Dissecting the functional interplay between SARS-CoV-2 viral RNAs and the host proteome

Charakterisierung der funktionalen Interaktionen zwischen SARS-CoV-2 RNA und dem Wirtszellproteom

Doctoral thesis for a doctoral degree at the Graduate School of Life Sciences

Julius-Maximilians-Universität Würzburg

Section Infection and Immunity

submitted by

Sabina Ganskih

from

Mannheim, Germany

Würzburg, 2023

Submitted on: _____

Members of the Thesis Committee

Chairperson: Prof. Dr. Christian Janzen

Primary Supervisor: Jun. Prof. Mathias Munschauer

Supervisor (Second): Prof. Dr. Utz Fischer

Supervisor (Third): Prof. Dr. Lars Dölken

Affidavit

I hereby confirm that my thesis entitled “Dissecting the functional interplay between SARS-CoV-2 viral RNAs and the host proteome” is the result of my own work. I did not receive any help or support from commercial consultants. All sources and / or materials applied are listed and specified in the thesis.

Furthermore, I confirm that this thesis has not yet been submitted as part of another examination process neither in identical nor in similar form.

Würzburg,

Place, Date Signature

Eidesstattliche Erklärung

Hiermit erkläre ich an Eides statt, die Dissertation „Charakterisierung der funktionalen Interaktionen zwischen SARS-CoV-2 RNA und dem Wirtszellproteom“ eigenständig, d.h. insbesondere selbständig und ohne Hilfe eines kommerziellen Promotionsberaters, angefertigt und keine anderen als die von mir angegebenen Quellen und Hilfsmittel verwendet zu haben.

Ich erkläre außerdem, dass die Dissertation weder in gleicher noch in ähnlicher Form bereits in einem anderen Prüfungsverfahren vorgelegen hat.

Würzburg,

Ort, Datum Unterschrift

Table of Content

Summary

The recent pandemic has reminded the public that basic research in virology is pivotal for human health. Understanding the mechanisms of successful viral replication and the role of host factors can help to combat viral infections and prevent future pandemics.

Our lab has published the first SARS-CoV-2 RNA-protein interaction atlas, laying the foundation to investigate the interplay between viral RNA and host RNA binding proteins (RBP). Based on this, my project created the largest collection of binding profiles of host and viral RBPs on SARS-CoV-2 RNA to date. This revealed the host protein SND1 as the first human RBP that specifically binds negative sense viral RNA at the 5´ end, a region associated with viral transcription initiation. The binding profile shares similarities with the viral RBP nsp9, which binds the 5´ ends of positive and negative sense SARS-CoV-2 RNA. Depletion of SND1 shows reduced levels of viral RNA revealing it as a proviral host factor. To decode the underlying molecular mechanism, I characterized the protein-protein interactions of SND1 in SARS-CoV-2 infected and uninfected cells. Infection remodels the protein interactors of SND1 from general RNA biology to membrane association and viral RNA synthesis. Upon infection, SND1 specifically interacts with nsp9, the RBP that shares the same binding region on the negative strand of SARS-CoV-2 RNA. Recent work demonstrates that nsp9 is NMPylated *in vitro* suggesting a functional role of nsp9 in priming of viral RNA synthesis. I was able to show that nsp9 is covalently linked to the 5´ ends of SARS-CoV-2 RNA during infection of human cells. Analysing the covalent bond of nsp9 with the viral RNA on nucleotide level shows close proximity to the initiation sites of viral RNA synthesis, suggesting that nsp9 acts as a protein-primer of SARS-CoV-2 RNA synthesis. SND1 modulates the distribution of nsp9 on the viral RNA, since depletion of SND1 results in imbalanced occupancy of nsp9 at the 5´ends of viral RNA.

This study is the first to provide evidence for the priming mechanism of SARS-CoV-2 in authentic viral replication and further reveals how this mechanism is modulated by the host RBP SND1. Detailed knowledge about priming of viral RNA synthesis can help to find targeted antivirals that could be used to fight coronaviral infections.

Zusammenfassung

Die letzte Pandemie zeigte erneut, das Grundlagenforschung im Bereich der Virologie essentiell für die Gesundheit des Menschen ist. Das Wissen über Schlüsselelemente erfolgreicher viraler Replikation und der Relevanz humaner Proteine darin kann helfen Infektionen zu bekämpfen und künftige Pandemien zu verhindern.

Unser Labor publizierte das erste SARS-CoV-2 RNA Protein-Interaktom und legte dabei den Grundstein für die Forschung am Zwischenspiel viraler RNA und humanen RNA Bindeproteinen (RBPs). Basierend darauf, generierte mein Projekt die bislang größte Sammlung an Bindeprofilen humaner sowie viraler RBPs auf der SARS-CoV-2 RNA. Dabei zeigte sich der Wirtsfaktor SND1 als das erste human RBP das in der Lage ist den Negativstrang der viral RNA zu binden, spezifisch an dessen 5´ Ende welches mit der Transkriptionsinitiierung assoziiert ist. Diese Bindestelle ist ähnlich zu dem viralen RBP nsp9, welches die 5´ Enden der positiv und negativ RNA bindet. Das Fehlen von SND1 in der Wirtszelle führt zu reduzierten Mengen viraler RNA und impliziert daher einen proviralen Einfluss von SND1. Um den zugrundeliegenden molekularen Mechanismus zu verstehen, betrachtete ich die Protein-Protein Interaktionen von SND1 in SARS-CoV-2 infizierten und uninfizierten Zellen. Dabei zeigte sich, dass durch die Infektion die Interaktionspartner von SND1 von genereller RNA Biologie zu Membranassoziierung sowie viraler RNA Synthese verschiebt. Mit Infektion der Zelle interagiert SND1 spezifisch mit nsp9, das RBP welches dieselbe Binderegion am Negativstrang mit SND1 auf der SARS-CoV-2 RNA teilt. Neuste *in vitro* Studien zeigen, dass nsp9 NMPyliert wird und deuten damit eine Relevanz von nsp9 in Priming an. Ich konnte im Kontext authentischer viraler Replikation zeigen, dass nsp9 kovalent an die 5´ Enden der SARS-CoV-2 RNA gebunden ist. Bei genauerer Untersuchung der kovalenten Bindung von nsp9 an der viralen RNA auf Nukleotidebene zeigt, dass diese Nahe der Initiationsstelle der Transkription liegen, was eine Relevanz von nsp9 als Protein-Primer in der SARS-CoV-2 RNA Synthese impliziert. Die Richtige Verteilung von nsp9 auf der viralen RNA wird von SND1 moduliert, da Abwesenheit von SND1 zu einem Ungleichgewicht von nsp9 an den 5´ Enden führt.

