with the observed reduction in the amount of NS5A coimmunoprecipitated with VAP-B, we confirm that the NS5A–VAP-B interaction was not dependent on S225 phosphorylation.

In order to further validate the interactions between NS5A and either NAP1L1 or Bin1, we applied two imaging techniques to examine their colocalization and interaction in Huh7 cells stably harboring SGRs. To complement conventional coimmunofluorescence, which can only determine colocalization, we used the proximity ligation assay (PLA), which allows the detection of direct protein-protein interactions in intact cells (37) PLA involves the use of oligonucleotides attached to secondary antibodies which guide the formation of circular DNA strands when bound in close proximity (approximately 5 to 30 nm). These DNA circles then template localized rolling-circle amplification (RCA), allowing individual interacting pairs of proteins to be visualized and enumerated in fixed samples (37, 38). We were able to apply the PLA technique only for NAP1L1 and Bin1, due to a lack of suitable antibody pairs for VAP-A and VAP-B; thus, we also restricted the coimmunofluorescence analysis to these two proteins.

Both NAP1L1 (Fig. 4a) and Bin1 (Fig. 4b) were distributed diffusely throughout the cytoplasm; in the context of both wild-type and S225D NS5A SGR, there was extensive colocalization with NS5A which was particularly noticeable close to the nucleus and not so apparent at the periphery of the cell (quantification shown in lower part of Fig. 4). S225A NS5A showed a different picture: NS5A and either NAP1L1 or Bin1 seemed to occupy mutually exclusive areas within the cytoplasm, and indeed, quantification showed no significant colocalization. As shown in Fig. 5, we observed strong fluorescence signals (red punctae) for both NS5A/NAP1L1 (Fig. 5a) and NS5A/Bin1 (Fig. 5b) in the cytoplasm of Huh7 cells harboring either wild-type or S225D SGR. In contrast, no PLA signal was observed for either NAP1L1 or Bin1 in cells harboring the S225A mutant SGR. These data confirm that NS5A interacts with both NAP1L1 and Bin1 in the cytoplasm of SGR-harboring cells in an S225 phosphorylation-dependent fashion.

S225 phosphorylation-dependent NS5A-interacting proteins are required for efficient viral genome replication. We previously demonstrated that the S225A mutation in the context of either an SGR or infectious virus resulted in a 1-log reduction in genome replication (19, 24). Because this mutation also disrupted binding of NS5A to NAP1L1, Bin1, VAP-A, and, to a lesser extent, VAP-B, we hypothesized that these cellular proteins might play a role in genome replication. To test this, we adopted an siRNA approach to ablate expression of each of these cellular proteins in Huh7 cells stably harboring a wild-type JFH-1 SGR. Cells were transfected with siRNA pools targeting different sites of NAP1L1, Bin1, VAP-A, or VAP-B or with a control nontargeting siRNA and harvested at 72 hours posttransfection (hpt), and protein expression levels were determined by Western blot analysis. As shown in Fig. 6a, ablation of NAP1L1 and Bin1 expression was efficient and resulted in a concomitant reduction in NS5A protein levels (Fig. 6b). The siRNAs for VAP-A and VAP-B were less efficient; however, VAP-A ablation did reduce NS5A expression, although not as effectively as NAP1L1 or Bin1. In contrast, we observed no significant effect of the VAP-B ablation on NS5A levels, possibly because, as reported previously, the role of VAP-B in HCV genome replication is mediated via interactions with both NS5B and NS5A (39). As NS5A expression levels are an indirect measure of genome replication, we also directly assessed the levels of SGR RNA in siRNA-transfected cells by real-time quantitative PCR (qRT-PCR) (Fig. 6c). Consistent with the effects of NAP1L1, Bin1, and VAP-A ablation, we observed reductions of between 100- and 1,000-fold in HCV-specific RNA levels compared to the levels in cells transfected with the control siRNA. Interestingly, despite the fact that VAP-B ablation had no effect on NS5A levels (Fig. 6b), it also significantly reduced HCV RNA levels, although not as effectively as the other 3 targets. Lastly, we confirmed that the ablation of NAP1L1, Bin1, VAP-A, or VAP-B expression had no effect on levels of S225A NS5A (Fig. 6d). These data confirm that the S225 phosphorylation-dependent interacting proteins NAP1L1, Bin1, and VAP-A (and, to a lesser extent, VAP-B) are involved in HCV genome replication.

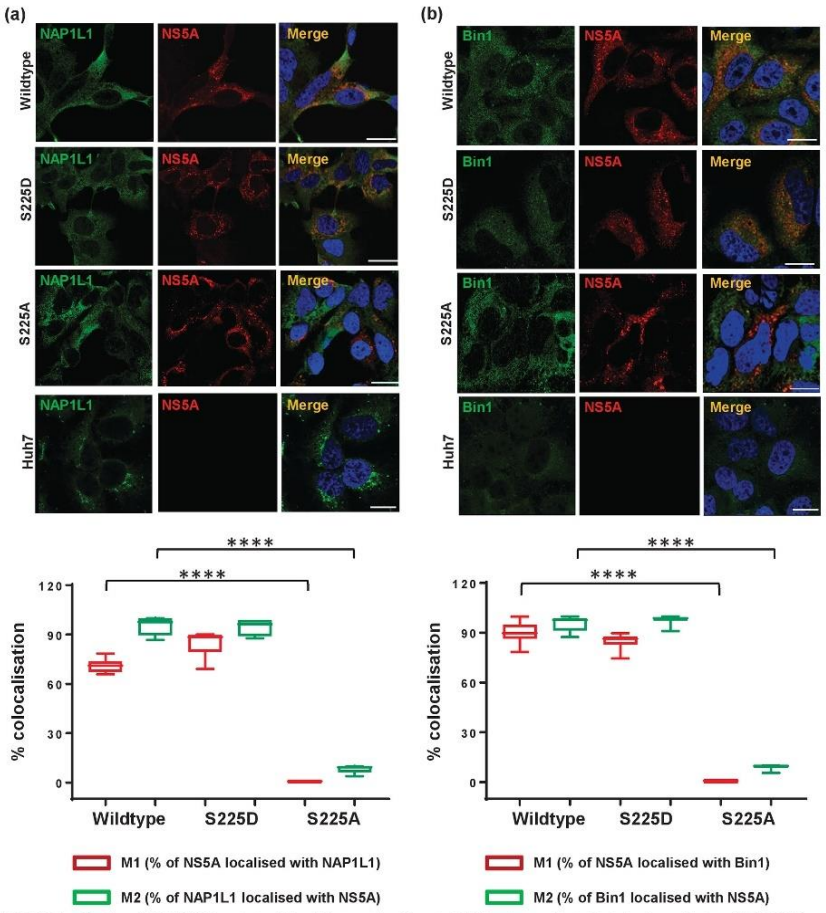


FIG 4 Colocalization of NS5A S225 mutants with cellular proteins. Parental Huh7 cells or cells stably harboring wild-type (WT), S225D, or S225A SGRs were seeded onto coverslips and incubated for 72 h prior to fixation. Cells were permeabilized and immunostained for NS5A (red) and either NAP1L1 (a, green), or Bin1 (b, green). Scale bar, 20 μm. Box-and-whisker plots show the percent colocalization between NS5A and either NAP1L1 or Bin1. Manders' overlap coefficients M1 and M2 were used to evaluate the degree of colocalization between two fluorescent labels (red and green) in 10 cells. ****, $P < 0.0001$.

The subcellular distribution of NS5A is regulated by S225 phosphorylation-dependent binding to NAP1L1, Bin1, and VAP-A. As we had shown previously (Fig. 1 and reference 24), the S225A mutant NS5A exhibited a perinuclear restricted distribution. We therefore considered that the loss of interaction with cellular factors might be, at least in part, responsible for this restricted distribution. To test this hypothesis, we again used siRNA to ablate expression of NAP1L1, Bin1, VAP-A, or VAP-B in Huh7 cells stably harboring a wild-type SGR. Cells transfected with siRNA were analyzed by immunofluorescence with antibodies to both NS5A and the four cellular proteins. As shown in Fig. 7a, NAP1L1 ablation led to an overall reduction in NS5A expression which was restricted to a perinuclear distribution compared with the negative-control siRNA-

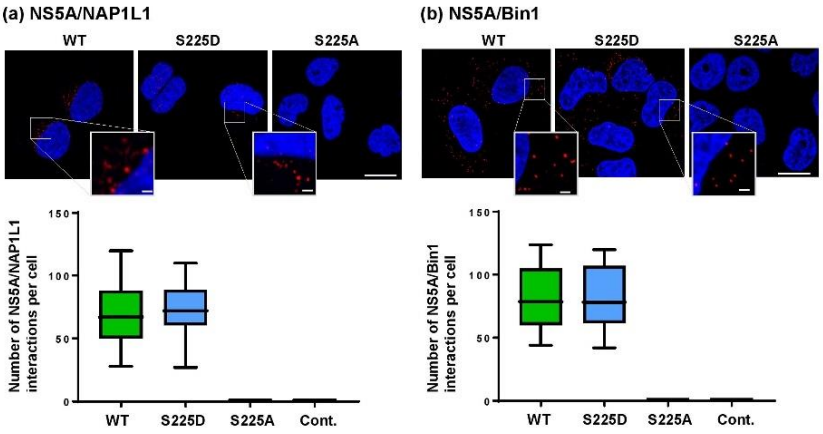


FIG 5 Proximity ligation assay for NS5A interactions with NAP1L1 and Bin1. Huh7 cells stably harboring wild-type, S225D, or S225A SGRs were seeded onto coverslips and incubated for 72 h prior to fixation and immunostaining for NS5A and either NAP1L1 or Bin1, followed by PLA probe MINUS and PLUS (Sigma-Aldrich) ligation and rolling-circle amplification (RCA). The hybridization probes were labeled with Alexa Fluor 555 (red), and the nuclei were counterstained with DAPI (blue). Scale bars, 20 μm and 2 μm. For each experiment, the PLA signals (per cell) of each sample were counted using Imaris software version 7.4 (Bitplane AG) from a minimum of 20 cells.

transfected (scrambled) or untransfected (mock) cells. To provide quantitative confirmation of these observations, we determined NS5A spatial distribution data for 12 randomly selected cells (Fig. 7a, graph). A similar pattern was observed when Bin1 (Fig. 7b) or VAP-A (Fig. 8a) expression was ablated. In contrast, ablation of VAP-B expression had only a modest effect on the distribution of NS5A which did not reach significance upon quantitative analysis (Fig. 8b). Lastly, we sought to confirm that the effects of NAP1L1, Bin1, or VAP-A ablation on NS5A distribution were not due to a reduction in levels of expression, as an indirect consequence of the inhibition of genome replication. To test this, Huh7-Lunet T7 cells were transfected with pTM-NS3-5B followed by NAP1L1, Bin1, or VAP-A siRNAs. As shown in Fig. 9, the restricted distribution of NS5A was also observed using this expression system, confirming that the phenotype was not dependent on genome replication or the level of NS5A expression. These data demonstrate that S225 phosphorylation is required for the interaction of NS5A with a number of cellular proteins and that these interactions control both the distribution of NS5A and HCV genome replication.

DISCUSSION

In this study, we used a proteomic approach to identify cellular proteins that interacted with NS5A in an S225 phosphorylation-dependent fashion, using affinity purification followed by liquid chromatography-tandem mass spectrometry (LC-MS/MS). We focused our attention on proteins that were enriched in samples derived from OST-tagged wild-type NS5A and that were significantly less abundant in samples from OST-tagged NS5A (S225A).

The most intriguing S225 phosphorylation-dependent interacting proteins were NAP1L1 and NAP1L4. These are members of the nucleosome assembly protein 1 (NAP1) family and, as their name suggests, are nuclear proteins primarily involved in chromatin assembly. However, they are also located in the cytoplasm, particularly during G_2, and mediate nucleocytoplasmic shuttling of histones during cell cycle progression (40). Of the five mammalian NAP1 family members (NAP1L1 to NAP1L5), only NAP1L1 and NAP1L4 are ubiquitously expressed, while the other three are expressed mainly in neurons (41). NAPs have been shown to interact with proteins of a number of DNA

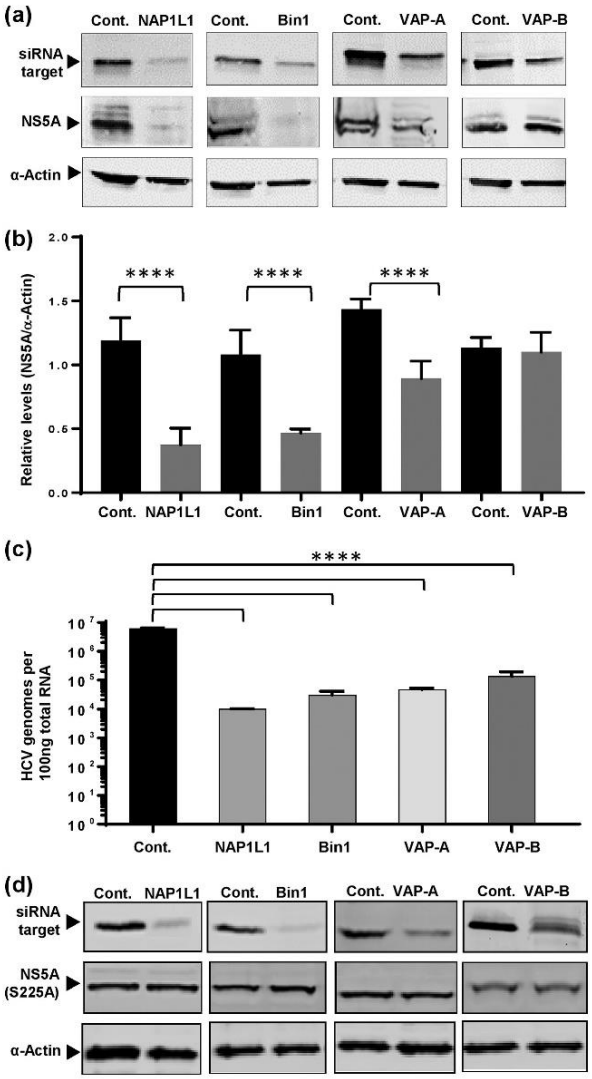


FIG 6 Effects of siRNA ablation of NS5A-interacting protein expression on NS5A expression and RNA replication. (a) Huh7 cells stably harboring a wild-type SGR were transfected with pooled siRNA targeting NAP1L1, Bin1, VAP-A, or VAP-B or a control scrambled siRNA (Santa Cruz) at a final concentration of 10 nM. Transfected cells were incubated in medium lacking G418, lysed at 72 hpt, and analyzed by Western blotting. (b) Total NS5A levels were quantified from fluorescent Western blots ($n = 3$). The Western blots are representative of three independent experiments. ****, $P < 0.0001$ compared to the value for the control. (c) At 72 hpt, the cells were harvested in TRIzol, total RNA was extracted, and qRT-PCR was conducted on 100 ng of total cellular RNA using 5′-UTR TaqMan primers (53). The data are from three independent experiments. (d) Huh7 cells stably harboring the S225A mutant SGR were transfected with indicated siRNA and analyzed by Western blotting at 72 hpt as described for panel a.

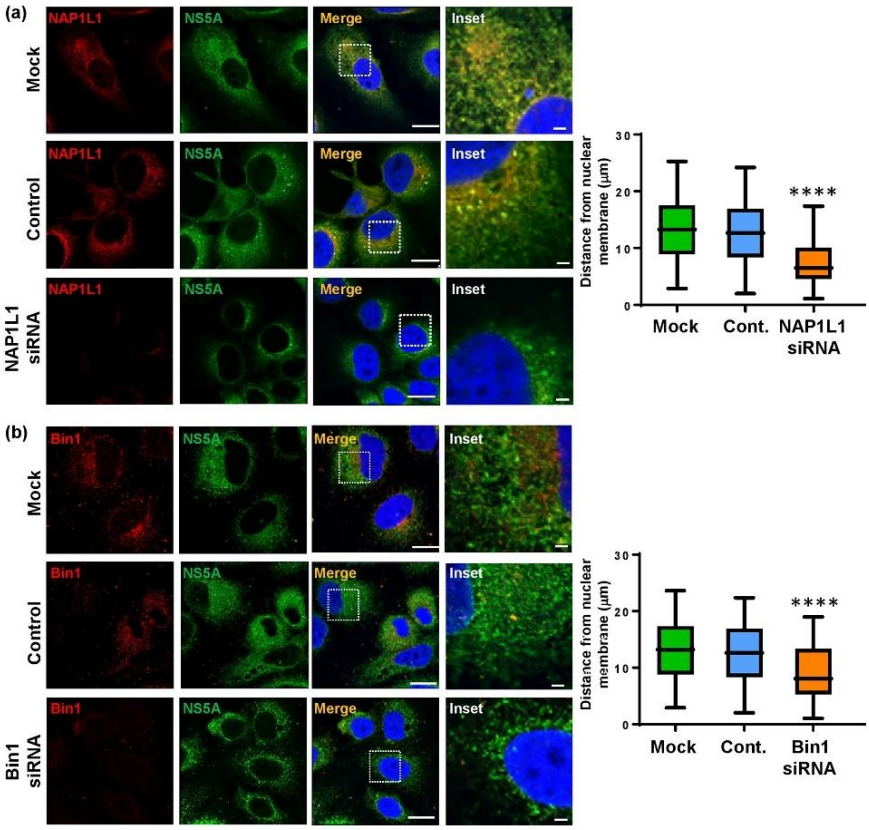


FIG 7 Effects of siRNA ablation of NAP1L1 or Bin1 on the distribution of NS5A. Huh7 cells stably harboring a wild-type SGR were left untransfected or were transfected with pooled siRNA targeting NAP1L1 or Bin1 or a control scrambled siRNA (Santa Cruz) at a final concentration of 10 nM. At 72 hpt, cells were permeabilized, immunostained for NS5A (green), NAP1L1 (a, red), or Bin1 (b, red), and imaged by confocal microscopy. Scale bars, 20 μm and 5 μm. Spatial data for NS5A were determined from 12 cells for each assay using the ImageJ software package. The data are representative of three independent experiments. ****, $P < 0.0001$ compared to the control value.

viruses, including papillomavirus E2 (42), Epstein-Barr virus EBNA1 (43), and KSHV LANA (35). NAP1L1 also binds to HIV-1 Tat and Rev, and these interactions regulate viral infectivity and the subcellular localization of the virus proteins (36, 44). Interestingly, both NAP1L1 and NAP1L4 were previously shown to interact with both the HCV core protein (33) and NS5A (34) when the viral proteins were expressed individually in HEK293 cells, although in neither study were these interactions validated. In contrast, our data reveal that in the context of the NS3-NS5B replication complex within Huh7 cells, both NAP1L1 and NAP1L4 interact with NS5A. Focusing on NAP1L1 (due to a lack of available immunological reagents for NAP1L4), we further showed that these interactions contribute to virus genome replication and the distribution of replication complexes throughout the cytoplasm. It will be intriguing to understand the precise role of these proteins in HCV genome replication; in this regard, NAP1 has been shown to oligomerize (45) and NAPs function to transport histones, two attributes that may be important for their function during the HCV life cycle.

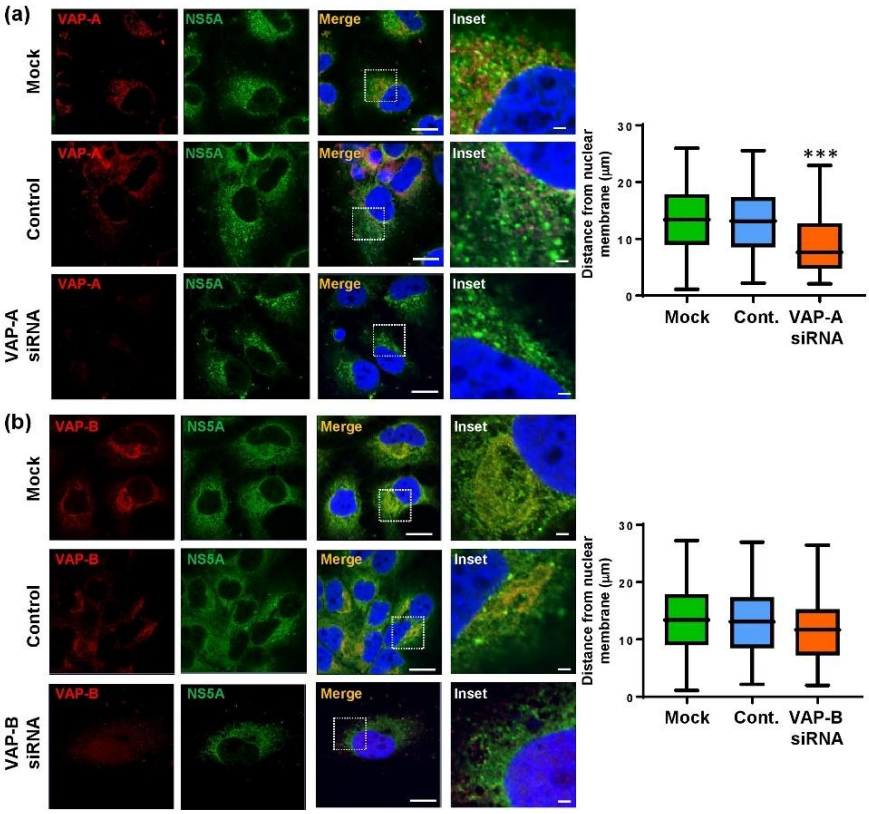


FIG 8 Effects of siRNA ablation of VAP-A or VAP-B on the distribution of NS5A. Huh7 cells stably harboring a wild-type SGR were left untransfected or were transfected with pooled siRNA targeting VAP-A or VAP-B or a control scrambled siRNA (Santa Cruz) at a final concentration of 10 nM. At 72 hpt, cells were permeabilized, immunostained for NS5A (green), VAP-A (a, red), or VAP-B (b, red) and imaged by confocal microscopy. Scale bars, 20 μm and 5 μm. Spatial data for NS5A were determined from 12 cells for each assay using the ImageJ software package. The data are representative of those from three independent experiments. ***, $P < 0.001$ compared to the control value.

We also identified Bin1 as an S225 phosphorylation-dependent NS5A interactor. Bin1 contains an N-terminal BAR domain, a C-terminal SH3 domain, and a central clathrin and adaptor binding domain (14, 26). Like other BAR domain-containing proteins, Bin1 is involved in membrane curvature and trafficking. It had previously been shown to interact via its SH3 domain with a polyproline motif (PxxPxR) located in NS5A LCSII (26, 27), and it also interacts with similar motifs in other viral proteins, such as alphavirus nsP3 (46). As LCSII is distal to LCSI, this suggests that S225 phosphorylation might effect a conformational change in NS5A that occludes the SH3 domain binding polyproline motif; alternatively, multiple contacts between the two proteins may be required to establish a stable interaction. The physiological role of Bin1 in regulating membrane curvature is consistent with its recruitment by NS5A to sites of replication where it may be involved in the rearrangements necessary to form the MW. As it is also involved in membrane trafficking, this provides a potential explanation for the perinuclear restriction of NS5A S225A replication complexes.

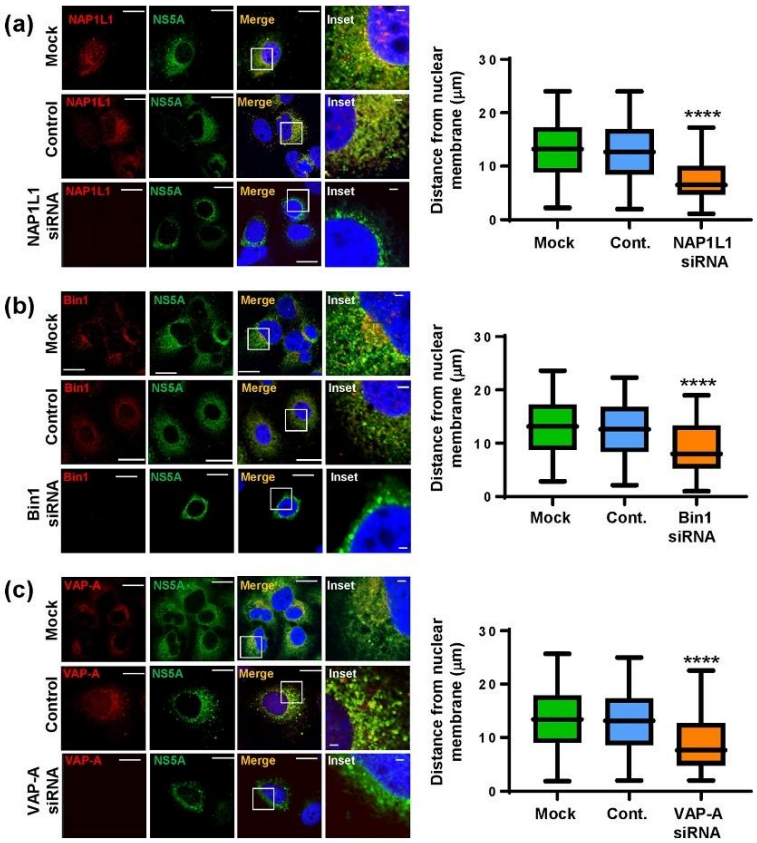


FIG 9 Effects of siRNA ablation of NAP1L1, Bin1, or VAP-A on the distribution of NS5A in the absence of genome replication. Huh7-Lunet T7 cells were left untransfected or were transfected with pooled siRNA targeting NAP1L1, Bin1, or VAP-A or a control scrambled siRNA (Santa Cruz) at a final concentration of 10 nM. At 48 hpt, cells were transfected with pTM-NS3-5B plasmid. After a further 24 h, cells were permeabilized, immunostained for NS5A (green), NAP1L1 (a, red), Bin1 (b, red), or VAP-A (c, red), and imaged by confocal microscopy. Scale bars, 20 μm and 5 μm. Spatial data for NS5A were determined from 20 cells for each assay using the ImageJ software package. ****, $P < 0.0001$ compared to the control value.

Both VAP-A and VAP-B have previously been shown to bind both NS5A and NS5B (30, 39). Interestingly, in the context of the genotype 1b Con1 isolate, residues T2185 and K2187 (corresponding to NS5A residues 215 and 217) were shown to be required for the interaction (32). Mutations of these residues abolished the interaction with VAP-A, although this was demonstrated using the yeast two-hybrid system. An additional culture-adaptive mutation (S2201I, corresponding to S229), could rescue the interaction, although this mutation resulted in a loss of hyperphosphorylation. Together with our data, these results suggest that LCSI and proximal sequences at the C terminus of domain I are required for the interaction with VAP-A. Furthermore, it was recently reported that the NS5A interaction with VAP-A was indirect and required a bridging protein, G protein pathway suppressor 2 (GPS2) (29). VAP-A and VAP-B are integral membrane proteins that have a range of reported functions pertaining to

vesicle trafficking and lipid transport (30, 47). VAP-A also interacts with the oxysterol binding protein (OSBP), which is recruited to the MW along with PI4K (48), and is reported to regulate cholesterol transport to the MW. Although OSBP was not identified in our proteomic study, these observations are consistent with a role for S225 phosphorylation in regulating MW formation and distribution via an interaction with VAP-A.

In conclusion, we propose that phosphorylation of NS5A at S225 within LCSI contributes to efficient genome replication by regulating interactions with key cellular factors, some of which we have identified and characterized in this study. It is important to note that multiple phosphorylation sites have been identified within LCSI, and it is conceivable therefore that the total pool of NS5A within a cell might contain many distinct phosphorylated species, each of which could also have a different function(s). Identifying these distinct NS5A species is a technical challenge which is limited by the sensitivity of mass spectrometric methods. In this context, we were previously able to unambiguously identify by mass spectrometry an NS5A species that was phosphorylated within LCSI only on S222 and S225 (18). The abundance of this species is consistent with a specific role during virus genome replication. Given the likely interdependence of different phosphorylation events, exemplified by recent evidence for hierarchical phosphorylation, i.e., the requirement for S235 to be phosphorylated prior to S238 phosphorylation (16), it will also be challenging to understand the functions of these different NS5A species. A complete understanding of either the complexity of NS5A phosphorylation or the functional consequences remains a distant and aspirational objective. By demonstrating the role of phosphorylation in regulating NS5A-host protein interactions, we believe that this study makes a significant contribution to this objective; however, it provides only a few pieces of the jigsaw puzzle! There is much yet to be discovered.

MATERIALS AND METHODS

HCV replicon constructs. A DNA construct of the neomycin phosphotransferase (Neo) containing subgenomic replicon pSGR-Neo-JFH-1-5A-OST, in which the One-Strep tag (OST) was introduced into the C terminus of NS5A domain III (25), was used in this study (Fig. 1a and b). Previously, the OST has been shown to have no effect on virus genome replication (25). Wild-type pSGR-Neo(*Rsr*II^{del})-JFH-1-5A-OST was first constructed by abolishing an RsrII site within the Neo gene using a Q5 site-directed mutagenesis kit (New England BioLabs [NEB]). Then the desired NS5A mutants from SGR-luc-JFH-1 (24) were cloned into pSGR-Neo(*Rsr*II^{del})-JFH-1-5A-OST via flanking NsiI/RsrII, and correct insertion was confirmed by sequencing.

Cell culture. Huh7 cells were maintained in Dulbecco's modified Eagle's medium (DMEM; Sigma-Aldrich) supplemented with 10% fetal bovine serum (FBS), 100 IU of penicillin/ml, 100 μg of streptomycin/ml, and 1% nonessential amino acids (NEAA) in a humidified incubator at 37°C with 5% CO_2. Huh7 cells carrying a subgenomic JFH-1 replicon (SGR-Neo-JFH-1) were maintained in the same medium supplemented with 300 μg/ml of G418 (BioPioneer). Huh7-Lunet T7 cells (expressing T7 RNA polymerase, a kind gift from Volker Lohmann) were maintained in the same medium supplemented with 5 μg/ml of Zeocin (Life Technologies).

Electroporation of replicon RNA and generation of stable cell lines. The preparation of *in vitro* transcripts and electroporations for pSGR-Neo-JFH-1-5A-OST-derived constructs were conducted as described previously (25). In brief, 2×10^6 Huh7 cells in diethyl pyrocarbonate (DEPC)–phosphate-buffered saline (PBS) were electroporated with 3 μg of *in vitro* RNA transcripts using a square-wave protocol at 260 V for 25 ms. Subsequently, cells were resuspended in complete DMEM and seeded at a culture area of 1×10^4 cells/cm^2 into 10-cm dishes. Forty-eight hours postelectroporation, the cells were selected with 300 μg/ml of G418. After 2 weeks, emerging colonies were pooled and kept under continuous G418 selection for 1 to 2 weeks in order to establish stable SGR-harboring cell lines. The maintenance of the S225 mutations was verified by RT-PCR and sequencing.

Affinity purification coupled to quantitative LC-MS/MS proteomics. To detect proteins bound to OST-tagged NS5A by affinity purification and mass spectrometry, cell pellets from Huh7 cells stably harboring NS5A wild-type (WT) and S225A and S225D mutant SGRs were prepared by snap-freezing cells in liquid nitrogen. Cell pellets were lysed in TAP lysis buffer (50 mM Tris-HCl [pH 7.5], 100 mM NaCl, 5% [vol/vol] glycerol, 0.2% [vol/vol] Nonidet P-40, 1.5 mM $MgCl_2$) in the presence of protease inhibitor cocktail (EDTA free, cOmplete; Roche), phosphatase inhibitor (PhosSTOP; Roche), and 750 U of Benzonase (Core Facility, Max Planck Institute of Biochemistry [MPI-B]) for 30 min on ice. After incubation on ice, cell lysates were sonicated using a Bioruptor with 15 alternating 30-s on/off cycles and clarified by centrifugation at 16,000 $\times$ *g*. Subsequently, Strep-Tactin Sepharose beads (Iba) were incubated with 2 mg of protein of clarified lysate for 60 min at 4°C and washed once with TAP lysis buffer and three times with TAP wash buffer (50 mM Tris-HCl [pH 7.5], 100 mM NaCl, 5% [vol/vol] glycerol, 1.5 mM $MgCl_2$) lacking Nonidet P-40 to remove residual detergent. Sample preparation and LC-MS/MS analysis were performed as described previously (49). Briefly, four independent affinity purifications were performed for each bait,

samples were sequentially digested with LysC (Wako Chemicals, USA) and trypsin (Promega), acidified with 0.1% trifluoroacetic acid (TFA), desalted with C18 stage tips, and analyzed by liquid chromatography coupled to mass spectrometry on an Orbitrap XL instrument (Thermo Fisher Scientific). Mass spectrometry raw files were processed with MaxQuant software version 1.5.5.1 (50) using the built-in Andromeda engine to search against human and JFH-1 proteomes containing forward and reverse sequences. Additionally, the label-free quantification (LFQ) (51) algorithm and Match Between Runs option were used. Perseus software version 1.5.5.1 was used to further process the data. In this manner, only proteins identified on the basis of at least two peptides and a minimum of three quantification events in at least one experimental group were considered. LFQ protein intensity values were log transformed and missing values filled by imputation. Significantly enriched proteins were determined by Welch's *t* test with permutation-based false-discovery rate (FDR) statistics, performing 250 permutations. The FDR threshold was set to 0.01 and S0 parameter was set to 0.1 to separate background from specifically enriched proteins. Results were plotted using R (https://www.R-project.org) and visually adapted using Adobe Illustrator.

SDS-PAGE and Western blotting. Cells were washed twice in PBS, lysed in 1× Glasgow lysis buffer (GLB; 1% [vol/vol] Triton X-100, 120 mM KCl, 30 mM NaCl, 5 mM $MgCl_2$, 10% [vol/vol] glycerol, and 10 mM PIPES-NaOH, [pH 7.2], with protease and phosphatase inhibitors) and harvested by centrifugation (2,800 × *g*, 10 min, and 4°C) before determination and normalization of protein concentration by bicinchoninic acid (BCA) assay (Pierce). Following separation by SDS-PAGE, proteins were transferred to a polyvinylidene difluoride (PVDF) membrane and blocked in 50% (vol/vol) Odyssey blocking (OB) buffer (LI-COR) in Tris-buffered saline (TBS). The membrane was incubated with primary antibodies overnight at 4°C, followed by secondary antibodies for 2 h at room temperature, both prepared in 25% OB buffer. Primary antibodies used were anti-NS5A (sheep, prepared in-house) at 1:4,000 (52), anti-NAP1L1 (rabbit; Santa Cruz) at 1:350, anti-Bin1 (rabbit; Generon) at 1:300, anti-VAP-A (rabbit; Generon) at 1:1,000, anti-VAP-B (rabbit; Generon) at 1:1,000, and anti-α-actin (mouse; Sigma) at 1:10,000. Secondary antibodies were anti-rabbit, anti-sheep (800 nm), or anti-mouse (700 nm) antibodies, used at 1:10,000 prior to imaging using a LI-COR Odyssey Sa infrared imaging system. Quantification of Western blots was carried out using Image Studio v3.1 (LI-COR) using a background subtraction method.

PLA. Cells were washed with PBS before fixation for 15 min at room temperature in 4% (wt/vol) paraformaldehyde (PFA); cells were subsequently permeabilized in 0.1% (vol/vol) Triton X-100–PBS and blocked with PBS-Tween (PBS-T) and 5% (wt/vol) bovine serum albumin (BSA) before immunostaining for anti-NS5A (mouse monoclonal, 1:1,000) and either anti-NAP1L1 (rabbit monoclonal, 1:100) or anti-Bin1 (rabbit monoclonal, 1:50) overnight at 4°C. Coverslips were washed 3 times for 5 min in PBS-T buffer under gentle shaking and incubated with proximity ligation assay (PLA) probes Duolink In Situ PLA Probe Anti-Mouse PLUS (DUO92001; Sigma-Aldrich) and Duolink In Situ PLA Probe Anti-Rabbit MINUS (DUO92005) for 2 h at 37°C. For PLA, all incubations were performed in a preheated humidity chamber and according the manufacturer's recommendations using a Duolink In Situ Detection Reagents Red kit (DUO92008). Coverslips were washed 3 times for 5 min in PBS-T buffer under gentle shaking and incubated with a DNA ligase previously diluted in ligation buffer for 30 min at 37°C. Coverslips were washed 3 times for 5 min in PBS-T buffer under gentle shaking and incubated with a DNA polymerase previously diluted in amplification buffer for 90 min at 37°C. Finally, coverslips were washed for 20 min under gentle shaking and then washed for 2 min with PBS and air dried. Coverslips were mounted with Duolink In Situ mounting medium with 4′,6-diamidino-2-phenylindole (DAPI), and fluorescence was visualized with a Zeiss LSM880 upright microscope.

siRNA ablation. SGR-harboring Huh7 cells or Huh7-Lunet T7 cells were transfected with 10 nM pooled siRNA (Santa Cruz) or 10 nM AllStars negative-control siRNA (Qiagen) using Lipofectamine 2000 (Invitrogen) according to the manufacturer's protocol. sc-75871 targets NAP1L1, sc-29804 targets Bin1, sc-61768 targets VAP-A, and sc-61770 targets VAP-B. For Huh7-Lunet T7 cell experiments, the cells were transfected with pTM-NS3-5B plasmid at 48 h posttransfection (hpt). Cells were incubated in DMEM supplemented with 5% fetal calf serum (FCS) and harvested at 72 hpt. Proteins and total RNA were isolated with TRIzol (Life Technologies), and subsequent Western blotting and real-time quantitative PCR (qRT-PCR) were performed.