Diese Studie ist die Erste, die Evidenzen für den Primingmechanismus von SARS-CoV-2 in authentischer viraler Replikation zeigt und wie diese durch SND1 moduliert wird. Detailliertes Wissen über das Priming viraler RNA Synthese kann dabei helfen gezielte nach antiviralen Substanzen zu suchen, die dabei helfen könnten Infektionen durch Coronaviren zu bekämpfen.

1. Introduction

1.1. Human pathogenic viruses

The wish to understand life-threatening diseases has driven humanity to identify and study their causative agents. The investigation of infectious diseases has led to the discovery of bacteria and subsequently viruses as causative agents of human diseases [1]. By understanding the biology of those pathogens, antimicrobial substances that hinder the growth of these pathogens were found, that help fight human diseases [2-4].

Viruses require a host cell to replicate and produce new virions, which are infectious viral particles. A virion contains nucleic acids that are sheltered by a protective shell made of proteins, called the capsid. Some virus species have an additional lipid membrane -the envelope- surrounding the capsid, which contains proteins relevant to the formation of viral particles and their attachment to the host cell (Figure 1A). While cells use deoxyribonucleic acid (DNA) as carrier of genetic information, viruses use either DNA or ribonucleic acid (RNA). The genome, which is the entirety of encoded genetic information, contains multiple genes that encode protein-coding and noncoding RNA. Noncoding RNA is functional by itself, while protein-coding RNAs, also called messenger RNAs (mRNAs), are translated into proteins by ribosomes [5]. Viral proteins can exert a multitude of functions like viral RNA replication, shutoff of host protein synthesis, or suppression of antiviral responses [6-13]. The genomes of viruses come in multiple forms. Apart from the type of genetic material used to encode their genome (DNA or RNA), the genome can be single-stranded (ss) or double-stranded (ds), linear or circular, segmented or non-segmented, or available for immediate translation (positive sense genome) or as template (negative sense genome) (Figure 1B) [14]. Despite all the differences in the genomic architecture, one fundamental task for a successful viral replication is to generate mRNA, to synthesize viral proteins, and to produce new virus particles.

To order the great variety of existing viruses in a simplified manner, David Baltimore published the first classification system that grouped viruses around the idea that every virus needs to create mRNA to initiate viral replication [15]. Based on the molecular

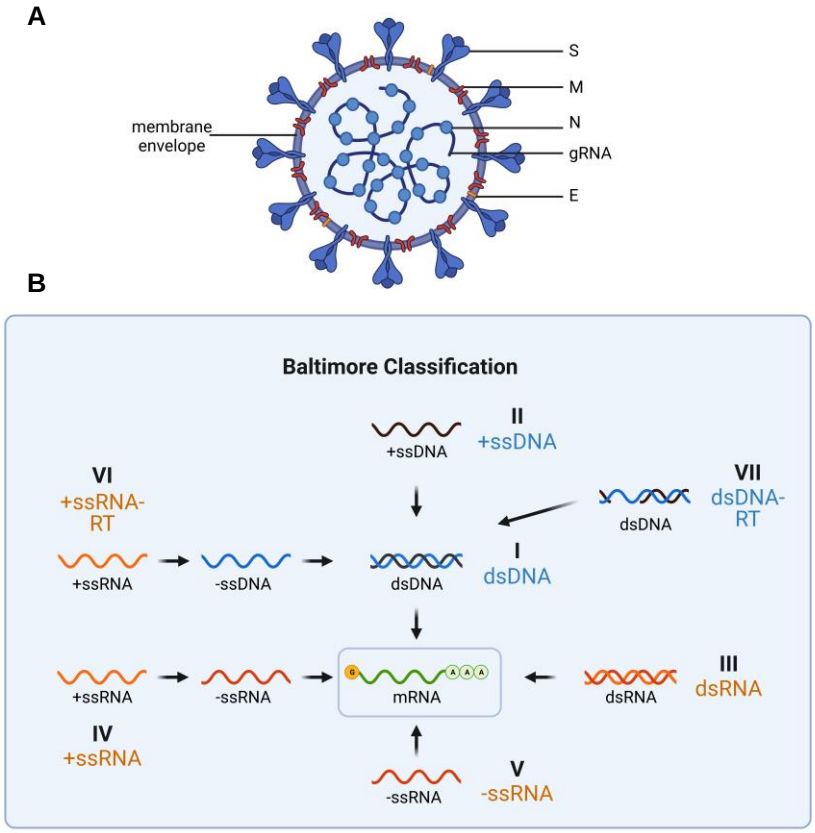

Figure 1 Schematic structure of the SARS-CoV-2 virus particle and the plethora of genome types