RNA extraction and qRT-PCR. To quantify the number of HCV genomes, total cell RNA was extracted using TRIzol reagent by following the manufacturer's instructions (Invitrogen). Total extracted cellular RNA (100 ng) was analyzed using a one-step qRT-PCR TaqMan-based kit (Eurogentec), with primers and probe designed against the 5′ untranslated region (UTR) as described previously (19, 53).

Immunofluorescence and confocal microscopy. Cells were washed with PBS before fixation for 20 min in 4% (wt/vol) PFA; cells were subsequently permeabilized in 0.1% (vol/vol) Triton X-100 and PBS and blocked with PBS-T and 5% (wt/vol) BSA before being immunostained with primary antibody as described above. Various fluorescently conjugated secondary antibodies were used at 1:500 (Life Technology). Nuclei were counterstained with DAPI. Confocal microscopy images were acquired on a Zeiss LSM880 upright microscope with Airyscan; postacquisition analysis was conducted using Zen software (Zen version 2015 black edition 2.3; Zeiss) or Fiji (version 1.49) software (54).

dSTORM. The direct stochastic optical reconstruction microscopy (dSTORM) system described previously (55) was modified with a cylindrical lens ($f = 150$ mm, where f is focal length; Thorlabs). Round glass 25-mm-diameter coverslips (Warner Instruments) were cleaned in a 1:1:5 solution of NH_3 (aq), H_2O_2, and H_2O at 80°C for 16 h. Cells stably harboring NS5A wild-type, S225D, and S225A mutant subgenomic replicons were seeded onto cleaned coverslips at 1×10^5 cells per well in six-well plates. Coverslips were fixed in 2% (wt/vol) PFA in normal medium for 10 min, permeabilized in 0.2% Triton X-100 in PBS for 10

min, and blocked in 1% (vol/vol) normal donkey serum (Sigma-Aldrich) for 1 h. Mouse monoclonal anti-NS5A antibody was directly labeled at an approximately 1:1 ratio using the carboxylic acid succinimidyl ester of the photoswitchable dye Alexa Fluor 647 (Life Technologies) in PBS containing 125 mM $NaHCO_3$ in the dark for 30 min. Unincorporated dye was removed by size exclusion using Zeba microspin desalting columns (Thermo Fisher Scientific). Immunostaining was conducted with directly labeled antibody (1:2,000) for 1 h. Coverslips were then treated with 0.01% poly-L-lysine (Sigma-Aldrich) for 10 min and incubated with a suspension of 150-nm gold nanoparticles (Sigma-Aldrich).

The image acquisition and processing software was used as described previously (56). Labels were stochastically activated with 642-nm laser excitation under wide-field illumination in the presence of fluorescence quenching buffer (glucose oxidase [10 U], catalase [50 U], 12.5 mg ml^{-1} of D-glucose, 1 mM 2-mercaptoethylamine in PBS [pH 8.0]). Data sets consisted of 11,000 image frames at a frame rate of 20 Hz. Localized emission events were binned into histograms for display and correction of image distortion by the cylindrical lens. Image smoothing was conducted in R using kernel density estimation to reflect *x-y-z* localization precisions measured from fiducial markers, and *z*-stacks were visualized in Fiji. Clustering analysis was conducted in Python using density-based spatial clustering of applications with noise (DBSCAN) (57) on the localization coordinates extracted from the palm3d software (MinPts = 30; ε = 150 nm). Cluster sizes were determined from the mean Euclidean distance between all localizations in identified clusters. Image analysis in R and Python used custom scripts (available on request), and statistical tests were conducted in GraphPad Prism using a one-way analysis of variance (ANOVA) with Tukey's multiple comparisons.

Quantification of NS5A distribution. For quantification of NS5A spatial arrangement, images were acquired with the same acquisition parameters, but with variable gain to ensure correct exposure. The spatial coordinates of NS5A were determined using the FindFoci function of the GDSC plugin for Fiji, with the nuclear envelope being manually outlined (utilizing the DAPI staining as a reference) and coordinates generated by Fiji. The distance from each NS5A to the nuclear envelope was then determined using trigonometry. NS5A spatial distribution data were generated for 12 randomly selected cells for each replicon variant and data combined into a box-and-whisker plot.

For colocalization analysis, Manders' overlap coefficient was calculated using ImageJ software with Just Another Co-localization Plugin (JACoP) (National Institutes of Health). Coefficient M1 reflects the fraction of the anti-NS5A signal that overlaps either the anti-NAP1L1 or anti-Bin1 signal. Coefficient M2 reflects the fraction of either the anti-NAP1L1 or anti-Bin1 signal that overlaps the anti-NS5A signal. Coefficient values range from 0 to 1, corresponding to nonoverlapping images and 100% colocalization images, respectively. Colocalization calculations were performed on >10 cells from at least two independent experiments.

Statistics. Data sets were analyzed using Student's *t* test assuming a two-tailed, unequal variance to determine statistical difference from the wild type (WT) (n = 3 or greater throughout).

SUPPLEMENTAL MATERIAL

Supplemental material for this article may be found at https://doi.org/10.1128/JVI.00805-17.

SUPPLEMENTAL FILE 1, XLSX file, 0.1 MB.

ACKNOWLEDGMENTS

We thank Takaji Wakita (National Institute for Infectious Diseases, Tokyo) for pSGR-Neo-JFH-1, Timothy Tellinghuisen (Scripps, Florida) for the NS5A monoclonal antibody, Volker Lohmann (Heidelberg) for the pTM-NS3-5B (JFH-1) plasmid and Huh7-Lunet T7 cells, and Douglas Ross-Thriepland (Astra-Zeneca) for the original pSGR-Neo-JFH-1-5A (OST) construct.

This work was funded by a Wellcome Trust Investigator Award to M.H. (grant number 096670). A.P. was funded by an ERC starting grant (iVIP; 311339), the Federal Ministry for Education and Research (ERA-Net grant ERASe), and the German Research Foundation (TRR179). The dSTORM microscope was funded by alumnus M. Beverly, in support of the University of Leeds Making a World of Difference campaign.

REFERENCES

1. Webster DP, Klenerman P, Dusheiko GM. 2015. Hepatitis C. Lancet 385:1124–1135. https://doi.org/10.1016/S0140-6736(14)62401-6.
2. Pawlotsky JM. 2016. Hepatitis C virus resistance to direct-acting antiviral drugs in interferon-free regimens. Gastroenterology 151:70–86. https://doi.org/10.1053/j.gastro.2016.04.003.
3. Tellinghuisen TL, Evans MJ, von Hahn T, You S, Rice CM. 2007. Studying hepatitis C virus: making the best of a bad virus. J Virol 81:8853–8867. https://doi.org/10.1128/JVI.00753-07.
4. Lohmann V, Korner F, Koch J, Herian U, Theilmann L, Bartenschlager R. 1999. Replication of subgenomic hepatitis C virus RNAs in a hepatoma cell line. Science 285:110–113. https://doi.org/10.1126/science.285.5424.110.
5. Paul D, Hoppe S, Saher G, Krijnse-Locker J, Bartenschlager R. 2013. Morphological and biochemical characterization of the membranous hepatitis C virus replication compartment. J Virol 87:10612–10627. https://doi.org/10.1128/JVI.01370-13.
6. Romero-Brey I, Merz A, Chiramel A, Lee JY, Chlanda P, Haselman U, Santarella-Mellwig R, Habermann A, Hoppe S, Kallis S, Walther P, Antony C, Krijnse-Locker J, Bartenschlager R. 2012. Three-dimensional architecture and biogenesis of membrane structures associated with hepatitis C

virus replication. PLoS Pathog 8:e1003056. https://doi.org/10.1371/journal.ppat.1003056.
7. Reiss S, Harak C, Romero-Brey I, Radujkovic D, Klein R, Ruggieri A, Rebhan I, Bartenschlager R, Lohmann V. 2013. The lipid kinase phosphatidylinositol-4 kinase III alpha regulates the phosphorylation status of hepatitis C virus NS5A. PLoS Pathog 9:e1003359. https://doi.org/10.1371/journal.ppat.1003359.
8. Berger KL, Kelly SM, Jordan TX, Tartell MA, Randall G. 2011. Hepatitis C virus stimulates the phosphatidylinositol 4-kinase III alpha-dependent phosphatidylinositol 4-phosphate production that is essential for its replication. J Virol 85:8870–8883. https://doi.org/10.1128/JVI.00059-11.
9. Lim YS, Hwang SB. 2011. Hepatitis C virus NS5A protein interacts with phosphatidylinositol 4-kinase type IIIalpha and regulates viral propagation. J Biol Chem 286:11290–11298. https://doi.org/10.1074/jbc.M110.194472.
10. Reiss S, Rebhan I, Backes P, Romero-Brey I, Erfle H, Matula P, Kaderali L, Poenisch M, Blankenburg H, Hiet MS, Longerich T, Diehl S, Ramirez F, Balla T, Rohr K, Kaul A, Buhler S, Pepperkok R, Lengauer T, Albrecht M, Eils R, Schirmacher P, Lohmann V, Bartenschlager R. 2011. Recruitment and activation 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.
11. Love RA, Brodsky O, Hickey MJ, Wells PA, Cronin CN. 2009. Crystal structure of a novel dimeric form of NS5A domain I protein from hepatitis C virus. J Virol 83:4395–4403. https://doi.org/10.1128/JVI.02352-08.
12. Lambert SM, Langley DR, Garnett JA, Angell R, Hedgethorne K, Meanwell NA, Matthews SJ. 2014. The crystal structure of NS5A domain 1 from genotype 1a reveals new clues to the mechanism of action for dimeric HCV inhibitors. Protein Sci 23:723–734. https://doi.org/10.1002/pro.2456.
13. Tellinghuisen TL, Marcotrigiano J, Rice CM. 2005. Structure of the zinc-binding domain of an essential component of the hepatitis C virus replicase. Nature 435:374–379. https://doi.org/10.1038/nature03580.
14. Feuerstein S, Solyom Z, Aladag A, Favier A, Schwarten M, Hoffmann S, Willbold D, Brutscher B. 2012. Transient structure and SH3 interaction sites in an intrinsically disordered fragment of the hepatitis C virus protein NS5A. J Mol Biol 420:310–323. https://doi.org/10.1016/j.jmb.2012.04.023.
15. Chong WM, Hsu SC, Kao WT, Lo CW, Lee KY, Shao JS, Chen YH, Chang J, Chen SS, Yu MJ. 2016. Phosphoproteomics identified an NS5A phosphorylation site involved in hepatitis C virus replication. J Biol Chem 291:3918–3931. https://doi.org/10.1074/jbc.M115.675413.
16. Hsu SC, Lo CW, Pan TC, Lee KY, Yu MJ. 2017. Serine 235 is the primary NS5A hyper-phosphorylation site responsible for HCV replication. J Virol https://doi.org/10.1128/jvi.00194-17.
17. Lee KY, Chen YH, Hsu SC, Yu MJ. 2016. Phosphorylation of serine 235 of the hepatitis C virus non-structural protein NS5A by multiple kinases. PLoS One 11:e0166763. https://doi.org/10.1371/journal.pone.0166763.
18. Lemay KL, Treadaway J, Angulo I, Tellinghuisen TL. 2013. A hepatitis C virus NS5A phosphorylation site that regulates RNA replication. J Virol 87:1255–1260. https://doi.org/10.1128/JVI.02154-12.
19. Ross-Thriepland D, Harris M. 2014. Insights into the complexity and functionality of hepatitis C virus NS5A phosphorylation. J Virol 88:1421–1432. https://doi.org/10.1128/JVI.03017-13.
20. Masaki T, Matsunaga S, Takahashi H, Nakashima K, Kimura Y, Ito M, Matsuda M, Murayama A, Kato T, Hirano H, Endo Y, Lemon SM, Wakita T, Sawasaki T, Suzuki T. 2014. Involvement of hepatitis C virus NS5A hyperphosphorylation mediated by casein kinase I-alpha in infectious virus production. J Virol 88:7541–7555. https://doi.org/10.1128/JVI.03170-13.
21. Eyre NS, Hampton-Smith RJ, Aloia AL, Eddes JS, Simpson KJ, Hoffmann P, Beard MR. 2016. Phosphorylation of NS5A serine-235 is essential to hepatitis C virus RNA replication and normal replication compartment formation. Virology 491:27–44. https://doi.org/10.1016/j.virol.2016.01.018.
22. Cordek DG, Croom-Perez TJ, Hwang J, Hargittai MR, Subba-Reddy CV, Han Q, Lodeiro MF, Ning G, McCrory TS, Arnold JJ, Koc H, Lindenbach BD, Showalter SA, Cameron CE. 2014. Expanding the proteome of an RNA virus by phosphorylation of an intrinsically disordered viral protein. J Biol Chem 289:24397–24416. https://doi.org/10.1074/jbc.M114.589911.
23. Tellinghuisen TL, Foss KL, Treadaway J. 2008. Regulation of hepatitis C virion production via phosphorylation of the NS5A protein. PLoS Pathog 4:e1000032. https://doi.org/10.1371/journal.ppat.1000032.
24. Ross-Thriepland D, Mankouri J, Harris M. 2015. Serine phosphorylation of the hepatitis C virus NS5A protein controls the establishment of replication complexes. J Virol 89:3123–3135. https://doi.org/10.1128/JVI.02995-14.
25. Amako Y, Sarkeshik A, Hotta H, Yates J 3rd, Siddiqui A. 2009. Role of oxysterol binding protein in hepatitis C virus infection. J Virol 83:9237–9246. https://doi.org/10.1128/JVI.00958-09.
26. Nanda SK, Herion D, Liang TJ. 2006. Src homology 3 domain of hepatitis C virus NS5A protein interacts with Bin1 and is important for apoptosis and infectivity. Gastroenterology 130:794–809. https://doi.org/10.1053/j.gastro.2005.12.030.
27. Zech B, Kurtenbach A, Krieger N, Strand D, Blencke S, Morbitzer M, Salassidis K, Cotten M, Wissing J, Obert S, Bartenschlager R, Herget T, Daub H. 2003. Identification and characterization of amphiphysin II as a novel cellular interaction partner of the hepatitis C virus NS5A protein. J Gen Virol 84:555–560. https://doi.org/10.1099/vir.0.18801-0.
28. Masumi A, Aizaki H, Suzuki T, DuHadaway JB, Prendergast GC, Komuro K, Fukazawa H. 2005. Reduction of hepatitis C virus NS5A phosphorylation through its interaction with amphiphysin II. Biochem Biophys Res Commun 336:572–578. https://doi.org/10.1016/j.bbrc.2005.08.142.
29. Xu G, Xin X, Zheng C. 2013. GPS2 is required for the association of NS5A with VAP-A and hepatitis C virus replication. PLoS One 8:e78195. https://doi.org/10.1371/journal.pone.0078195.
30. Tu H, Gao L, Shi ST, Taylor DR, Yang T, Mircheff AK, Wen Y, Gorbalenya AE, Hwang SB, Lai MM. 1999. Hepatitis C virus RNA polymerase and NS5A complex with a SNARE-like protein. Virology 263:30–41. https://doi.org/10.1006/viro.1999.9893.
31. Gao L, Aizaki H, He JW, Lai MM. 2004. Interactions between viral nonstructural proteins and host protein hVAP-33 mediate the formation of hepatitis C virus RNA replication complex on lipid raft. J Virol 78:3480–3488. https://doi.org/10.1128/JVI.78.7.3480-3488.2004.
32. Evans MJ, Rice CM, Goff SP. 2004. Phosphorylation of hepatitis C virus nonstructural protein 5A modulates its protein interactions and viral RNA replication. Proc Natl Acad Sci U S A 101:13038–13043. https://doi.org/10.1073/pnas.0405152101.
33. Germain MA, Chatel-Chaix L, Gagne B, Bonneil E, Thibault P, Pradezynski F, de Chassey B, Meyniel-Schicklin L, Lotteau V, Baril M, Lamarre D. 2014. Elucidating novel hepatitis C virus-host interactions using combined mass spectrometry and functional genomics approaches. Mol Cell Proteomics 13:184–203. https://doi.org/10.1074/mcp.M113.030155.
34. Pichlmair A, Kandasamy K, Alvisi G, Mulhern O, Sacco R, Habjan M, Binder M, Stefanovic A, Eberle CA, Goncalves A, Burckstummer T, Muller AC, Fauster A, Holze C, Lindsten K, Goodbourn S, Kochs G, Weber F, Bartenschlager R, Bowie AG, Bennett KL, Colinge J, Superti-Furga G. 2012. Viral immune modulators perturb the human molecular network by common and unique strategies. Nature 487:486–490. https://doi.org/10.1038/nature11289.
35. Gupta N, Thakker S, Verma SC. 2016. KSHV encoded LANA recruits nucleosome assembly protein NAP1L1 for regulating viral DNA replication and transcription. Sci Rep 6:32633. https://doi.org/10.1038/srep32633.
36. De Marco A, Dans PD, Knezevich A, Maiuri P, Pantano S, Marcello A. 2010. Subcellular localization of the interaction between the human immunodeficiency virus transactivator Tat and the nucleosome assembly protein 1. Amino Acids 38:1583–1593. https://doi.org/10.1007/s00726-009-0378-9.
37. Söderberg O, Gullberg M, Jarvius M, Ridderstrale K, Leuchowius KJ, Jarvius J, Wester K, Hydbring P, Bahram F, Larsson LG, Landegren U. 2006. Direct observation of individual endogenous protein complexes in situ by proximity ligation. Nat Methods 3:995–1000. https://doi.org/10.1038/nmeth947.
38. Gullberg M, Gustafsdottir SM, Schallmeiner E, Jarvius J, Bjarnegard M, Betsholtz C, Landegren U, Fredriksson S. 2004. Cytokine detection by antibody-based proximity ligation. Proc Natl Acad Sci U S A 101:8420–8424. https://doi.org/10.1073/pnas.0400552101.
39. Hamamoto I, Nishimura Y, Okamoto T, Aizaki H, Liu M, Mori Y, Abe T, Suzuki T, Lai MM, Miyamura T, Moriishi K, Matsuura Y. 2005. Human VAP-B is involved in hepatitis C virus replication through interaction with NS5A and NS5B. J Virol 79:13473–13482. https://doi.org/10.1128/JVI.79.21.13473-13482.2005.
40. Miyaji-Yamaguchi M, Kato K, Nakano R, Akashi T, Kikuchi A, Nagata K. 2003. Involvement of nucleocytoplasmic shuttling of yeast Nap1 in mitotic progression. Mol Cell Biol 23:6672–6684. https://doi.org/10.1128/MCB.23.18.6672-6684.2003.
41. Okuwaki M, Kato K, Nagata K. 2010. Functional characterization of

3 CONCLUDING REMARKS AND OUTLOOK

Using mass spectrometry and biochemical methods, I identified the largely uncharacterized protein C17orf85 (now named NCBP3) as a cap-binding protein that, together with NCBP1, forms the alternative CBC enabling export of mRNA from the nucleus. Under physiological conditions, the alternative and canonical CBC function in a redundant manner. Through the identification of the alternative CBC, I identified a critical additional part for mRNA processing and showed that mRNA export is more complex than previously thought. Supportively for that notion, other recent studies have also suggested previously unappreciated complexity in the process of mRNA export. It became evident that besides the constitutive export of bulk mRNA, export of transcripts can be highly selective, allowing priority export for some mRNAs over others [221]. The identification of the alternative CBC raises many questions and facilitates further regulatory possibilities related to mRNA export and quality control. Current questions vary from the molecular composition and cooperation of the CBC subunits and their binding to the cap structure as well as the mechanism behind the engagement of other macromolecule complexes like the EJC or TREX complexes to the functional requirement of two redundant CBCs.

Investigating the molecular composition and cooperation of the CBC subunits is of fundamental importance to understand the synergies and discrepancies in binding of the subunits to each other and their interaction with the cap structure. The interaction between NCBP1 and NCBP2 as well as between NCBP2 and the cap structure has been studied using structural approaches [348,349]. Comparative modelling for NCBP3 using HHpred revealed a structural model based on the RRM of poly(A)-specific ribonuclease (PARN). The similarity to the RRM of PARN suggests a similar cap-binding mode for NCBP3 engaging the cap structure by stacking the cap on a single aromatic amino acid (tryptophan). In contrast, NCBP2 as well as the cytoplasmic cap-binding protein eIF4E sandwich the RNA cap structure between two aromatic amino acids (Figure 14). For NCBP3, the structure prediction model allowed to identify critical residues and suggested similar binding as seen for PARN. However, it is likely that other aromatic amino acids contribute to the cap-binding of NCBP3. More detailed structural analysis using x-ray crystallography or cryo-electron microscopy combined with mutational and functional experiments may be necessary to identify the exact binding mode of capped RNA to NCBP3.

An atomic structure of the NCBP1/3 complex in comparison with the crystal structure of the NCBP1/2 complex could help to understand whether the binding of NCBP3 is mutually exclusive or whether a trimeric complex of NCBP1/2/3 exists. Moreover, structural information could

support the understanding of RNA selectivity by the two CBCs. Further investigation of the influence of NCBP2 on the binding between NCBP1 and -3 and vice versa, the influence of NCBP3 on NCBP1/2 interaction, would be needed to understand the synergies and discords of the two complexes in eukaryotes.

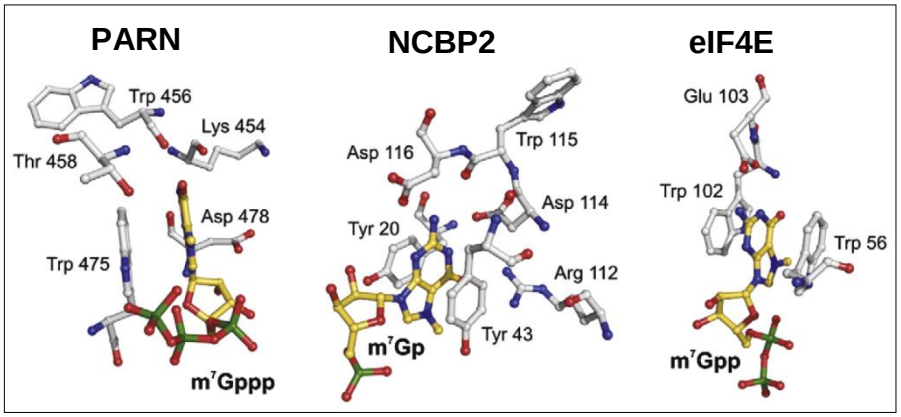


Figure 14: Comparison of cap-binding pockets for PARN, NCBP2 and eIF4E.

Amino acid residues of poly(A)-specific ribonuclease (PARN), nuclear cap-binding protein 2 (NCBP2) and eukaryotic initiation factor 4E (eIF4E) involved in binding to the m^7G cap structure. Amino acids are shown in ball-and-stick mode. Red, oxygen; green, phosphor; blue, nitrogen; yellow, carbon. Schematic and figure legend were adapted from reference [350].

The binding of the CBCs is required to recruit cellular components that are fundamental for RNA processing. However, it is still unclear how the two CBCs select the binding of the right RNA processing factors for the given type of RNA. We showed that NCBP2 and NCBP3 associate with EJC and TREX factors to different extents, preferentially binding to NCBP3. However, PHAX, which is the export adaptor for snRNAs export, only associated with NCBP2. EJC, TREX and PHAX are recruited to mRNA via NCBP1, a common protein of the canonical and the alternative CBC. It is therefore likely that association of NCBP2 or NCBP3 to NCBP1 exposes or blocks binding sites on NCBP1 for RNA processing factors and thereby mediates selectivity in recruiting specific RNA processing factors. Compared to NCBP2, NCBP3 is a relatively large protein of ~70 kDa with a disordered C-terminal domain. Through interaction with NCBP1 or other factors, the C-terminal domain may fold and engage mRNA processing factors like EJC and TREX subunits. The contribution of NCBP3`s C-terminal domain for the recruitment of mRNA processing factors could be addressed by combining structural analysis with functional and mutational biochemical experiments.

Altogether, the molecular basis of the interplay between the subunits of the alternative and canonical CBC and their engagement of auxiliary mRNA processing factors will allow to better understand the apparent redundancy of the two CBCs. We questioned why the human organism encodes for two CBCs with apparently redundant function. Different reasons could be possible:

(1) The two CBCs may be separated in labour and ensure proper processing of all needed RNAs. Naturally, the alternative and canonical CBC handle different types of RNA (e.g. snRNA and mRNA), however, during loss of one complex the other can compensate for the loss establishing a form of fallback mechanism. (2) Two CBCs may be necessary under conditions demanding regulation of RNA processing, e.g. during conditions of environmental stress or developmental processes. (3) The two CBCs may be regulated in their activity, which opens the possibility to promote usage of one CBC over the other.

In my second first author publication, investigating the relevance of the alternative CBC during virus infection, I could demonstrate that the alternative CBC is required to constrain viral infections. This supports the possibility that the two CBCs evolved to regulate RNA turnover during environmental stress conditions such as viral infections. Although I could show that the expression of a subset of ISGs is regulated by NCBP3, the exact mechanism how the alternative CBC is selected for export of this subset of mRNA is still unclear. In this scenario, a regulation of the CBCs by post-translational modifications could be likely. One commonly found modification is the phosphorylation of serine, threonine or tyrosine. Several phospho-sites were identified in NCBP1, -2 and -3 having the ability to be regulated by environmental stimuli [351–353,353–358]. Interestingly, mass spectrometry based identification of phosphorylation changes (preliminary, unpublished data) revealed that defined residues in NCBP1 and NCBP3 change their phosphorylation state in response to stimulation with IFN. This supports the possibility that the alternative CBC is regulated upon environmental stimuli by post-translational modifications. To elucidate the underlying mechanism, further experiments involving mutational and functional experiments are required, which could demonstrate that gene expression is controlled on a yet unappreciated level.

Respiratory infections with influenza A virus are a major cause of death worldwide [359–361]. Therefore, it is of outstanding interest to understand how severity of disease and pathogenicity towards respiratory virus infections is regulated. Infection of Ncbp3 deficient mice with IAV caused severe lung pathology resulting in high morbidity highlighting the influence of Ncbp3 on the outcome of disease. The expression of antiviral programs needs proper regulation to immediate and timely express cytokines and antiviral proteins as well as to express negative regulators to prevent overshooting of the system [362–365]. This regulation could be maintained by the function of Ncbp3 and thereby loss of Ncbp3 results in the dysregulation of cellular antiviral programs resulting in lung pathology and mortality. Further research is needed to understand the

exact mechanism how Ncbp3 influences the severity of disease, which could help to prevent and/or treat virus-associated respiratory diseases.

Viruses evolved mechanisms to manipulate cellular mRNA processing and export machineries to either dampen innate immune responses and/or to hijack cellular resources for viral replication. IAV and VSV are highly potent in disrupting cellular mRNA processing pathways. Until now, the influence of the two viruses on mRNA export was only shown by interrupting functions of nuclear pore associated proteins. A direct influence of viral proteins on Ncbp3 to prevent expression of innate immune regulated genes may also be possible. However, this hypothesis needs further investigations by firstly investigating if viral proteins potentially can interact with Ncbp3 and if so, by secondly examining the underlying cellular mechanism of viral NCBP3 manipulation.

HCV infection is a major cause of acute hepatitis and chronic liver disease affecting more than 170 million people and being a widespread problem [366]. Understanding the molecular basis of HCV infection and replication is of major importance to prevent disease and improve disease outcome. Basic research on HCV life cycle lead to the development of directly acting antiviral (DAA) drugs currently used for the treatments of HCV infection [367]. NS5A is a highly phosphorylated protein and it will be challenging to descramble the unique function and the functional consequences of all NS5A phosphorylations. However, our finding that phosphorylation of serine 225 of HCV NS5A protein contributes to efficient HCV genome replication through the interaction with the cellular proteins NAPL1, Bin1 and VAP-A, shed light onto a very complex system and refined our understanding of the complexity of post-translational modifications for virus-host interaction. Since HCV genetics are highly variable and drug resistant strains are developing very rapidly [368], it is of great importance to break down the molecular basis for viral replication to establish new DAA agents. Recently, a class of NS5A targeting agents were developed and are very promising for the treatment of HCV infections [369]. In this regard, the insights we gained investigating NS5A S225-dependent cellular interactions may be useful for the generation and selection of new antiviral drugs.

In summary, the results presented in my thesis emphasize the importance to continuously reassess molecular mechanisms that are thought to be well established. Basic research still reveals yet undescribed cellular mechanisms, which are fundamental for our basic understanding of life, which can be used for the discovery of new and effective drugs that may be used for therapeutic purposes in the future.

REFERENCES

1. Kabachinski, G. & Schwartz, T. U. The nuclear pore complex – structure and function at a glance. *J. Cell Sci.* **128,** (2015).
2. Delaleau, M. & Borden, K. L. B. Multiple Export Mechanisms for mRNAs. *Cells* **4,** 452–73 (2015).
3. Ohno, M., Fornerod, M. & Mattaj, I. W. Nucleocytoplasmic transport: the last 200 nanometers. *Cell* **92,** 327–36 (1998).
4. Kuersten, S., Ohno, M. & Mattaj, I. W. Nucleocytoplasmic transport: Ran, beta and beyond. *Trends in Cell Biology* **11,** 497–503 (2001).
5. Köhler, A. & Hurt, E. Exporting RNA from the nucleus to the cytoplasm. *Nat. Rev. Mol. Cell Biol.* **8,** 761–73 (2007).
6. Hopper, A. K. & Phizicky, E. M. tRNA transfers to the limelight. *Genes Dev.* **17,** 162–180 (2003).
7. Wolin, S. L. & Matera, A. G. The trials and travels of tRNA. *Genes Dev.* **13,** 1–10 (1999).
8. Melton, D. A., De Robertis, E. M. & Cortese, R. Order and intracellular location of the events involved in the maturation of a spliced tRNA. *Nature* **284,** 143–148 (1980).
9. De Robertis, E. M., Black, P. & Nishikura, K. Intranuclear location of the tRNA splicing enzymes. *Cell* **23,** 89–93 (1981).
10. Arts, G. J., Kuersten, S., Romby, P., Ehresmann, B. & Mattaj, I. W. The role of exportin-t in selective nuclear export of mature tRNAs. *EMBO J.* **17,** 7430–41 (1998).
11. Kutay, U. *et al.* Identification of a tRNA-specific nuclear export receptor. *Mol. Cell* **1,** 359–69 (1998).
12. Lipowsky, G. *et al.* Coordination of tRNA nuclear export with processing of tRNA. *RNA* **5,** 539–49 (1999).
13. Lund, E. & Dahlberg, J. E. Proofreading and aminoacylation of tRNAs before export from the nucleus. *Science* **282,** 2082–5 (1998).
14. Shaheen, H. H. & Hopper, A. K. Retrograde movement of tRNAs from the cytoplasm to the nucleus in Saccharomyces cerevisiae. *Proc. Natl. Acad. Sci.* **102,** 11290–11295 (2005).
15. Takano, A., Endo, T. & Yoshihisa, T. tRNA Actively Shuttles Between the Nucleus and Cytosol in Yeast. *Science (80-.).* **309,** 140–142 (2005).
16. Yoshihisa, T., Yunoki-Esaki, K., Ohshima, C., Tanaka, N. & Endo, T. Possibility of cytoplasmic pre-tRNA splicing: the yeast tRNA splicing endonuclease mainly localizes on the mitochondria. *Mol. Biol. Cell* **14,** 3266–79 (2003).
17. Murthi, A. *et al.* Regulation of tRNA Bidirectional Nuclear-Cytoplasmic Trafficking in Saccharomyces cerevisiae. *Mol. Biol. Cell* **21,** 639–649 (2010).
18. Whitney, M. L., Hurto, R. L., Shaheen, H. H. & Hopper, A. K. Rapid and reversible nuclear accumulation of cytoplasmic tRNA in response to nutrient availability. *Mol. Biol. Cell* **18,** 2678–86 (2007).
19. Hurto, R. L., Tong, A. H. Y., Boone, C. & Hopper, A. K. Inorganic Phosphate Deprivation Causes tRNA Nuclear Accumulation via Retrograde Transport in Saccharomyces cerevisiae. *Genetics* **176,** (2007).

20. Pierce, J. B., Eswara, M. B. K. & Mangroo, D. The ins and outs of nuclear re-export of retrogradely transported tRNAs in Saccharomyces cerevisiae. *Nucleus* **1,** 224–230 (2010).

21. Eswara, M. B. K., McGuire, A. T., Pierce, J. B. & Mangroo, D. Utp9p facilitates Msn5p-mediated nuclear reexport of retrograded tRNAs in Saccharomyces cerevisiae. *Mol. Biol. Cell* **20,** 5007–25 (2009).

22. Kramer, E. B. & Hopper, A. K. Retrograde transfer RNA nuclear import provides a new level of tRNA quality control in Saccharomyces cerevisiae. *Proc. Natl. Acad. Sci. U. S. A.* **110,** 21042–7 (2013).

23. Hopper, A. K. Transfer RNA Post-Transcriptional Processing, Turnover, and Subcellular Dynamics in the Yeast Saccharomyces cerevisiae. *Genetics* **194,** 43–67 (2013).

24. Shaheen, H. H. *et al.* Retrograde nuclear accumulation of cytoplasmic tRNA in rat hepatoma cells in response to amino acid deprivation. *Proc. Natl. Acad. Sci.* **104,** 8845–8850 (2007).

25. Pierce, J. B. & Mangroo, D. *Schizosaccharomyces pombe* , unlike *Saccharomyces cerevisiae* , may not directly regulate nuclear-cytoplasmic transport of spliced tRNAs in response to nutrient availability. *Biochem. Cell Biol.* **89,** 554–561 (2011).

26. Chafe, S. C., Pierce, J. B., Eswara, M. B. K., McGuire, A. T. & Mangroo, D. Nutrient stress does not cause retrograde transport of cytoplasmic tRNA to the nucleus in evolutionarily diverse organisms. *Mol. Biol. Cell* **22,** 1091–103 (2011).

27. Johnstone, A. D., Mullen, R. T. & Mangroo, D. Plants, like mammals, but unlike *Saccharomyces,* do not regulate nuclear-cytoplasmic tRNA trafficking in response to nutrient stress. *Plant Signal. Behav.* **6,** 1183–1188 (2011).

28. Miyagawa, R., Mizuno, R., Watanabe, K. & Ijiri, K. Formation of tRNA granules in the nucleus of heat-induced human cells. *Biochem. Biophys. Res. Commun.* **418,** 149–155 (2012).

29. Arts, G.-J., Fornerod, M. & Mattaj, lain W. Identification of a nuclear export receptor for tRNA. *Curr. Biol.* **8,** 305–314 (1998).

30. Calado, A., Treichel, N., Müller, E.-C., Otto, A. & Kutay, U. Exportin-5-mediated nuclear export of eukaryotic elongation factor 1A and tRNA. *EMBO J.* **21,** 6216–24 (2002).

31. Bohnsack, M. T. *et al.* Exp5 exports eEF1A via tRNA from nuclei and synergizes with other transport pathways to confine translation to the cytoplasm. *EMBO J.* **21,** 6205–15 (2002).

32. Hellmuth, K. *et al.* Yeast Los1p has properties of an exportin-like nucleocytoplasmic transport factor for tRNA. *Mol. Cell. Biol.* **18,** 6374–86 (1998).

33. Bartel, D. P. MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell* **116,** 281–97 (2004).

34. Kim, V. N. MicroRNA biogenesis: coordinated cropping and dicing. *Nat. Rev. Mol. Cell Biol.* **6,** 376–385 (2005).

35. Valencia-Sanchez, M. A., Liu, J., Hannon, G. J. & Parker, R. Control of translation and mRNA degradation by miRNAs and siRNAs. *Genes Dev.* **20,** 515–524 (2006).

36. Leung, A. K. L. The Whereabouts of microRNA Actions: Cytoplasm and Beyond. *Trends Cell Biol.* **25,** 601–610 (2015).

37. Lee, Y. *et al.* MicroRNA genes are transcribed by RNA polymerase II. *EMBO J.* **23,** 4051–60 (2004).

38. CAI, X., Hagedorn, C. H. & Cullen, B. R. Human microRNAs are processed from capped,

polyadenylated transcripts that can also function as mRNAs. *RNA* **10,** 1957–1966 (2004).

39. Borchert, G. M., Lanier, W. & Davidson, B. L. RNA polymerase III transcribes human microRNAs. *Nat. Struct. Mol. Biol.* **13,** 1097–1101 (2006).

40. Lee, Y. *et al.* The nuclear RNase III Drosha initiates microRNA processing. *Nature* **425,** 415–419 (2003).

41. Han, J. *et al.* Molecular Basis for the Recognition of Primary microRNAs by the Drosha-DGCR8 Complex. *Cell* **125,** 887–901 (2006).