A Schematic representation of the SARS-CoV-2 virus particle. The genomic RNA (gRNA) is coated by nucleoprotein (N) and embedded in a lipid bilayer membrane called envelope. The envelope contains the structural proteins envelope (E), membrane (M) and the surface receptor required for host cell infection Spike (S). **B** The Baltimore Classification groups viruses based on the molecular intermediates required to create messenger RNA (mRNA) [14,15]. Created with BioRender.com

intermediates needed for the viral genome to create mRNA, Baltimore proposed six virus classes that upon discovery of the gapped hepadnaviral genome were extended to a seventh (Figure 1B). With the introduction of new technologies, like next-generation sequencing and metagenomics, and combining them with phylogenetics, the relationship between different viruses nowadays is better understood. Based on that newly acquired knowledge, the International Committee on Taxonomy of Viruses (ICTV) introduced a new classification system that is guided by the taxonomic rank used for cellular organisms [16]. In those taxonomic ranks, the species are grouped based on their genetic similarity in a hierarchical manner, with higher ranks representing the oldest shared ancestor. Providing

a clear and comprehensible overview of the necessary steps to initiate replication, the Baltimore classification is still commonly used to complement the virus taxonomy.

As obligatory intracellular pathogens, viruses need the host cell and its biosynthesis machinery, since they encode for a minimal set of genes to replicate. To find which protein-coding genes in the human genome are relevant for viral replication, assays have been developed to study the relevance of each of these genes with parallel screening methods. Using the clustered regularly interspaced short palindromic repeats (CRISPR) - CRISPR associated protein 9 (Cas9) systems, a single gene per cell is disrupted to investigate its relevance in viral replication. These genome-wide CRISPR screens exist for a plethora of human pathogenic viruses revealing the influence of the host proteins on viral replication [17-25]. To favor the production of the viral macromolecules needed to generate new virus particles, viruses use RNA-protein and protein-protein interactions to modulate the function of certain host proteins in favor of viral replication [10, 26-31]. Since viral infections pose a danger to the host, the infected cell elicits an antiviral response to hinder the successful replication of viruses [32-36]. Thus, to maintain the ability to infect the host cell, viruses in turn evolve to circumvent these antiviral pathways, creating a constant thug of war between virus and host [36].

1.2. RNA-viruses

As their name suggests, RNA viruses carry their genetic information on an RNA genome. The genome can be in positive (pos) or negative (neg) sense orientation and single-stranded (ssRNA) or double-stranded (dsRNA). A common trade of each member in this realm is the necessity to encode an RNA-dependent polymerase for successful replication. The realm of *Riboviria* is of special concern to human health since throughout history zoonotic transmissions have been a frequent cause of epidemics and pandemics [37, 38]. Some of the most infamous RNA viruses are the human immunodeficiency virus-1 (HIV-1), influenza A virus (IAV), as well as the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) [37-40]. The impact on health, society, and economy caused by the diseases of those viruses is tremendous and requires a better understanding of their molecular biology. Especially with the recent coronavirus disease 2019 (COVID-19) pandemic, basic research is trying to understand the details of coronaviral replication and

the human immune response to this infection. Especially the high risk of recombination and new zoonotic transmission make coronaviruses an interesting target for basic research [41-43].

1.3. Coronaviridae

Coronaviruses are non-segmented, pos sense RNA viruses containing a membrane envelope and according to the Baltimore classification belong to the virus class IV. When investigating the viral particles under an electron microscope, the spike protein (S) creates a halo that resembles a solar corona, hence the name [43, 44]. The subfamily *Orthocoronavirinae* divides into four genera: *Alpha*-, *Beta*-, *Gamma*-, and *Deltacoronavirus*. While the common genome pool of gamma-and deltacoronaviruses can be found in birds and pigs, alpha-and betacoronaviruses are mostly found in bats [45]. The seven known human coronaviruses (hCoV) are members of alpha- or betacoronaviruses, and genetically close to bat coronaviruses. In recent history, three species of the genus *Betacoronavirus* were transmitted to humans from animal reservoirs, a phenomenon called zoonosis. In 2002/2003, SARS-CoV-1 infected 8,098 people, leading to the severe acute respiratory syndrome (SARS), a disease impacting the lower respiratory tract with a case fatality rate of ~10 % [46, 47]. SARS-CoV-1 is closely related to coronaviruses found in the Chinese horseshoe bat *Rhinolophus sinicus*, which could have evolved to infect humans through palm civets [48-54]. In 2012, the Middle East respiratory syndrome-related coronavirus (MERS-CoV) was the source of an even more lethal coronavirus-caused disease. 2,604 cases and 936 deaths are reported since 2012, resulting in a case fatality rate of ~36 % [55]. While the closest viral ancestor of MERS-CoV is of bat origin, dromedary camels used as livestock are considered the main source of transmission between animals and humans [56, 57]. The latest *Betacoronavirus* spillover was in late 2019 by SARS-CoV-2. Despite the relatively low fatality rate of ~1 %, SARS-CoV-2 has infected 767,726,861 humans causing more deaths than SARS-CoV-1 and MERS-CoV combined (WHO, Jul 11th, 2023) [58, 59]. The exact route through which SARS-CoV-2 was introduced into the human populace remains unsolved [60]. It is quite likely, that multiple severe acute respiratory syndrome-related coronaviruses (SARSr-CoV) through recombination could have provided the genetic basis of SARS-CoV-2 which managed to adapt to humans as new hosts [61]. Even before the SARS-CoV-2 pandemic, experts had

already warned about the zoonotic risk that coronaviruses cause and the necessity to investigate this pathogen [38, 62]. With the risk of future zoonotic transmissions or recombination of existing coronavirus strains, it is pivotal to understand this virus family and find common pathways as targets for antiviral treatments.