42. Denli, A. M., Tops, B. B. J., Plasterk, R. H. A., Ketting, R. F. & Hannon, G. J. Processing of primary microRNAs by the Microprocessor complex. *Nature* **432,** 231–235 (2004).

43. Gregory, R. I. *et al.* The Microprocessor complex mediates the genesis of microRNAs. *Nature* **432,** 235–240 (2004).

44. Lund, E., Güttinger, S., Calado, A., Dahlberg, J. E. & Kutay, U. Nuclear Export of MicroRNA Precursors. *Science (80-.).* **303,** 95–98 (2004).

45. Bohnsack, M. T., Czaplinski, K. & Gorlich, D. Exportin 5 is a RanGTP-dependent dsRNA-binding protein that mediates nuclear export of pre-miRNAs. *RNA* **10,** 185–91 (2004).

46. Yi, R., Qin, Y., Macara, I. G. & Cullen, B. R. Exportin-5 mediates the nuclear export of pre-microRNAs and short hairpin RNAs. *Genes Dev.* **17,** 3011–6 (2003).

47. Okada, C. *et al.* A High-Resolution Structure of the Pre-microRNA Nuclear Export Machinery. *Science (80-.).* **326,** 1275–1279 (2009).

48. Bernstein, E., Caudy, A. A., Hammond, S. M. & Hannon, G. J. Role for a bidentate ribonuclease in the initiation step of RNA interference. *Nature* **409,** 363–366 (2001).

49. Schwarz, D. S. *et al.* Asymmetry in the assembly of the RNAi enzyme complex. *Cell* **115,** 199–208 (2003).

50. Khvorova, A., Reynolds, A. & Jayasena, S. D. Functional siRNAs and miRNAs exhibit strand bias. *Cell* **115,** 209–16 (2003).

51. Ha, M. & Kim, V. N. Regulation of microRNA biogenesis. *Nat. Rev. Mol. Cell Biol.* **15,** 509–524 (2014).

52. Berezikov, E., Chung, W.-J., Willis, J., Cuppen, E. & Lai, E. C. Mammalian Mirtron Genes. *Mol. Cell* **28,** 328–336 (2007).

53. Okamura, K., Hagen, J. W., Duan, H., Tyler, D. M. & Lai, E. C. The Mirtron Pathway Generates microRNA-Class Regulatory RNAs in Drosophila. *Cell* **130,** 89–100 (2007).

54. Babiarz, J. E., Ruby, J. G., Wang, Y., Bartel, D. P. & Blelloch, R. Mouse ES cells express endogenous shRNAs, siRNAs, and other Microprocessor-independent, Dicer-dependent small RNAs. *Genes Dev.* **22,** 2773–2785 (2008).

55. Chong, M. M. W. *et al.* Canonical and alternate functions of the microRNA biogenesis machinery. *Genes Dev.* **24,** 1951–1960 (2010).

56. Xie, M. *et al.* Mammalian 5'-capped microRNA precursors that generate a single microRNA. *Cell* **155,** 1568–80 (2013).

57. Cheloufi, S., Dos Santos, C. O., Chong, M. M. W. & Hannon, G. J. A dicer-independent miRNA biogenesis pathway that requires Ago catalysis. *Nature* **465,** 584–589 (2010).

58. Cifuentes, D. *et al.* A Novel miRNA Processing Pathway Independent of Dicer Requires Argonaute2 Catalytic Activity. *Science (80-.).* **328,** 1694–1698 (2010).

59. Yang, J.-S. *et al.* Conserved vertebrate mir-451 provides a platform for Dicer-independent, Ago2-mediated microRNA biogenesis. *Proc. Natl. Acad. Sci.* **107,** 15163–15168 (2010).

60. Yoda, M. *et al.* Poly(A)-Specific Ribonuclease Mediates 3′-End Trimming of Argonaute2-Cleaved Precursor MicroRNAs. *Cell Rep.* **5,** 715–726 (2013).

61. Heo, I. *et al.* Mono-Uridylation of Pre-MicroRNA as a Key Step in the Biogenesis of Group II let-7 MicroRNAs. *Cell* **151,** 521–532 (2012).

62. Fromont-Racine, M., Senger, B., Saveanu, C. & Fasiolo, F. Ribosome assembly in eukaryotes. *Gene* **313,** 17–42 (2003).

63. Tschochner, H. & Hurt, E. Pre-ribosomes on the road from the nucleolus to the cytoplasm. *Trends Cell Biol.* **13,** 255–63 (2003).

64. Fatica, A. & Tollervey, D. Making ribosomes. *Curr. Opin. Cell Biol.* **14,** 313–8 (2002).

65. Rouquette, J., Choesmel, V. & Gleizes, P.-E. Nuclear export and cytoplasmic processing of precursors to the 40S ribosomal subunits in mammalian cells. *EMBO J.* **24,** 2862–2872 (2005).

66. Schäfer, T., Strauss, D., Petfalski, E., Tollervey, D. & Hurt, E. The path from nucleolar 90S to cytoplasmic 40S pre-ribosomes. *EMBO J.* **22,** 1370–80 (2003).

67. Ferreira-Cerca, S., Pöll, G., Gleizes, P.-E., Tschochner, H. & Milkereit, P. Roles of Eukaryotic Ribosomal Proteins in Maturation and Transport of Pre-18S rRNA and Ribosome Function. *Mol. Cell* **20,** 263–275 (2005).

68. Moy, T. I. & Silver, P. A. Nuclear export of the small ribosomal subunit requires the ran-GTPase cycle and certain nucleoporins. *Genes Dev.* **13,** 2118–33 (1999).

69. Trotta, C. R., Lund, E., Kahan, L., Johnson, A. W. & Dahlberg, J. E. Coordinated nuclear export of 60S ribosomal subunits and NMD3 in vertebrates. *EMBO J.* **22,** 2841–51 (2003).

70. Thomas, F. & Kutay, U. Biogenesis and nuclear export of ribosomal subunits in higher eukaryotes depend on the CRM1 export pathway. *J. Cell Sci.* **116,** (2003).

71. Gadal, O. *et al.* Nuclear export of 60s ribosomal subunits depends on Xpo1p and requires a nuclear export sequence-containing factor, Nmd3p, that associates with the large subunit protein Rpl10p. *Mol. Cell. Biol.* **21,** 3405–15 (2001).

72. Ho, J. H., Kallstrom, G. & Johnson, A. W. Nmd3p is a Crm1p-dependent adapter protein for nuclear export of the large ribosomal subunit. *J. Cell Biol.* **151,** 1057–66 (2000).

73. Stage-Zimmermann, T., Schmidt, U. & Silver, P. A. Factors affecting nuclear export of the 60S ribosomal subunit in vivo. *Mol. Biol. Cell* **11,** 3777–89 (2000).

74. Hurt, E. *et al.* A novel in vivo assay reveals inhibition of ribosomal nuclear export in ran-cycle and nucleoporin mutants. *J. Cell Biol.* **144,** 389–401 (1999).

75. Nissan, T. A., Bassler, J., Petfalski, E., Tollervey, D. & Hurt, E. 60S pre-ribosome formation viewed from assembly in the nucleolus until export to the cytoplasm. *EMBO J.* **21,** 5539–47 (2002).

76. West, M., Hedges, J. B., Chen, A. & Johnson, A. W. Defining the order in which Nmd3p and Rpl10p load onto nascent 60S ribosomal subunits. *Mol. Cell. Biol.* **25,** 3802–13 (2005).

77. Hedges, J., West, M. & Johnson, A. W. Release of the export adapter, Nmd3p, from the 60S ribosomal subunit requires Rpl10p and the cytoplasmic GTPase Lsg1p. *EMBO J.* **24,** 567–579 (2005).

78. Lebreton, A. *et al.* A functional network involved in the recycling of nucleocytoplasmic pre-60S

factors. *J. Cell Biol.* **173,** 349–60 (2006).

79. Hung, N.-J. & Johnson, A. W. Nuclear Recycling of the Pre-60S Ribosomal Subunit-Associated Factor Arx1 Depends on Rei1 in Saccharomyces cerevisiae. *Mol. Cell. Biol.* **26,** 3718–3727 (2006).

80. Belaya, K., Tollervey, D. & Kos, M. FLIPing heterokaryons to analyze nucleo-cytoplasmic shuttling of yeast proteins. *RNA* **12,** 921–30 (2006).

81. Bradatsch, B. *et al.* Arx1 functions as an unorthodox nuclear export receptor for the 60S preribosomal subunit. *Mol. Cell* **27,** 767–79 (2007).

82. Squatrito, M., Mancino, M., Donzelli, M., Areces, L. B. & Draetta, G. F. EBP1 is a nucleolar growth-regulating protein that is part of pre-ribosomal ribonucleoprotein complexes. *Oncogene* **23,** 4454–4465 (2004).

83. Matera, A. G., Terns, R. M. & Terns, M. P. Non-coding RNAs: lessons from the small nuclear and small nucleolar RNAs. *Nat Rev Mol Cell Biol* **8,** 209–220 (2007).

84. Hernandez, N. Small nuclear RNA genes: a model system to study fundamental mechanisms of transcription. *J. Biol. Chem.* **276,** 26733–6 (2001).

85. Kunkel, G. R., Maser, R. L., Calvet, J. P. & Pederson, T. U6 small nuclear RNA is transcribed by RNA polymerase III. *Proc. Natl. Acad. Sci. U. S. A.* **83,** 8575–9 (1986).

86. Schramm, L. & Hernandez, N. Recruitment of RNA polymerase III to its target promoters. *Genes Dev.* **16,** 2593–620 (2002).

87. Matera, A. G. & Wang, Z. A day in the life of the spliceosome. *Nat. Rev. Mol. Cell Biol.* **15,** 108–21 (2014).

88. Yong, J., Wan, L. & Dreyfuss, G. Why do cells need an assembly machine for RNA–protein complexes? *Trends Cell Biol.* **14,** 226–232 (2004).

89. Hamm, J. *et al.* Monomethylated cap structures facilitate RNA export from the nucleus. *Cell* **63,** 109–18 (1990).

90. Medlin, J. E., Uguen, P., Taylor, A., Bentley, D. L. & Murphy, S. The C-terminal domain of pol II and a DRB-sensitive kinase are required for 3' processing of U2 snRNA. *EMBO J.* **22,** 925–934 (2003).

91. Cuello, P., Boyd, D. C., Dye, M. J., Proudfoot, N. J. & Murphy, S. Transcription of the human U2 snRNA genes continues beyond the 3**J** box in vivo. *EMBO J.* **18,** 2867–2877 (1999).

92. Suzuki, T., Izumi, H. & Ohno, M. Cajal body surveillance of U snRNA export complex assembly. *J. Cell Biol.* **190,** 603–12 (2010).

93. Ohno, M., Segref, A., Bachi, A., Wilm, M. & Mattaj, I. W. PHAX, a Mediator of U snRNA Nuclear Export Whose Activity Is Regulated by Phosphorylation. *Cell* **101,** 187–198 (2000).

94. Izaurralde, E. *et al.* A cap-binding protein complex mediating U snRNA export. *Nature* **376,** 709–12 (1995).

95. Segref, A., Mattaj, I. W. & Ohno, M. The evolutionarily conserved region of the U snRNA export mediator PHAX is a novel RNA-binding domain that is essential for U snRNA export. *RNA* **7,** 351–60 (2001).

96. Mouaikel, J., Verheggen, C., Bertrand, E., Tazi, J. & Bordonné, R. Hypermethylation of the cap structure of both yeast snRNAs and snoRNAs requires a conserved methyltransferase that is localized to the nucleolus. *Mol. Cell* **9,** 891–901 (2002).

97. Huang, Q. & Pederson, T. A human U2 RNA mutant stalled in 3' end processing is impaired in

nuclear import. *Nucleic Acids Res.* **27,** 1025–1031 (1999).

98. Huber, J. *et al.* Snurportin1, an m3G-cap-specific nuclear import receptor with a novel domain structure. *EMBO J.* **17,** 4114–4126 (1998).

99. Narayanan, U., Achsel, T., Lührmann, R. & Matera, A. G. Coupled In Vitro Import of U snRNPs and SMN, the Spinal Muscular Atrophy Protein. *Mol. Cell* **16,** 223–234 (2004).

100. Palacios, I., Hetzer, M., Adam, S. A. & Mattaj, I. W. Nuclear import of U snRNPs requires importin beta. *EMBO J.* **16,** 6783–92 (1997).

101. Matera, A. G. & Shpargel, K. B. Pumping RNA: nuclear bodybuilding along the RNP pipeline. *Curr. Opin. Cell Biol.* **18,** 317–324 (2006).

102. Sleeman, J. E. & Lamond, A. I. Newly assembled snRNPs associate with coiled bodies before speckles, suggesting a nuclear snRNP maturation pathway. *Curr. Biol.* **9,** 1065–1074 (1999).

103. Staněk, D., Rader, S. D., Klingauf, M. & Neugebauer, K. M. Targeting of U4/U6 small nuclear RNP assembly factor SART3/p110 to Cajal bodies. *J. Cell Biol.* **160,** 505–16 (2003).

104. Nesic, D. A role for Cajal bodies in the final steps of U2 snRNP biogenesis. *J. Cell Sci.* (2004). doi:10.1242/jcs.01308

105. Schaffert, N., Hossbach, M., Heintzmann, R., Achsel, T. & Lührmann, R. RNAi knockdown of hPrp31 leads to an accumulation of U4/U6 di-snRNPs in Cajal bodies. *EMBO J.* **23,** 3000–9 (2004).

106. Staněk, D. & Neugebauer, K. M. Detection of snRNP assembly intermediates in Cajal bodies by fluorescence resonance energy transfer. *J. Cell Biol.* **166,** 1015–25 (2004).

107. Decroly, E., Ferron, F., Lescar, J. & Canard, B. Conventional and unconventional mechanisms for capping viral mRNA. *Nat. Rev. Microbiol.* **10,** 51–65 (2011).

108. Egloff, S. & Murphy, S. Cracking the RNA polymerase II CTD code. *Trends Genet.* **24,** 280–288 (2008).

109. Topisirovic, I., Svitkin, Y. V, Sonenberg, N. & Shatkin, A. J. Cap and cap-binding proteins in the control of gene expression. *Wiley Interdiscip. Rev. RNA* **2,** 277–98 (2011).

110. Ramanathan, A., Robb, G. B. & Chan, S. H. mRNA capping: Biological functions and applications. *Nucleic Acids Res.* **44,** 7511–7526 (2016).

111. Cowling, V. H. & Cole, M. D. Regulation of mRNA Cap Methylation. *Genes Cancer* **1,** 576–579 (2010).

112. Gonatopoulos-Pournatzis, T. & Cowling, V. H. Cap-binding complex (CBC). *Biochem. J.* **457,** 231–42 (2014).

113. Daffis, S. *et al.* 2′-O methylation of the viral mRNA cap evades host restriction by IFIT family members. *Nature* **468,** 452–456 (2010).

114. Ohno, M. Size matters in RNA export. *RNA Biol.* **9,** 1413–7 (2012).

115. Masuyama, K., Taniguchi, I., Kataoka, N. & Ohno, M. RNA length defines RNA export pathway. *Genes Dev.* **18,** 2074–85 (2004).

116. König, J. *et al.* iCLIP reveals the function of hnRNP particles in splicing at individual nucleotide resolution. *Nat. Struct. Mol. Biol.* **17,** 909–15 (2010).

117. McAfee, J. G., Shahied-Milam, L., Soltaninassab, S. R. & LeStourgeon, W. M. A major determinant of hnRNP C protein binding to RNA is a novel bZIP-like RNA binding domain. *RNA*

2, 1139–52 (1996).

118. Huang, M. *et al.* The C-protein tetramer binds 230 to 240 nucleotides of pre-mRNA and nucleates the assembly of 40S heterogeneous nuclear ribonucleoprotein particles. *Mol. Cell. Biol.* **14,** 518–33 (1994).

119. Weighardt, F., Biamonti, G. & Riva, S. The roles of heterogeneous nuclear ribonucleoproteins (hnRNP) in RNA metabolism. *Bioessays* **18,** 747–56 (1996).

120. Müller-McNicoll, M. & Neugebauer, K. M. How cells get the message: dynamic assembly and function of mRNA–protein complexes. *Nat. Rev. Genet.* **14,** 275–287 (2013).

121. Listerman, I., Sapra, A. K. & Neugebauer, K. M. Cotranscriptional coupling of splicing factor recruitment and precursor messenger RNA splicing in mammalian cells. *Nat. Struct. Mol. Biol.* **13,** 815–22 (2006).

122. Wahl, M. C., Will, C. L. & Lührmann, R. The Spliceosome: Design Principles of a Dynamic RNP Machine. *Cell* **136,** 701–718 (2009).

123. Bentley, D. L. Rules of engagement: co-transcriptional recruitment of pre-mRNA processing factors. *Curr. Opin. Cell Biol.* **17,** 251–6 (2005).

124. Inoue, K., Ohno, M., Sakamoto, H. & Shimura, Y. Effect of the cap structure on pre-mRNA splicing in Xenopus oocyte nuclei. *Genes Dev.* **3,** 1472–9 (1989).

125. Izaurralde, E. *et al.* A nuclear cap binding protein complex involved in pre-mRNA splicing. *Cell* **78,** 657–668 (1994).

126. Konarska, M. M., Padgett, R. A. & Sharp, P. A. Recognition of cap structure in splicing in vitro of mRNA precursors. *Cell* **38,** 731–6 (1984).

127. Edery, I. & Sonenberg, N. Cap-dependent RNA splicing in a HeLa nuclear extract. *Proc. Natl. Acad. Sci. U. S. A.* **82,** 7590–4 (1985).

128. Ohno, M., Sakamoto, H. & Shimura, Y. Preferential excision of the 5' proximal intron from mRNA precursors with two introns as mediated by the cap structure. *Proc. Natl. Acad. Sci. U. S. A.* **84,** 5187–91 (1987).

129. Patzelt, E., Thalmann, E., Hartmuth, K., Blaas, D. & Kuechler, E. Assembly of pre-mRNA splicing complex is cap dependent. *Nucleic Acids Res.* **15,** 1387–99 (1987).

130. Lewis, J. D., Izaurralde, E., Jarmolowski, A., McGuigan, C. & Mattaj, I. W. A nuclear cap-binding complex facilitates association of U1 snRNP with the cap-proximal 5' splice site. *Genes Dev.* **10,** 1683–98 (1996).

131. Pabis, M. *et al.* The nuclear cap-binding complex interacts with the U4/U6·U5 tri-snRNP and promotes spliceosome assembly in mammalian cells. *RNA* **19,** 1054–63 (2013).

132. Jiao, X., Chang, J. H., Kilic, T., Tong, L. & Kiledjian, M. A Mammalian Pre-mRNA 5′ End Capping Quality Control Mechanism and an Unexpected Link of Capping to Pre-mRNA Processing. *Mol. Cell* **50,** 104–115 (2013).

133. Tange Thomas Øand Nott, A., Moore, M. J., Tange, T. Ø. & Nott, A. The ever-increasing complexities of the exon junction complex. *Curr. Opin. Cell Biol.* **16,** 279–284 (2004).

134. Le Hir, H., Izaurralde, E., Maquat, L. E. & Moore, M. J. The spliceosome deposits multiple proteins 20-24 nucleotides upstream of mRNA exon-exon junctions. *EMBO J.* **19,** 6860–9 (2000).

135. Chang, Y.-F., Imam, J. S. & Wilkinson, M. F. The nonsense-mediated decay RNA surveillance pathway. *Annu. Rev. Biochem.* **76,** 51–74 (2007).

136. Alexandrov, A., Colognori, D., Shu, M.-D. & Steitz, J. A. Human spliceosomal protein CWC22 plays a role in coupling splicing to exon junction complex deposition and nonsense-mediated decay. *Proc. Natl. Acad. Sci.* **109,** 21313–21318 (2012).

137. Barbosa, I. *et al.* Human CWC22 escorts the helicase eIF4AIII to spliceosomes and promotes exon junction complex assembly. *Nat. Struct. Mol. Biol.* **19,** 983–90 (2012).

138. Steckelberg, A.-L., Boehm, V., Gromadzka, A. M. & Gehring, N. H. CWC22 Connects Pre-mRNA Splicing and Exon Junction Complex Assembly. *Cell Rep.* **2,** 454–461 (2012).

139. Steckelberg, A.-L., Altmueller, J., Dieterich, C. & Gehring, N. H. CWC22-dependent pre-mRNA splicing and eIF4A3 binding enables global deposition of exon junction complexes. *Nucleic Acids Res.* **43,** 4687–700 (2015).

140. Bessonov, S., Anokhina, M., Will, C. L., Urlaub, H. & Lührmann, R. Isolation of an active step I spliceosome and composition of its RNP core. *Nature* **452,** 846–850 (2008).

141. Buchwald, G., Schüssler, S., Basquin, C., Le Hir, H. & Conti, E. Crystal structure of the human eIF4AIII-CWC22 complex shows how a DEAD-box protein is inhibited by a MIF4G domain. *Proc. Natl. Acad. Sci. U. S. A.* **110,** E4611-8 (2013).

142. Gehring, N. H., Lamprinaki, S., Hentze, M. W. & Kulozik, A. E. The hierarchy of exon-junction complex assembly by the spliceosome explains key features of mammalian nonsense-mediated mRNA decay. *PLoS Biol.* **7,** e1000120 (2009).

143. Andersen, C. B. F. *et al.* Structure of the Exon Junction Core Complex with a Trapped DEAD-Box ATPase Bound to RNA. *Science (80-.).* **313,** 1968–1972 (2006).

144. Ballut, L. *et al.* The exon junction core complex is locked onto RNA by inhibition of eIF4AIII ATPase activity. *Nat. Struct. Mol. Biol.* **12,** 861–869 (2005).

145. Nielsen, K. H. *et al.* Mechanism of ATP turnover inhibition in the EJC. *RNA* **15,** 67–75 (2008).

146. Bono, F., Ebert, J., Lorentzen, E. & Conti, E. The Crystal Structure of the Exon Junction Complex Reveals How It Maintains a Stable Grip on mRNA. *Cell* **126,** 713–725 (2006).

147. Merz, C., Urlaub, H., Will, C. L. & Lührmann, R. Protein composition of human mRNPs spliced in vitro and differential requirements for mRNP protein recruitment. *RNA* **13,** 116–28 (2007).

148. Singh, G. *et al.* The cellular EJC interactome reveals higher-order mRNP structure and an EJC-SR protein nexus. *Cell* **151,** 750–64 (2012).

149. Tange, T. Ø., Shibuya, T., Jurica, M. S., Moore, M. J. & Tange Thomas Øand Shibuya, T. Biochemical analysis of the EJC reveals two new factors and a stable tetrameric protein core. *RNA* **11,** 1869–1883 (2005).

150. Schmidt, U. *et al.* Assembly and mobility of exon-exon junction complexes in living cells. *RNA* **15,** 862–76 (2009).

151. Le Hir, H. & Andersen, G. R. Structural insights into the exon junction complex. *Curr. Opin. Struct. Biol.* **18,** 112–9 (2008).

152. Schwerk, C. *et al.* ASAP, a novel protein complex involved in RNA processing and apoptosis. *Mol. Cell. Biol.* **23,** 2981–90 (2003).

153. Li, C., Lin, R.-I., Lai, M.-C., Ouyang, P. & Tarn, W.-Y. Nuclear Pnn/DRS protein binds to spliced mRNPs and participates in mRNA processing and export via interaction with RNPS1. *Mol. Cell. Biol.* **23,** 7363–76 (2003).

154. Murachelli, A. G., Ebert, J., Basquin, C., Le Hir, H. & Conti, E. The structure of the ASAP core

complex reveals the existence of a Pinin-containing PSAP complex. *Nat. Struct. Mol. Biol.* **19,** 378–86 (2012).

155. Schmidt, C. *et al.* Mass spectrometry-based relative quantification of proteins in precatalytic and catalytically active spliceosomes by metabolic labeling (SILAC), chemical labeling (iTRAQ), and label-free spectral count. *RNA* **20,** 406–420 (2014).

156. Boehm, V. & Gehring, N. H. Exon Junction Complexes : Supervising the Gene Expression Assembly Line. *Trends Genet.* **32,** 724–735 (2016).

157. Chan, C. C. *et al.* eIF4A3 is a novel component of the exon junction complex. *RNA* **10,** 200–9 (2004).

158. Degot, S. *et al.* Association of the breast cancer protein MLN51 with the exon junction complex via its speckle localizer and RNA binding module. *J. Biol. Chem.* **279,** 33702–15 (2004).

159. Ferraiuolo, M. A. *et al.* A nuclear translation-like factor eIF4AIII is recruited to the mRNA during splicing and functions in nonsense-mediated decay. *Proc. Natl. Acad. Sci. U. S. A.* **101,** 4118–23 (2004).

160. Gatfield, D. & Izaurralde, E. REF1/Aly and the additional exon junction complex proteins are dispensable for nuclear mRNA export. *J. Cell Biol.* **159,** 579–88 (2002).

161. Kataoka, N. *et al.* Pre-mRNA splicing imprints mRNA in the nucleus with a novel RNA-binding protein that persists in the cytoplasm. *Mol. Cell* **6,** 673–82 (2000).

162. Kataoka, N., Diem, M. D., Kim, V. N., Yong, J. & Dreyfuss, G. Magoh, a human homolog of Drosophila mago nashi protein, is a component of the splicing-dependent exon-exon junction complex. *EMBO J.* **20,** 6424–33 (2001).

163. Le Hir, H., Gatfield, D., Braun, I. C., Forler, D. & Izaurralde, E. The protein Mago provides a link between splicing and mRNA localization. *EMBO Rep.* **2,** 1119–24 (2001).

164. Le Hir, H., Gatfield, D., Izaurralde, E. & Moore, M. J. The exon-exon junction complex provides a binding platform for factors involved in mRNA export and nonsense-mediated mRNA decay. *EMBO J.* **20,** 4987–97 (2001).

165. Luo, M.-J. L. *et al.* Pre-mRNA splicing and mRNA export linked by direct interactions between UAP56 and Aly. *Nature* **413,** 644–7 (2001).

166. Lykke-Andersen, J., Shu, M. D. & Steitz, J. A. Human Upf proteins target an mRNA for nonsense-mediated decay when bound downstream of a termination codon. *Cell* **103,** 1121–31 (2000).

167. Lykke-Andersen, J., Shu, M. D. & Steitz, J. A. Communication of the position of exon-exon junctions to the mRNA surveillance machinery by the protein RNPS1. *Science* **293,** 1836–9 (2001).

168. Palacios, I. M., Gatfield, D., St Johnston, D. & Izaurralde, E. An eIF4AIII-containing complex required for mRNA localization and nonsense-mediated mRNA decay. *Nature* **427,** 753–7 (2004).

169. Shibuya, T., Tange, T. Ø., Sonenberg, N. & Moore, M. J. eIF4AIII binds spliced mRNA in the exon junction complex and is essential for nonsense-mediated decay. *Nat. Struct. Mol. Biol.* **11,** 346–51 (2004).

170. Zhou, Z. *et al.* The protein Aly links pre-messenger-RNA splicing to nuclear export in metazoans. *Nature* **407,** 401–5 (2000).

171. Valencia, P., Dias, A. P. & Reed, R. Splicing promotes rapid and efficient mRNA export in mammalian cells. *Proc. Natl. Acad. Sci. U. S. A.* **105,** 3386–91 (2008).

172. Masuda, S. *et al.* Recruitment of the human TREX complex to mRNA during splicing. *Genes Dev.* **19,** 1512–7 (2005).

173. Rodríguez-Navarro, S. & Hurt, E. Linking gene regulation to mRNA production and export. *Curr. Opin. Cell Biol.* **23,** 302–9 (2011).

174. Strässer, K. *et al.* TREX is a conserved complex coupling transcription with messenger RNA export. *Nature* **417,** 304–8 (2002).

175. Dufu, K. *et al.* ATP is required for interactions between UAP56 and two conserved mRNA export proteins, Aly and CIP29, to assemble the TREX complex. *Genes Dev.* **24,** 2043–53 (2010).

176. Folco, E. G., Lee, C.-S., Dufu, K., Yamazaki, T. & Reed, R. The proteins PDIP3 and ZC11A associate with the human TREX complex in an ATP-dependent manner and function in mRNA export. *PLoS One* **7,** e43804 (2012).

177. Chang, C.-T. *et al.* Chtop is a component of the dynamic TREX mRNA export complex. *EMBO J.* **32,** 473–86 (2013).

178. Hautbergue, G. M. *et al.* UIF, a New mRNA export adaptor that works together with REF/ALY, requires FACT for recruitment to mRNA. *Curr. Biol.* **19,** 1918–24 (2009).

179. Viphakone, N. *et al.* Luzp4 defines a new mRNA export pathway in cancer cells. *Nucleic Acids Res.* **43,** 2353–2366 (2015).

180. Yamazaki, T. *et al.* The Closely Related RNA helicases, UAP56 and URH49, Preferentially Form Distinct mRNA Export Machineries and Coordinately Regulate Mitotic Progression. *Mol. Biol. Cell* **21,** 2953–2965 (2010).

181. Cheng, H. *et al.* Human mRNA export machinery recruited to the 5' end of mRNA. *Cell* **127,** 1389–400 (2006).

182. Viphakone, N. *et al.* TREX exposes the RNA-binding domain of Nxf1 to enable mRNA export. *Nat. Commun.* **3,** 1006 (2012).

183. Hautbergue, G. M., Hung, M.-L. M.-L., Golovanov, A. P., Lian, L.-Y. & Wilson, S. A. Mutually exclusive interactions drive handover of mRNA from export adaptors to TAP. *Proc. Natl. Acad. Sci. U. S. A.* **105,** 5154–9 (2008).

184. Proudfoot, N. J. Ending the message: poly(A) signals then and now. *Genes Dev.* **25,** 1770–82 (2011).

185. Proudfoot, N. J., Furger, A. & Dye, M. J. Integrating mRNA Processing with Transcription. *Cell* **108,** 501–512 (2002).

186. Eckmann, C. R., Rammelt, C. & Wahle, E. Control of poly(A) tail length. *Wiley Interdiscip. Rev. RNA* **2,** 348–361 (2011).

187. Johnson, S. A., Kim, H., Erickson, B. & Bentley, D. L. The export factor Yra1 modulates mRNA 3' end processing. *Nat. Struct. Mol. Biol.* **18,** 1164–71 (2011).

188. Katahira, J. *et al.* Human TREX component Thoc5 affects alternative polyadenylation site choice by recruiting mammalian cleavage factor I. *Nucleic Acids Res.* **41,** 7060–7072 (2013).

189. Tran, D. D. H. *et al.* THOC5 controls 3′end-processing of immediate early genes via interaction with polyadenylation specific factor 100 (CPSF100). *Nucleic Acids Res.* **42,** 12249–12260 (2014).

190. Lou, H., Neugebauer, K. M., Gagel, R. F. & Berget, S. M. Regulation of alternative polyadenylation by U1 snRNPs and SRp20. *Mol. Cell. Biol.* **18,** 4977–85 (1998).

191. Proudfoot, N. J. Ending the message: poly(A) signals then and now. *Genes Dev.* **25,** 1770–1782

(2011).

192. Berg, M. G. *et al.* U1 snRNP Determines mRNA Length and Regulates Isoform Expression. *Cell* **150,** 53–64 (2012).

193. Kaida, D. *et al.* U1 snRNP protects pre-mRNAs from premature cleavage and polyadenylation. *Nature* **468,** 664–668 (2010).

194. Ruepp, M.-D., Schümperli, D. & Barabino, S. M. L. mRNA 3′ end processing and more-multiple functions of mammalian cleavage factor I-68. *Wiley Interdiscip. Rev. RNA* **2,** 79–91 (2011).

195. Li, Y. *et al.* An intron with a constitutive transport element is retained in a Tap messenger RNA. *Nature* **443,** 234–237 (2006).

196. Segref, A. *et al.* Mex67p, a novel factor for nuclear mRNA export, binds to both poly(A)+ RNA and nuclear pores. *EMBO J.* **16,** 3256–3271 (1997).

197. Grüter, P. *et al.* TAP, the Human Homolog of Mex67p, Mediates CTE-Dependent RNA Export from the Nucleus. *Mol. Cell* **1,** 649–659 (1998).

198. Katahira, J. *et al.* The Mex67p-mediated nuclear mRNA export pathway is conserved from yeast to human. *EMBO J.* **18,** 2593–2609 (1999).

199. Katahira, J. & Yoneda, Y. Roles of the TREX complex in nuclear export of mRNA ND ES OS NO ST ND ES OS NO. *RNA Biol.* **6,** 149–152 (2009).

200. Katahira, J. mRNA export and the TREX complex. *Biochim. Biophys. Acta* **1819,** 507–13 (2012).

201. Golovanov, A. P., Hautbergue, G. M., Tintaru, A. M., Lian, L.-Y. & Wilson, S. A. The solution structure of REF2-I reveals interdomain interactions and regions involved in binding mRNA export factors and RNA. *RNA* **12,** 1933–48 (2006).

202. Okamura, M., Inose, H. & Masuda, S. RNA Export through the NPC in Eukaryotes. *Genes (Basel).* **6,** 124–149 (2015).

203. Wickramasinghe, V. O. *et al.* mRNA Export from Mammalian Cell Nuclei Is Dependent on GANP. *Curr. Biol.* **20,** 25–31 (2010).

204. Jani, D. *et al.* Functional and structural characterization of the mammalian TREX-2 complex that links transcription with nuclear messenger RNA export. *Nucleic Acids Res.* **40,** 4562–73 (2012).

205. Wickramasinghe, V. O., Stewart, M. & Laskey, R. A. GANP enhances the efficiency of mRNA nuclear export in mammalian cells. *Nucleus* **1,** 393–396 (2010).

206. Umlauf, D. *et al.* The human TREX-2 complex is stably associated with the nuclear pore basket. *J. Cell Sci.* **126,** 2656–67 (2013).

207. Wickramasinghe, V. O. *et al.* Selective nuclear export of specific classes of mRNA from mammalian nuclei is promoted by GANP. *Nucleic Acids Res.* **42,** 1–13 (2014).

208. Bachi, A. *et al.* The C-terminal domain of TAP interacts with the nuclear pore complex and promotes export of specific CTE-bearing RNA substrates. *RNA* **6,** 136–58 (2000).

209. Blevins, M. B., Smith, A. M., Phillips, E. M. & Powers, M. A. Complex Formation among the RNA Export Proteins Nup98, Rae1/Gle2, and TAP. *J. Biol. Chem.* **278,** 20979–20988 (2003).

210. Mor, A. *et al.* Dynamics of single mRNP nucleocytoplasmic transport and export through the nuclear pore in living cells. *Nat. Cell Biol.* **12,** 543–552 (2010).

211. Sheinberger, J. & Shav-Tal, Y. The dynamic pathway of nuclear RNA in eukaryotes. *Nucleus* **4,** 195–205 (2013).

212. Oeffinger, M. & Zenklusen, D. To the pore and through the pore: A story of mRNA export kinetics. *Biochim. Biophys. Acta - Gene Regul. Mech.* **1819,** 494–506 (2012).

213. Ben-Ari, Y. *et al.* The life of an mRNA in space and time. *J. Cell Sci.* **123,** (2010).

214. Movement of nuclear poly(A) RNA throughout the interchromatin space in living cells. *Curr. Biol.* **9,** 285–291 (1999).

215. Shav-Tal, Y. *et al.* Dynamics of single mRNPs in nuclei of living cells. *Science* **304,** 1797–800 (2004).

216. Krull, S. *et al.* Protein Tpr is required for establishing nuclear pore-associated zones of heterochromatin exclusion. *EMBO J.* **29,** 1659–1673 (2010).

217. Shibata, S., Matsuoka, Y. & Yoneda, Y. Nucleocytoplasmic transport of proteins and poly(A)+ RNA in reconstituted Tpr-less nuclei in living mammalian cells. *Genes Cells* **7,** 421–34 (2002).

218. Wu, J., Matunis, M. J., Kraemer, D., Blobel, G. & Coutavas, E. Nup358, a cytoplasmically exposed nucleoporin with peptide repeats, Ran-GTP binding sites, zinc fingers, a cyclophilin A homologous domain, and a leucine-rich region. *J. Biol. Chem.* **270,** 14209–13 (1995).

219. Forler, D. *et al.* RanBP2/Nup358 provides a major binding site for NXF1-p15 dimers at the nuclear pore complex and functions in nuclear mRNA export. *Mol. Cell. Biol.* **24,** 1155–67 (2004).

220. Bernad, R., van der Velde, H., Fornerod, M. & Pickersgill, H. Nup358/RanBP2 attaches to the nuclear pore complex via association with Nup88 and Nup214/CAN and plays a supporting role in CRM1-mediated nuclear protein export. *Mol. Cell. Biol.* **24,** 2373–84 (2004).