1.4. SARS-CoV-2

The genome of SARS-CoV-2 is roughly 30,000 kb in length, making it one of the largest pathogenic RNA viruses (Figure 3A). The viral replicase proteins maintain this genome size through proofreading, preventing the accumulation of deleterious mutations [63, 64]. In total, the genome encodes 31 proteins, which are divided into the non-structural proteins (nsp), the structural proteins and the accessory proteins. Immediately after infection, ribosomes initiate translation on the genomic RNA (gRNA) and synthesize a large polyprotein (pp1a and pp1ab) that auto-proteolytically cleaves into 16 nsps (Figure 2). In principle, a gene is transcribed into an mRNA, which in the form of tri-nucleotide sequences (codons) encodes a protein. The start codon defines the reading frame in which the ribosome translates the codon into the corresponding amino acid. A programmed frameshift can change the reading frame, resulting in the synthesis of an alternative protein. To generate the two separate polyproteins, the ribosome translates pp1ab in the first reading frame. A pseudoknot structure in the gene causes a -1 frameshifting event that changes the reading frame of the ribosome and leads to the synthesis of pp1ab [65]. Belonging to the order of “nested viruses” (*nido* = nested*, virales* = viruses*)*, SARS-CoV-2 produces subgenomic RNAs (sgRNAs) that translate into 15 viral proteins, which are further discriminated into accessory and structural proteins. The role of the accessory proteins ORF3a-d, ORF6, ORF7a-b, ORF8, ORF9b-c, and ORF10 remains poorly understood and often is transferred from other SARSr-CoV. Lastly, the structural proteins nucleocapsid (N), S, membrane (M), and envelope (E) are relevant for viral particle formation. The protein N interacts with viral gRNA as well as with M to package the gRNA into newly formed viral particles [66, 67]. The E protein integrates into host membranes and together with M, regulates the localization of S to the endoplasmic-reticulum-Golgi intermediate compartment (ERGIC) [68].

Since the COVID-19 outbreak in late 2019, SARS-CoV-2 has caused 6,948,764 deaths (WHO, Jul 11th 2023). Reaching an endemic phase as a permanently present infectious agent, SARS-CoV-2 together with other hCoVs will remain a constant burden on health care and economy [58, 59, 69, 70]. Understanding the basic principles of the virus-host interplay on a molecular level will provide new possibilities for targeted antivirals, which will help to reduce the impact hCoVs have on human society.

1.5. Coronaviral replication cycle

When infecting a host cell, coronaviruses use the host's biosynthesis pathways, as well as its nutrients for replication. By estimation from Murine Hepatitis Virus (MHV), a virion during a single replication cycle can multiply 10 to 100 times [52, 71]. While all members of the *Nidovirales* follow the same principle of viral replication, the detailed description of the viral life cycle below is exemplified by SARS-CoV-2 (Figure 2).

1.5.1. Virus attachment and entry

The first step of infection involves interaction between the receptor-binding protein S and the angiotensin-converting enzyme 2 (ACE2), a surface membrane protein of the host cell [40, 72-74]. ACE2 is a highly conserved surface protein among different animal species, which explains the high tropism of SARS-CoV-2 and the risk of transmissions between different species [75, 76]. The S protein consists of two subunits, S1 and S2, which are separated by a furin cleavage site that needs to be cleaved for successful infection [72, 77, 78]. S1 contains the recognition binding domain (RBD) which interacts with ACE2 stabilizing the position of the viral particle on the host cell [79]. Subsequently, the fusion peptide (FP) in S2 is inserted into the cellular membrane, triggering the interaction of the heptapeptide repeat sequence 1 (HR1) and HR2 and the subsequent fusion of the viral and host membranes [80].

For SARS-CoV-2, two separate pathways can activate S through cleavage. The first involves proteolytic cleavage by the transmembrane protease serine 2 (TMPRSS2) at the cell surface [81, 82]. Low levels of TMPRSS2 expression or improper cleavage can lead to clathrin-mediated endocytosis of the virus particle. This alternative route of entry requires cleavage of S2 by cathepsins in the endosomes [82-84]. Both events lead to the fusion of the viral and host cell membranes, releasing the encapsulated viral genome into

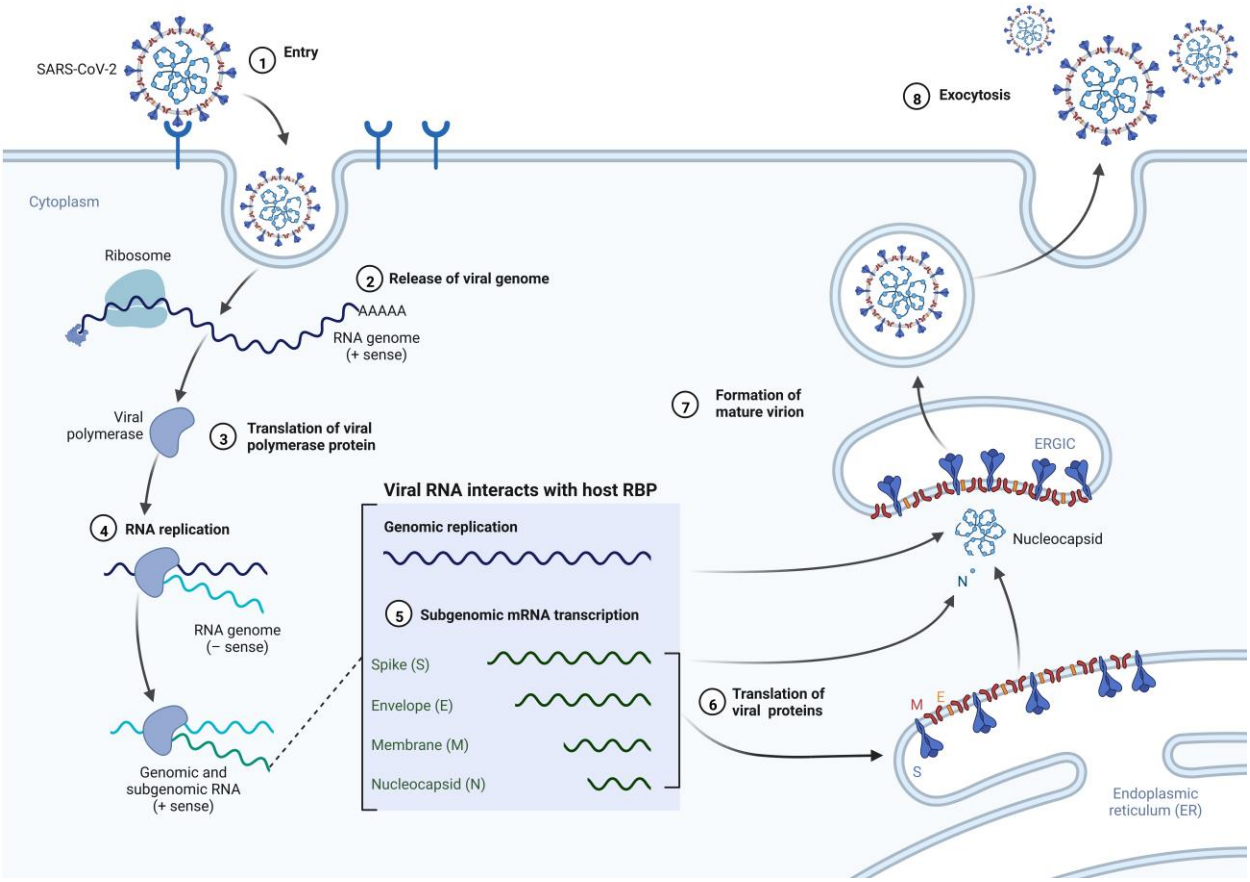

Figure 2 Simplified replication cycle of *Coronaviridae* using SARS-CoV-2 as example

1 After attachment to the host surface receptor, the viral envelope fuses with the host cell membrane **2** to release the viral genome into the cytoplasm. **3** The capped and polyadenylated genomic RNA (gRNA) is directly translated by the ribosomes generating the viral polyproteins pp1a and pp1ab, which encodes for the viral RNA replication-transcription complex (RTC). **4** The RTC synthesizes the viral negative (-) sense gRNA and subgenomic RNA (sgRNA) **5** that serve as templates for the subsequent genomic replication as well as the transcription of viral mRNAs. **6** The structural proteins which are the basic building blocks for new virus particle formation are translated at the endoplasmic reticulum (ER) and transported to the ER-Golgi intermediate compartment (ERGIC). **7** The structural protein nucleocapsid (N) covers the gRNA for packaging at the ERGIC. **8** The fully formed viral particle is released via exocytosis to infect the next host cell. Adapted template from Benjamin Goldman-Israelow (Creator) and Ginny Fulford on BioRender.com. [72, 336-340]

the cytosol. Here, the 5´ capped and polyadenylated (poly(A)) genomic RNA (gRNA) is translated by cellular ribosomes [65].

1.5.2. gRNA translation and formation of replication compartments

After releasing the gRNA, ribosomes start translating the ORF1a gene, encoding for the polyprotein pp1a. The papain-like proteases (PL^{pro}) in nsp3 and the main protease in nsp5 (M^{pro}) cleave the polyprotein into the nonstructural proteins (nsp) 1-11 [85-87]. The same

gene locus contains a slippery sequence, which can lead to a -1 ribosomal frameshift resulting in the translation of pp1ab that cleaves into nsp1-10 and nsp12-16 [65, 88, 89]. The early viral proteins nsp3 and nsp4 are the main factors, which supported by nsp6 form double membrane vesicles (DMV) as the replication platform from the endoplasmic reticulum (ER) [90-92]. At the ER, nsp3 and nsp4 integrate through their transmembrane domains in the ER-membrane and form DMVs as well as convoluted membranes [91]. At the same time, those proteins form a pore complex, which connects the DMV lumen with the cytosol [90-92].