221. Wickramasinghe, V. O. & Laskey, R. A. Control of mammalian gene expression by selective mRNA export. *Nat. Rev. Mol. Cell Biol.* **16,** 431–442 (2015).

222. Katahira, J., Inoue, H., Hurt, E. & Yoneda, Y. Adaptor Aly and co-adaptor Thoc5 function in the Tap-p15-mediated nuclear export of HSP70 mRNA. *EMBO J.* **28,** 556–67 (2009).

223. Wang, L. *et al.* The THO Complex Regulates Pluripotency Gene mRNA Export and Controls Embryonic Stem Cell Self-Renewal and Somatic Cell Reprogramming. *Cell Stem Cell* **13,** 676–690 (2013).

224. Huang, Y., Gattoni, R., Stévenin, J. & Steitz, J. A. SR splicing factors serve as adapter proteins for TAP-dependent mRNA export. *Mol. Cell* **11,** 837–43 (2003).

225. Long, J. C. & Caceres, J. F. The SR protein family of splicing factors: master regulators of gene expression. *Biochem. J.* **417,** 15–27 (2009).

226. Huang, Y. & Steitz, J. A. SRprises along a Messenger's Journey. *Mol. Cell* **17,** 613–615 (2005).

227. Wickramasinghe, V. O. *et al.* Human Inositol Polyphosphate Multikinase Regulates Transcript-Selective Nuclear mRNA Export to Preserve Genome Integrity. *Mol. Cell* **51,** 737–750 (2013).

228. Smulevitch, S. *et al.* RTE and CTE mRNA export elements synergistically increase expression of unstable, Rev-dependent HIV and SIV mRNAs. *Retrovirology* **3,** 6 (2006).

229. Lindtner, S. *et al.* RNA-binding Motif Protein 15 Binds to the RNA Transport Element RTE and Provides a Direct Link to the NXF1 Export Pathway. *J. Biol. Chem.* **281,** 36915–36928 (2006).

230. Aibara, S., Katahira, J., Valkov, E. & Stewart, M. The principal mRNA nuclear export factor NXF1:NXT1 forms a symmetric binding platform that facilitates export of retroviral CTE-RNA. *Nucleic Acids Res.* **43,** 1883–93 (2015).

231. Kraut-Cohen, J. & Gerst, J. E. Addressing mRNAs to the ER: cis sequences act up! *Trends Biochem. Sci.* **35,** 459–469 (2010).

232. Palazzo, A. F. *et al.* The Signal Sequence Coding Region Promotes Nuclear Export of mRNA. *PLoS Biol.* **5,** e322 (2007).

233. Faria, A. M. C. *et al.* The nucleoporin Nup96 is required for proper expression of interferon-regulated proteins and functions. *Immunity* **24,** 295–304 (2006).

234. Chakraborty, P. *et al.* Nucleoporin Levels Regulate Cell Cycle Progression and Phase-Specific Gene Expression. *Dev. Cell* **15,** 657–667 (2008).

235. Peng, S. S., Chen, C. Y., Xu, N. & Shyu, A. B. RNA stabilization by the AU-rich element binding protein, HuR, an ELAV protein. *EMBO J.* **17,** 3461–70 (1998).

236. Brennan, C. M., Gallouzi, I. E. & Steitz, J. A. Protein ligands to HuR modulate its interaction with target mRNAs in vivo. *J. Cell Biol.* **151,** 1–14 (2000).

237. Barreau, C., Paillard, L. & Osborne, H. B. AU-rich elements and associated factors: are there unifying principles? *Nucleic Acids Res.* **33,** 7138–7150 (2005).

238. Fries, B. *et al.* Analysis of Nucleocytoplasmic Trafficking of the HuR Ligand APRIL and Its Influence on CD83 Expression. *J. Biol. Chem.* **282,** 4504–4515 (2007).

239. Kimura, T., Hashimoto, I., Nagase, T. & Fujisawa, J.-I. CRM1-dependent, but not ARE-mediated, nuclear export of IFN- 1 mRNA. *J. Cell Sci.* **117,** 2259–2270 (2004).

240. Hodge, D. L. *et al.* IFN-gamma AU-rich element removal promotes chronic IFN-gamma expression and autoimmunity in mice. *J. Autoimmun.* **53,** 33–45 (2014).

241. Yang, J., Bogerd, H. P., Wang, P. J., Page, D. C. & Cullen, B. R. Two closely related human nuclear export factors utilize entirely distinct export pathways. *Mol. Cell* **8,** 397–406 (2001).

242. Culjkovic, B., Topisirovic, I., Skrabanek, L., Ruiz-Gutierrez, M. & Borden, K. L. B. B. eIF4E is a central node of an RNA regulon that governs cellular proliferation. *J. Cell Biol.* **175,** 415–26 (2006).

243. Topisirovic, I. *et al.* Molecular dissection of the eukaryotic initiation factor 4E (eIF4E) export-competent RNP. *EMBO J.* **28,** 1087–98 (2009).

244. Culjkovic-Kraljacic, B. *et al.* Combinatorial targeting of nuclear export and translation of RNA inhibits aggressive B-cell lymphomas. *Blood* **127,** 858–868 (2016).

245. Culjkovic, B. & Borden, K. L. Understanding and Targeting the Eukaryotic Translation Initiation Factor eIF4E in Head and Neck Cancer. *J. Oncol.* **2009,** 1–12 (2009).

246. Culjkovic-Kraljacic, B., Baguet, A., Volpon, L., Amri, A. & Borden, K. L. B. The Oncogene eIF4E Reprograms the Nuclear Pore Complex to Promote mRNA Export and Oncogenic Transformation. *Cell Rep.* **2,** 207–215 (2012).

247. Mettenleiter, T. C., Müller, F., Granzow, H. & Klupp, B. G. The way out: what we know and do not know about herpesvirus nuclear egress. *Cell. Microbiol.* **15,** 170–178 (2013).

248. Mettenleiter, T. C. Herpesvirus assembly and egress. *J. Virol.* **76,** 1537–47 (2002).

249. Speese, S. D. *et al.* Nuclear Envelope Budding Enables Large Ribonucleoprotein Particle Export during Synaptic Wnt Signaling. *Cell* **149,** 832–846 (2012).

250. Hatch, E. & Hetzer, M. Breaching the nuclear envelope in development and disease. *J. Cell Biol.* **205,** 133–141 (2014).

251. Montpetit, B. & Weis, K. Cell biology. An alternative route for nuclear mRNP export by membrane budding. *Science* **336,** 809–10 (2012).

252. Samji, T. Influenza A: understanding the viral life cycle. *Yale J. Biol. Med.* **82,** 153–9 (2009).

253. von Appen, A. & Beck, M. Structure Determination of the Nuclear Pore Complex with Three-Dimensional Cryo electron Microscopy. *J. Mol. Biol.* **428,** 2001–10 (2016).

254. Gao, S. *et al.* Characteristics of Nucleocytoplasmic Transport of H1N1 Influenza A Virus Nuclear Export Protein. *J. Virol.* **88,** 7455–7463 (2014).

255. Mühlbauer, D. *et al.* Influenza Virus-Induced Caspase-Dependent Enlargement of Nuclear Pores Promotes Nuclear Export of Viral Ribonucleoprotein Complexes. *J. Virol.* **89,** 6009–6021 (2015).

256. Rivas, H. G., Schmaling, S. K. & Gaglia, M. M. Shutoff of host gene expression in influenza A virus and herpesviruses: Similar mechanisms and common themes. *Viruses* **8,** 1–26 (2016).

257. Bercovich-Kinori, A. *et al.* A systematic view on influenza induced host shutoff. *Elife* **5,** 1–12 (2016).

258. Yarbrough, M. L., Mata, M. A., Sakthivel, R. & Fontoura, B. M. A. Viral Subversion of Nucleocytoplasmic Trafficking. *Traffic* **15,** 127–140 (2014).

259. Krug, R. M. Functions of the influenza A virus NS1 protein in antiviral defense. *Curr. Opin. Virol.* **12,** 1–6 (2015).

260. Hale, B. G., Randall, R. E., Ortin, J. & Jackson, D. The multifunctional NS1 protein of influenza A viruses. *J. Gen. Virol.* **89,** 2359–2376 (2008).

261. García-Sastre, A. *et al.* Influenza A virus lacking the NS1 gene replicates in interferon-deficient systems. *Virology* **252,** 324–30 (1998).

262. Egorov, A. *et al.* Transfectant influenza A viruses with long deletions in the NS1 protein grow efficiently in Vero cells. *J. Virol.* **72,** 6437–41 (1998).

263. Nemeroff, M. E., Barabino, S. M., Li, Y., Keller, W. & Krug, R. M. Influenza virus NS1 protein interacts with the cellular 30 kDa subunit of CPSF and inhibits 3'end formation of cellular pre-mRNAs. *Mol. Cell* **1,** 991–1000 (1998).

264. Chen, Z., Li, Y. & Krug, R. M. Influenza A virus NS1 protein targets poly(A)-binding protein II of the cellular 3'-end processing machinery. *EMBO J.* **18,** 2273–2283 (1999).

265. Satterly, N. *et al.* Influenza virus targets the mRNA export machinery and the nuclear pore complex. *Proc. Natl. Acad. Sci.* **104,** 1853–1858 (2007).

266. Kuss, S. K., Mata, M. a, Zhang, L. & Fontoura, B. M. a. Nuclear imprisonment: viral strategies to arrest host mRNA nuclear export. *Viruses* **5,** 1824–49 (2013).

267. Kainov, D. E. *et al.* Differential Effects of NS1 Proteins of Human Pandemic H1N1/2009, Avian Highly Pathogenic H5N1, and Low Pathogenic H5N2 Influenza A Viruses on Cellular Pre-mRNA Polyadenylation and mRNA Translation. *J. Biol. Chem.* **286,** 7239–7247 (2011).

268. Hale, B. G. *et al.* Inefficient Control of Host Gene Expression by the 2009 Pandemic H1N1 Influenza A Virus NS1 Protein. *J. Virol.* **84,** 6909–6922 (2010).

269. Kochs, G., García-Sastre, A. & Martínez-Sobrido, L. Multiple anti-interferon actions of the influenza A virus NS1 protein. *J. Virol.* **81,** 7011–21 (2007).

270. Ayllon, J. *et al.* A Single Amino Acid Substitution in the Novel H7N9 Influenza A Virus NS1 Protein Increases CPSF30 Binding and Virulence. *J. Virol.* **88,** 12146–12151 (2014).

271. Twu, K. Y., Kuo, R.-L., Marklund, J. & Krug, R. M. The H5N1 Influenza Virus NS Genes Selected after 1998 Enhance Virus Replication in Mammalian Cells. *J. Virol.* **81,** 8112–8121 (2007).

272. Poon, L. L., Pritlove, D. C., Fodor, E. & Brownlee, G. G. Direct evidence that the poly(A) tail

of influenza A virus mRNA is synthesized by reiterative copying of a U track in the virion RNA template. *J. Virol.* **73,** 3473–6 (1999).

273. Firth, A. E. *et al.* Ribosomal frameshifting used in influenza A virus expression occurs within the sequence UCC_UUU_CGU and is in the +1 direction. *Open Biol.* **2,** 120109 (2012).

274. Jagger, B. W. *et al.* An Overlapping Protein-Coding Region in Influenza A Virus Segment 3 Modulates the Host Response. *Science (80-.).* **337,** 199–204 (2012).

275. Dias, A. *et al.* The cap-snatching endonuclease of influenza virus polymerase resides in the PA subunit. *Nature* **458,** 914–918 (2009).

276. Yuan, P. *et al.* Crystal structure of an avian influenza polymerase PAN reveals an endonuclease active site. *Nature* **458,** 909–913 (2009).

277. Oishi, K., Yamayoshi, S. & Kawaoka, Y. Mapping of a Region of the PA-X Protein of Influenza A Virus That Is Important for Its Shutoff Activity. *J. Virol.* **89,** 8661–8665 (2015).

278. Sun, Y. *et al.* Twenty amino acids at the C-terminus of PA-X are associated with increased influenza A virus replication and pathogenicity. *J. Gen. Virol.* **96,** 2036–2049 (2015).

279. Hayashi, T., MacDonald, L. A. & Takimoto, T. Influenza A Virus Protein PA-X Contributes to Viral Growth and Suppression of the Host Antiviral and Immune Responses. *J. Virol.* **89,** 6442–6452 (2015).

280. Gao, H. *et al.* The contribution of PA-X to the virulence of pandemic 2009 H1N1 and highly pathogenic H5N1 avian influenza viruses. *Sci. Rep.* **5,** 8262 (2015).

281. Hu, J. *et al.* PA-X Decreases the Pathogenicity of Highly Pathogenic H5N1 Influenza A Virus in Avian Species by Inhibiting Virus Replication and Host Response. *J. Virol.* **89,** 4126–4142 (2015).

282. Khaperskyy, D. A., Schmaling, S., Larkins-Ford, J., McCormick, C. & Gaglia, M. M. Selective Degradation of Host RNA Polymerase II Transcripts by Influenza A Virus PA-X Host Shutoff Protein. *PLOS Pathog.* **12,** e1005427 (2016).

283. Khaperskyy, D. A. *et al.* Influenza A Virus Host Shutoff Disables Antiviral Stress-Induced Translation Arrest. *PLoS Pathog.* **10,** e1004217 (2014).

284. von Kobbe C *et al.* Vesicular stomatitis virus matrix protein inhibits host cell gene expression by targeting the nucleoporin Nup98. *Mol. Cell* **6,** 1243–52 (2000).

285. Faria, P. A. *et al.* VSV Disrupts the Rae1/mrnp41 mRNA Nuclear Export Pathway. *Mol. Cell* **17,** 93–102 (2005).

286. Quan, B., Seo, H.-S., Blobel, G. & Ren, Y. Vesiculoviral matrix (M) protein occupies nucleic acid binding site at nucleoporin pair (Rae1•Nup98). doi:10.1073/pnas.1409076111

287. Petersen, J. M., Her, L. S., Varvel, V., Lund, E. & Dahlberg, J. E. The matrix protein of vesicular stomatitis virus inhibits nucleocytoplasmic transport when it is in the nucleus and associated with nuclear pore complexes. *Mol. Cell. Biol.* **20,** 8590–601 (2000).

288. Castello, 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).

289. Enninga, J., Levy, D. E., Blobel, G. & Fontoura, B. M. A. Role of Nucleoporin Induction in Releasing an mRNA Nuclear Export Block. *Science (80-.).* **295,** 1523–1525 (2002).

290. Rajani, K. R. *et al.* Complexes of Vesicular Stomatitis Virus Matrix Protein with Host Rae1 and Nup98 Involved in Inhibition of Host Transcription. *PLoS Pathog.* **8,** e1002929 (2012).

291. Kalverda, B., Pickersgill, H., Shloma, V. V. & Fornerod, M. Nucleoporins Directly Stimulate Expression of Developmental and Cell-Cycle Genes Inside the Nucleoplasm. *Cell* **140,** 360–371 (2010).

292. Capelson, M. *et al.* Chromatin-Bound Nuclear Pore Components Regulate Gene Expression in Higher Eukaryotes. *Cell* **140,** 372–383 (2010).

293. Blower, M. D., Nachury, M., Heald, R. & Weis, K. A Rae1-Containing Ribonucleoprotein Complex Is Required for Mitotic Spindle Assembly. *Cell* **121,** 223–234 (2005).

294. Cross, M. K. & Powers, M. A. Nup98 regulates bipolar spindle assembly through association with microtubules and opposition of MCAK. *Mol. Biol. Cell* **22,** 661–672 (2011).

295. Chakraborty, P. *et al.* Vesicular stomatitis virus inhibits mitotic progression and triggers cell death. *EMBO Rep.* **10,** 1154–1160 (2009).

296. Hastie, E. & Grdzelishvili, V. Z. Vesicular stomatitis virus as a flexible platform for oncolytic virotherapy against cancer. *J. Gen. Virol.* **93,** 2529–2545 (2012).

297. te Velthuis, A. J. W. & Fodor, E. Influenza virus RNA polymerase: insights into the mechanisms of viral RNA synthesis. *Nat. Rev. Microbiol.* **14,** 479–493 (2016).

298. Resa-Infante, P., Jorba, N., Coloma, R. & Ortin, J. The influenza virus RNA synthesis machine: advances in its structure and function. *RNA biology* **8,** 207–215 (2011).

299. Yamanaka, K., Ishihama, A. & Nagata, K. Reconstitution of influenza virus RNA-nucleoprotein complexes structurally resembling native viral ribonucleoprotein cores. *J. Biol. Chem.* **265,** 11151–5 (1990).

300. Baudin, F., Bach, C., Cusack, S. & Ruigrok, R. W. Structure of influenza virus RNP. I. Influenza virus nucleoprotein melts secondary structure in panhandle RNA and exposes the bases to the solvent. *EMBO J.* **13,** 3158–65 (1994).

301. Chan, W.-H. *et al.* Functional analysis of the influenza virus H5N1 nucleoprotein tail loop reveals amino acids that are crucial for oligomerization and ribonucleoprotein activities. *J. Virol.* **84,** 7337–45 (2010).

302. Honda, A., Uéda, K., Nagata, K. & Ishihama, A. RNA polymerase of influenza virus: role of NP in RNA chain elongation. *J. Biochem.* **104,** 1021–6 (1988).

303. Resa-Infante, P., Recuero-Checa, M. A., Zamarreño, N., Llorca, O. & Ortín, J. Structural and functional characterization of an influenza virus RNA polymerase-genomic RNA complex. *J. Virol.* **84,** 10477–87 (2010).

304. Lee, M. T. M. *et al.* Definition of the minimal viral components required for the initiation of unprimed RNA synthesis by influenza virus RNA polymerase. *Nucleic Acids Res.* **30,** 429–38 (2002).

305. Newcomb, L. L. *et al.* Interaction of the Influenza A Virus Nucleocapsid Protein with the Viral RNA Polymerase Potentiates Unprimed Viral RNA Replication. *J. Virol.* **83,** 29–36 (2009).

306. Zhang, S. *et al.* Biochemical and kinetic analysis of the influenza virus RNA polymerase purified from insect cells. *Biochem. Biophys. Res. Commun.* **391,** 570–574 (2010).

307. Turrell, L., Lyall, J. W., Tiley, L. S., Fodor, E. & Vreede, F. T. The role and assembly mechanism of nucleoprotein in influenza A virus ribonucleoprotein complexes. *Nat. Commun.* **4,** (2013).

308. Kawaguchi, A., Momose, F. & Nagata, K. Replication-coupled and host factor-mediated encapsidation of the influenza virus genome by viral nucleoprotein. *J. Virol.* **85,** 6197–204 (2011).