1.5.3. Viral RNA synthesis

The DMVs are the site of viral RNA synthesis, where the RNA-dependent RNA polymerase (RdRp) synthesizes new viral gRNA and sgRNA [93]. The whole replication transcription complex (RTC) consisting of all replicase proteins is encoded in the nsps 2-16. The replicase complex contains the proteins required for the synthesis of viral RNA. The complex is composed of nsp12, nsp7, two nsp8 as well as the capping machinery consisting of nsp13, -14, -16 as well as the cofactor nsp10 [94, 95]. From the gRNA, the replicase complex produces full-length neg sense RNA that serves as a template for the synthesis of new gRNA (Figure 3B). Within the replicase complex, nsp12 acts as the catalytic subunit for viral RNA synthesis, while nsp7 and nsp8 are co-factors that enhance the synthesis [96]. The polymerase domain of nsp12 comprises of finger, palm and thumb domains [97]. At the N-terminus, nsp12 carries another domain that is unique in *Nidovirales*. This N-terminal domain has an essential nucleotidylation function and is termed the nidovirus RdRp-associated nucleotidyltransferase (NiRAN) domain [98]. The NiRAN domain of nsp12 covalently attaches a nucleoside monophosphate (NMP) to nsp9, a process termed NMPylation [95, 99-102]. In this reaction, the NiRAN domain uses nucleoside triphosphate (NTP) as a substrate to transfer the NMP to the primary amine of the nsp9 N-terminus, creating a phosphoramidate bond. A similar mechanism has been observed in poliovirus [103, 104]. The viral infection results into a covalent linkage of the viral protein genome-linked (VPg) to the pos sense 5´end of the viral genome to prime for RNA replication [105]. Currently, it is suggested that in addition to NMPylation, the NiRAN domain is also able to de-NMPylate nsp9, leaving an RNA stretch with the core cap structure, to allow the formation of the cap-0 and the subsequent cap-1 structures through

nsp14 and nsp16, with nsp10 as cofactor [100, 101]. Those studies show mechanistic evidence of NMPylation *in vitro* and suggest the relevance of NMPylated nsp9 in capping which is compatible with the potential priming mechanism of nsp9 [95, 99-102]. The significance of this reaction so far has not been demonstrated *in vivo.*

While it is widely accepted, that the replicase complex is relevant for primed RNA synthesis and elongation, the precise mechanism of initiation is still under debate. Early reports from te Velthuis *et al.* and Imbert *et al.* suggest that nsp7 and nsp8 form a multimeric structure, that *in vitro* can initiate *de novo* priming of neg strand synthesis [106, 107]. Structural analysis of the hexadecameric structure of eight nsp7 and eight nsp8 shows, that it forms a ring-like structure containing a positively charged inner circle, that would be optimal for dsRNA binding [108, 109]. More recent work investigating the nsp7, -8, -12 complex could not confirm the previously suggested hypothesis of *de novo* synthesis [96, 97]. A study investigating the genomic replication in a member of coronaviruses has shown that nsp8 and nsp9 adapt by mutation to compensate for a replication-deficient deletion mutant of the 3´ untranslated region (UTR) [110]. The authors, therefore, suggest a relevance for nsp8 and nsp9 in the initiation of replication through the RNA binding capabilities of nsp9 and the previously assumed primase function of nsp8. For a long time, the role of nsp9 remained elusive claiming to be an unspecific RNA and DNA binding protein [111, 112]. The newest data suggesting NMPylation of nsp9 give further suggestions for a relevance of nsp9 in priming [95, 99-102, 113].

As a member of the *Nidovirales* SARS-CoV-2 shares the common attribute to synthesize a nested set of sgRNAs that share the same 3´ UTR as well as the same 5’ leader sequence (Figure 3) [114-117]. To synthesize this set of sgRNA, the RTC performs discontinuous synthesis, which involves a template switch [118, 119]. Transcription of the neg sense RNA initiates at the 3´ UTR and elongates transcription up to the transcription regulatory sequence (TRS) (Figure 3B). The TRS is located upstream of each gene body (TRS-B) and downstream of the leader sequence (TRS-L) (Figure 3A). These TRS contain a core sequence of 6 nucleotides that are identical and used for template switching [117, 119]. The newly synthesized antisense sequence of TRS-B (anti TRS-B) disassociates from the initial template and reanneals to the TRS-L to continue transcription using the leader sequence as the new template. This template switch results in a shorter neg sense

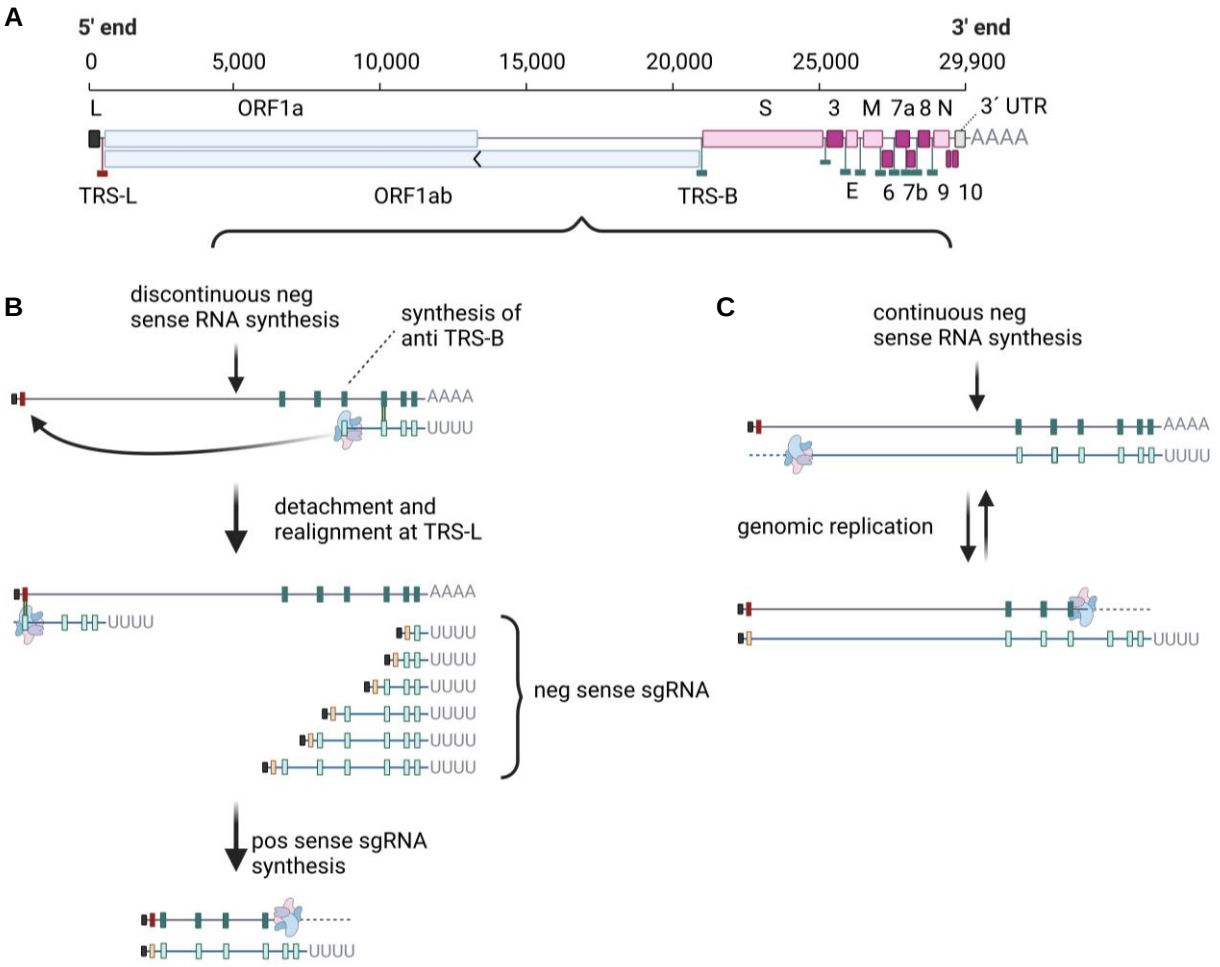

Figure 3 The SARS-CoV-2 genome architecture and the synthesis of viral RNA

A The SARS-CoV-2 genome is roughly 30,000 kb in size and has a 5´ leader sequence (L) and a 3´ untranslated region (3´ UTR), which is shared by all subgenomic RNAs (sgRNA). Upstream of each gene body (B) lies a transcription regulatory sequence (TRS), which has the same sequence as the TRS downstream of the leader sequence. **B** Discontinuous neg sense RNA synthesis leads to the transcription of nested sgRNAs. When the RTC synthesizes the neg strand RNA, it also creates the anti TRS-B, a sequence that base pairs with the TRS-B and the TRS-L. The newly synthesized neg strand RNA detaches from the template RNA and realigns at the TRS-L, which results in a shorter neg strand sgRNA that contains the neg sense sequence of the 5´ terminus. Now, the neg sense sgRNA serves as a template for pos sense sgRNA synthesis to create the viral mRNAs. **C** The replication of new genomic RNA (gRNA) relies on the continuous negative (neg) sense RNA synthesis by the RTC. Created with BioRender.com, adapted from Malone *et al*. (2022) [119]

RNA that serves as a template for the synthesis of accessory and structural mRNAs. Studies using direct RNA sequencing and ribosome profiling have expanded the view on viral SARS-CoV-2 sgRNAs and provided evidence for unprecedented non-canonical sgRNAs [114, 117, 120, 121]. The long-distance interaction between the TRSs that is

required for template switching is most likely supported through regulatory RNA sequences and RNA-protein interactions with viral as well as host proteins [122]. The resulting RNA-protein complex (RPC) could then regulate template switching. How exactly the discontinuous RNA synthesis is regulated remains elusive. In addition to the discontinuous RNA synthesis, the continuous RNA synthesis generates the “anti-genome” that serves as a template for gRNA replication (Figure 3C).

Coronaviral neg sense RNAs serve as a template for the synthesis of mRNAs through the RTC, which are capped by the viral capping machinery. This machinery consists of nsp14, nsp16, and the cofactor nsp10 adding the N7- and 2′-*O*-methylated cap ($^{me7}G_o$pppA$_{1m}$) leaving the cap-1 structure on viral sgRNAs [123-125]. Furthermore, a potential involvement of nsp12 in capping through the covalent addition of NMP to nsp9 has been proposed recently [100, 126]. The final viral mRNA is polyadenylated but the precise mechanisms of polyadenylation remains poorly understood. Research on bovine coronavirus (BCoV) has shown varying lengths of the poly(A) tail during infection [127]. This suggests that the length of the poly(A) tail is regulated by viral proteins, host factors or both [128]. By acquiring a cap and poly(A) tail, the SARS-CoV-2 RNA mimics the host mRNA providing structures to be recognized by host RBPs relevant for translation [31, 129, 130].

When searching for drugs against pathogens, targeting factors unique to the pathogen is beneficial to reduce drug-induced side effects. The synthesis of neg sense RNA from an RNA template is a unique point in the viral replication cycle that is not performed by the host and would therefore be a good targeting point for potential drug designs. Understanding the details of sgRNA synthesis and which viral and host proteins are involved could help to design sophisticated antiviral drugs.

1.5.4. Virus particle assembly and release

Upon synthesis, the structural proteins S, E, and M are initially integrated into the ER membrane. Subsequently, they are transported to the ERGIC, where particle assembly takes place [66-68]. Virus-like particle (VLP) systems have revealed that expression of the structural proteins S, E, and M are sufficient to produce and release VLPs [66, 131-133]. N covers newly synthesized gRNA tightly and through interaction with M condenses the

gRNA for complete particle formation [66, 67]. How exactly viral gRNA is specifically recruited for viral particle formation remains still elusive. While for MHV a genome packaging sequence in the gene locus of nsp15 was found, no similar mechanism has been discovered for SARS-CoV-2 so far [134]. After assembly, the full particles are released by exocytosis, where they spread to initiate a new infection cycle.