309. Vreede, F. T., Jung, T. E. & Brownlee, G. G. Model suggesting that replication of influenza virus

is regulated by stabilization of replicative intermediates. *J. Virol.* **78,** 9568–72 (2004).

310. Ng, A. K.-L. *et al.* Structure of the influenza virus A H5N1 nucleoprotein: implications for RNA binding, oligomerization, and vaccine design. *FASEB J.* **22,** 3638–3647 (2008).

311. Ye, Q., Krug, R. M. & Tao, Y. J. The mechanism by which influenza A virus nucleoprotein forms oligomers and binds RNA. *Nature* **444,** 1078–1082 (2006).

312. Turrell, L., Hutchinson, E. C., Vreede, F. T. & Fodor, E. Regulation of influenza A virus nucleoprotein oligomerization by phosphorylation. *J. Virol.* **89,** 1452–5 (2015).

313. Mondal, A. *et al.* Phosphorylation at the Homotypic Interface Regulates Nucleoprotein Oligomerization and Assembly of the Influenza Virus Replication Machinery. *PLOS Pathog.* **11,** e1004826 (2015).

314. Boulo, S. *et al.* Human importin alpha and RNA do not compete for binding to influenza A virus nucleoprotein. *Virology* **409,** 84–90 (2011).

315. Chenavas, S. *et al.* Monomeric Nucleoprotein of Influenza A Virus. *PLoS Pathog.* **9,** e1003275 (2013).

316. Moradpour, D., Penin, F. & Rice, C. M. Replication of hepatitis C virus. *Nat. Rev. Microbiol.* **5,** 453–463 (2007).

317. Paul, D., Hoppe, S., Saher, G., Krijnse-Locker, J. & Bartenschlager, R. Morphological and biochemical characterization of the membranous hepatitis C virus replication compartment. *J. Virol.* **87,** 10612–27 (2013).

318. Romero-Brey, I. *et al.* Three-Dimensional Architecture and Biogenesis of Membrane Structures Associated with Hepatitis C Virus Replication. *PLoS Pathog.* **8,** e1003056 (2012).

319. Lindenbach, B. D. & Rice, C. M. Unravelling hepatitis C virus replication from genome to function. *Nat. 2005 4367053* (2005).

320. Cheng, J. C., Chang, M. F. & Chang, S. C. Specific interaction between the hepatitis C virus NS5B RNA polymerase and the 3' end of the viral RNA. *J. Virol.* **73,** 7044–7049 (1999).

321. O'Farrell, D., Trowbridge, R., Rowlands, D. & Jäger, J. Substrate complexes of hepatitis C virus RNA polymerase (HC-J4): Structural evidence for nucleotide import and De-novo initiation. *J. Mol. Biol.* **326,** 1025–1035 (2003).

322. Huang, Y., Staschke, K., De Francesco, R. & Tan, S.-L. Phosphorylation of hepatitis C virus NS5A nonstructural protein: A new paradigm for phosphorylation-dependent viral RNA replication? *Virology* **364,** 1–9 (2007).

323. Hughes, M., Griffin, S. & Harris, M. Domain III of NS5A contributes to both RNA replication and assembly of hepatitis C virus particles. *J. Gen. Virol.* **90,** 1329–1334 (2009).

324. Macdonald, A. & Harris, M. Hepatitis C virus NS5A: tales of a promiscuous protein. *J. Gen. Virol.* **85,** 2485–2502 (2004).

325. Masaki, T. *et al.* Interaction of Hepatitis C Virus Nonstructural Protein 5A with Core Protein Is Critical for the Production of Infectious Virus Particles. *J. Virol.* **82,** 7964–7976 (2008).

326. Tellinghuisen, T. L., Foss, K. L. & Treadaway, J. Regulation of hepatitis C virion production via phosphorylation of the NS5A protein. *PLoS Pathog.* **4,** e1000032 (2008).

327. Ascher, D. B. *et al.* Potent hepatitis C inhibitors bind directly to NS5A and reduce its affinity for RNA. *Sci. Rep.* **4,** (2014).

328. Lim, P. J. *et al.* Correlation between NS5A Dimerization and Hepatitis C Virus Replication *.

(2012). doi:10.1074/jbc.M112.376822

329. Toshana, *et al.* All Three Domains of the Hepatitis C Virus Nonstructural NS5A Protein Contribute to RNA Binding. *J. Virol.* **84,** 9267–9277 (2010).

330. Huang, L. *et al.* Hepatitis C Virus Nonstructural Protein 5A (NS5A) Is an RNA-binding Protein *. (2005). doi:10.1074/jbc.M508175200

331. Love, R. A., Brodsky, O., Hickey, M. J., Wells, P. A. & Cronin, C. N. Crystal structure of a novel dimeric form of NS5A domain I protein from hepatitis C virus. *J. Virol.* **83,** 4395–403 (2009).

332. Tellinghuisen, T. L., Marcotrigiano, J. & Rice, C. M. Structure of the zinc-binding domain of an essential component of the hepatitis C virus replicase. doi:10.1038/nature03580

333. Quinkert, D., Bartenschlager, R. & Lohmann, V. Quantitative analysis of the hepatitis C virus replication complex. *J. Virol.* **79,** 13594–605 (2005).

334. Miyanari, Y. *et al.* Hepatitis C virus non-structural proteins in the probable membranous compartment function in viral genome replication. *J. Biol. Chem.* **278,** 50301–8 (2003).

335. Neufeldt, C. J. *et al.* The Hepatitis C Virus-Induced Membranous Web and Associated Nuclear Transport Machinery Limit Access of Pattern Recognition Receptors to Viral Replication Sites. *PLoS Pathog.* **12,** (2016).

336. Appel, N. *et al.* Essential Role of Domain III of Nonstructural Protein 5A for Hepatitis C Virus Infectious Particle Assembly. *PLoS Pathog.* **4,** e1000035 (2008).

337. Kim, Y. K., Kim, C. S., Lee, S. H. & Jang, S. K. Domains I and II in the 5′ Nontranslated Region of the HCV Genome Are Required for RNA Replication. *Biochem. Biophys. Res. Commun.* **290,** 105–112 (2002).

338. Alaoui-Lsmaili, M. H. *et al.* The hepatitis C virus NS5B RNA-dependent RNA polymerase activity and susceptibility to inhibitors is modulated by metal cations. *J. Hum. Virol.* **3,** 306–16

339. Quezada, E. M. & Kane, C. M. The Hepatitis C Virus NS5A Stimulates NS5B During In Vitro RNA Synthesis in a Template Specific Manner. *Open Biochem. J.* **3,** 39–48 (2009).

340. Ross-Thriepland, D. & Harris, M. Insights into the complexity and functionality of hepatitis C virus NS5A phosphorylation. *J. Virol.* **88,** 1421–32 (2014).

341. Masaki, T. *et al.* Involvement of hepatitis C virus NS5A hyperphosphorylation mediated by casein kinase I-α in infectious virus production. *J. Virol.* **88,** 7541–55 (2014).

342. Lemay, K. L., Treadaway, J., Angulo, I. & Tellinghuisen, T. L. A hepatitis C virus NS5A phosphorylation site that regulates RNA replication. *J. Virol.* **87,** 1255–60 (2013).

343. Ross-Thriepland, D., Mankouri, J. & Harris, M. Serine phosphorylation of the hepatitis C virus NS5A protein controls the establishment of replication complexes. *J. Virol.* **89,** 3123–35 (2015).

344. Ohno, M., Kataoka, N. & Shimura, Y. A nuclear cap binding protein from HeLa cells. *Nucleic Acids Res.* **18,** 6989–6995 (1990).

345. Kittler, R. *et al.* Genome-scale RNAi profiling of cell division in human tissue culture cells. *Nat. Cell Biol.* **9,** 1401–12 (2007).

346. Luo, B. *et al.* Highly parallel identification of essential genes in cancer cells. *Proc. Natl. Acad. Sci. U. S. A.* **105,** 20380–5 (2008).

347. Cheung, H. W. *et al.* Systematic investigation of genetic vulnerabilities across cancer cell lines reveals lineage-specific dependencies in ovarian cancer. *Proc. Natl. Acad. Sci.* **108,** 12372–12377

(2011).

348. Mazza, C., Segref, A., Mattaj, I. W. & Cusack, S. Large-scale induced fit recognition of an m(7)GpppG cap analogue by the human nuclear cap-binding complex. *EMBO J.* **21,** 5548–57 (2002).

349. Mazza, C., Ohno, M., Segref, A., Mattaj, I. W. & Cusack, S. Crystal Structure of the Human Nuclear Cap Binding Complex. *Mol. Cell* **8,** 383–396 (2001).

350. Monecke, T., Schell, S., Dickmanns, A. & Ficner, R. Crystal structure of the RRM domain of poly(A)-specific ribonuclease reveals a novel m(7)G-cap-binding mode. *J. Mol. Biol.* **382,** 827–834 (2008).

351. Bian, Y. *et al.* An enzyme assisted RP-RPLC approach for in-depth analysis of human liver phosphoproteome. *J. Proteomics* **96,** 253–262 (2014).

352. Rigbolt, K. T. G. *et al.* System-Wide Temporal Characterization of the Proteome and Phosphoproteome of Human Embryonic Stem Cell Differentiation. *Sci. Signal.* **4,** rs3-rs3 (2011).

353. Olsen, J. V. *et al.* Quantitative Phosphoproteomics Reveals Widespread Full Phosphorylation Site Occupancy During Mitosis. *Sci. Signal.* **3,** ra3-ra3 (2010).

354. Cantin, G. T. *et al.* Combining Protein-Based IMAC, Peptide-Based IMAC, and MudPIT for Efficient Phosphoproteomic Analysis. *J. Proteome Res.* **7,** 1346–1351 (2008).

355. Zhou, H. *et al.* Toward a Comprehensive Characterization of a Human Cancer Cell Phosphoproteome. *J. Proteome Res.* **12,** 260–271 (2013).

356. Mayya, V. *et al.* Quantitative Phosphoproteomic Analysis of T Cell Receptor Signaling Reveals System-Wide Modulation of Protein-Protein Interactions. *Sci. Signal.* **2,** ra46-ra46 (2009).

357. Dephoure, N. *et al.* A quantitative atlas of mitotic phosphorylation. *Proc. Natl. Acad. Sci.* **105,** 10762–10767 (2008).

358. Olsen, J. V. *et al.* Global, In Vivo, and Site-Specific Phosphorylation Dynamics in Signaling Networks. *Cell* **127,** 635–648 (2006).

359. Young-Xu, Y., van Aalst, R., Russo, E., Lee, J. K. H. & Chit, A. The Annual Burden of Seasonal Influenza in the US Veterans Affairs Population. *PLoS One* **12,** e0169344 (2017).

360. Molinari, N.-A. M. *et al.* The annual impact of seasonal influenza in the US: Measuring disease burden and costs. *Vaccine* **25,** 5086–5096 (2007).

361. Plass, D. *et al.* Trends in disease burden in Germany: results, implications and limitations of the Global Burden of Disease study. *Dtsch. Arztebl. Int.* **111,** 629–38 (2014).

362. Ivashkiv, L. B. & Donlin, L. T. Regulation of type I interferon responses. *Nat. Rev. Immunol.* **14,** 36–49 (2013).

363. Quicke, K. M., Diamond, M. S. & Suthar, M. S. Negative regulators of the RIG-I-like receptor signaling pathway. *Eur. J. Immunol.* **47,** 615–628 (2017).

364. Mogensen, T. H. Pathogen recognition and inflammatory signaling in innate immune defenses. *Clinical Microbiology Reviews* **22,** 240–273 (2009).

365. Cao, X. Self-regulation and cross-regulation of pattern-recognition receptor signalling in health and disease. *Nature Reviews Immunology* **16,** 35–50 (2016).

366. WHO | World Health Organization. *WHO* (2018).

367. Origa, R. *et al.* Treatment of hepatitis C virus infection with direct-acting antiviral drugs is safe and effective in patients with hemoglobinopathies. *Am. J. Hematol.* **92,** 1349–1355 (2017).

368. Choo, Q. L. *et al.* Genetic organization and diversity of the hepatitis C virus. *Proc. Natl. Acad. Sci. U. S. A.* **88,** 2451–5 (1991).

369. Gitto, S., Gamal, N. & Andreone, P. NS5A inhibitors for the treatment of hepatitis C infection. *J. Viral Hepat.* **24,** 180–186 (2017).

ABBREVIATIONS

Selected protein abbreviations are included in the list.

	4ESE	4E-sensetive element
A	AAs	Amino acids
	ALREX	Alternative mRNA export
	ALYREF	THO complex subunit 4
	APA	Alternative polyadenylation
	AREs	AU-rich elements
	ASAP	Apoptosis and splicing associated protein complex
	ATP	Adenosine triphosphate
C	CBC	Cap-binding complex
	CE	Capping enzyme
	CPA	Cleavage and polyadenylation
	CPSF	Cellular and polyadenylation specific factor
	CRM1	Chromosome region maintenance 1
	cRNA	complementary RNA
	CTD	C-terminal domain
	CTE	Constitutive transport element
D	Da	Dalton
	DAA	Directly acting antiviral
	DNA	Deoxyribonucleic acid
	ds	Double-stranded
E	eIF	Eukaryotic translation initiation factor
	EJC	Exon junction complex
	ER	Endoplasmic reticulum
F	FG	Phenylalanine-glycine
G	GAP	GTPase activating protein
	GDP	Guanosine diphosphate
	GEF	Guanine nucleotide exchange factor
	Gp	Guanosine monophosphat
	GTase	RNA guanylyltransferase
	GTP	Guanosine triphosphate
H	HCV	Hepatitis C virus

	hnRNP	Heterogeneous nuclear ribonucleoprotein particles
I	IAV	Influenza A virus
	IFN	Interferon
	IL6	Interleukin 6
	IP_6	Inositol hexakiphosphate
	IPMK	Inositol polyphosphate multikinase
	IRES	Internal ribosomal entry site
	ITE	Intronless transcript element
L	LCS	Low-complexity regions
	LRR	Leucine-rich repeat
M	M	Matrix
	m^7G	N7-methylguanosine
	miRNA	Micro RNA
	mRNA	Messenger RNA
	mRNP	Messenger ribonucleoprotein particles
N	N7MTase	Guanine-N7 methyltransferase
	NCBP	Nuclear cap-binding protein
	NCR	Non-coding region
	NE	Nuclear envelope
	NES	Nuclear export signal
	NLS	Nuclear localization signal
	NMD	Nonsense-mediated decay
	NP	Nucleoprotein
	NPC	Nuclear pore complex
	NS1	Non-structural protein 1
	NS5A	Non-structural 5A protein
	nt	Nucleotide
	Nup	Nucleoporins
	NXF1	Nuclear RNA export factor 1
	NXT1	NTF2-related export protein 1
O	ORF	Open reading frame
P	PABP	Poly(A)-binding protein
	PAMPs	Pathogen associated molecular pattern
	PAP	Poly(A) polymerase
	PARN	Poly(A)-specific ribonuclease

	PAS	Poly(A) signal
	PHAX	Phosphorylated adapter RNA export protein
	pi	Primary
	Poly(A)	Polyadenylated
	PR8	Mouse adapted laboratory influenza strain A/PuertoRico/8/1934
	pre	Precursor
	pri	Primary
	PRR	Pattern recognition receptor
R	RBD	RNA-binding domain
	RdRp	RNA-dependent RNA polymerase
	RISC	RNA-induced silencing
	RNA	Ribonucleic acid
	RNGTT	RNA guanylyltransferase and 5′triphosphatase
	RNP	Ribonucleoprotein particle
	RRM	RNA recognition motif
	rRNA	Ribosomal RNA
	rRNP	Ribosomal ribonucleoprotein particles
	RTE	RNA transport element
	RTPase	RNA triphosphatase
S	S	Svedberg unit
	shRNA	short hairpinRNA
	siRNA	Small interfering RNA
	Sm	Sphingomyelin
	SMN	Survival of motor neuron
	snoRNA	Small nucleolar RNA
	snRNA	Small nuclear RNA
	snRNP	Small nuclear ribonucleoprotein particles
	SR	Serine-arginine
	ss	Single-stranded
	SSCR	Signal sequence-coding region
T	TGS1	Trimethylguanosine synthase
	TMG	Trimethylguanosine
	TREX	Transcription export
	tRNA	Transfer RNA
	TUTase	Terminal uridylyl transferase

U	U/G	Uridylate- and guanylate-rich
	UBA	Ubiquitin-associated
	UTR	Untranslated region
V	vRNA	Viral RNA
	vRNP	Viral ribonucleoprotein particles
	VSV	Vesicular stomatitis virus
X	XPO	Exportin

ACKNOWLEDGMENT

I would like to express my gratitude to Prof. Dr. Elena Conti who supervised my Ph.D. thesis. Thank you for your interest in the project, critical comments and your supporting attitude.

Furthermore, I would like to thank my direct supervisor Prof. Dr. Andreas Pichlmair for the huge opportunity to work in his motivating research group as well as for all his enthusiasm and interest in my project. Thank you personally and scientifically for your support over past years.

Since I could not have done the work on the projects alone, I would like to thank all my collaboration partners for their great work and support in the projects.

A special gratitude I would like to give to Dr. Matthias Habjan, a former lab member and great mentor. I always appreciated your patient and caring personality as well as your intellectual and practical knowledge.

My further appreciation goes to all my "Innate" lab members for their support and interest in my project, fruitful discussion, all the help in the lab and for thesis proof-reading. Thank you for the great time in the lab and all the delicious cakes ☺. I would also like to thank the Mann department for the friendly and supportive environment, a lot of reagents and for all the nice department events.

My dear lab girls – Bea, Cathleen, Angelika, Darya and Renate – the last years would not have been so enjoyable without you. Thank you so much for every day we spend together, for so much fun and making the lab not only a work place but also a place, I loved to go!

My dearest Bea, I am more than grateful that our paths crossed and I can´t imagine any work place in the future without you. You became such a close friend! Along this, I would like to thank my dear school friends, Marina and Sabrina, for all we experienced together, for keeping me harmonious and supporting me in all I`m doing. Additionally, I would like to mention my "UNI-Crew" – I´m so grateful that we studied together and thank you all for your friendships.

My dearly beloved Michi, you know me the best with all my good and not so good sides ☺. I cannot describe how grateful I am for your love and support - thank you so much for sharing your life with me, listening to my problems and complaints, cheering me up after an exhausting day and that you are always there for me!

My deepest gratitude and eternal love goes to my family – my beloved dad and the best sisters in the world – who supported me my whole life and completed me to the person I am today – I owe it all to you and will be forever grateful for your love. Mami, your soul and love were and will always be on my side and help me in all I`m doing!