1.6. RNA binding proteins

RNA binding proteins (RBPs) interact directly with RNAs through specific modular domains or charged disordered regions, forming a complex of RNA and protein that is referred to as ribonucleoprotein (RNP). Current estimates assume that humans express more than 1,500 RBPs [135, 136]. RBPs interact with nucleic acids using RNA binding domains (RBD) like the RNA recognition motif (RRM) or zinc finger (ZnF) domains [137]. With recent efforts that increased our knowledge about RBDs, it has been shown, that intrinsically disordered regions (IDR) can act as RNA binding modules [135, 138, 139] This interaction between RNA and protein can change the function, structure, or stability of the target [140, 141]. One example specifically for SARS-CoV-2 would be the viral helicase nsp13, which binds the RNA-duplex with a zinc-binding domain and unwinds the dsRNA for replication [142]. Aside from the role of RNA as a carrier of information, RNA can exert an immediate effector function and modulate the function of proteins. An exemplary case in SARS-CoV-2 infected cells would be the leader sequence and its effect on the viral protein nsp1. This protein inhibits the translation of host mRNA by interacting with the 40S ribosomal subunit and induces endonucleolytic cleavage of the mRNA [10]. Since this would also abolish the translation of viral RNAs, the highly structured leader sequence on viral RNAs binds to nsp1 resulting in its disassociations from the 40S ribosomal subunit allowing the translation of viral mRNAs specifically.

The sum of all proteins interacting with a specific RNA, the RNA-protein interactome, is highly context-dependent and changes relative to the cell's environment or health status. One commonly used strategy used to investigate direct RNA-protein interactions is the induction of a covalent link between a ribonucleotide and an amino acid using UV light, which only occurs at "zero distance" [143-145]. This covalent link allows the purification of an RPC under stringent conditions, disrupting any interaction that is not mediated through

a covalent bond. With the ability to covalently link RNA with the directly interacting protein, as well as technical advances in mass spectrometry the systematic discovery of new RBPs and non-canonical RNA binding domains has become possible [135, 146, 147].

1.7. The fate of (viral) mRNAs in the host cell

As already elaborated above, every virus requires to synthesize viral mRNA for the production of new infectious particles. One possibility for viral mRNAs to use the host translation machinery is by mimicking the structure of host mRNAs by capping and polyadenylation. To achieve this, viruses either use the host's RNA synthesis machinery, encode their own capping and polyadenylation enzymes or transfer the cap of a host mRNA to their own mRNA [123, 148, 149].

In host cells, RBPs bind to newly synthesized RNA in the nucleus. RBPs can add nucleotide modification to the RNA, orchestrate splicing of the primary transcript, or change the location of the RNA [150-155]. Thus, upon release into the cytosol, the RBPs have modified the transcript synthesized by the RNA polymerase before translocation into the cytosol. Modifications on the mRNA, as well as the composition of the bound RBPs, can alter the fate of an mRNA in the cytosol [150, 154, 155]. To provide an example, the YTH domain-containing family proteins (YTHDF1-3) are reader proteins of N^6-methyladenosine (m^6A) modifications on mRNA that recognize m^6A modified mRNA, bind those resulting in their destabilization [156, 157]. Several RBPs of mRNAs recognize and bind the 5´ cap structure, the poly(A) tail, or cis-regulatory elements in the 3´ UTR [135, 158-160]. Depending on the specific RBP complexes associated with the RNA, factors for translation initiation are recruited to drive translation and produce the encoded protein. The host protein eukaryotic translation initiation factor 4B (eIF4B) for example binds directly to mRNA and acts as a cofactor for eIF1 and eIF1A induced unwinding of 5´ terminal secondary structures [135, 161].

An erroneous mRNA or environmental stress on a cell can lead to translational inhibition induced by RBPs. Translational inhibition results in two different fates of an mRNA: 1. The RNA is silenced for later translation or degradation [162, 163] or 2. the RNA is decapped and degraded [164, 165]. These translationally arrested mRNAs are stored or degraded in membraneless organelles like processing bodies (P bodies) or stress granules.

Many of the RBPs recognize general features, like the 5´ cap, poly(A) tail, modifications, or secondary structures of mRNA. Those RBPs could also target viral mRNAs carrying the same features and aside of promoting translation induce translational silencing or degradation of the viral mRNA. Investigating which RBPs interact with the viral RNAs and how these interactions influence the fate of the viral RNA is pivotal to understand how viruses maintain the translation of viral proteins and how they protect their RNA from degradation.

1.8. Methods of investigating RNA-protein interactions

To investigate RNA-protein interactions many different strategies of cross-linking and purification methods were established in recent years, each with their benefits and limitations (Table 1). Using an RNA-centric approach followed by mass spectrometry reveals RBPs that interact with an RNA of interest (Table 2). A protein-centric approach typically takes advantage of next-generation cDNA sequencing to map the RNAs bound by the RBP and its binding site or binding preference on the RNA (Table 3). Combining both approaches provides valuable insight into the RNA-bound proteome and a first indication of involved cellular pathways. All these methods are mainly divided by their cross-linking and purification strategy.

1.8.1. Cross-linking of RNA-protein complexes

To investigate the direct interaction between an RBP and an RNA, fixation of the complex helps to reduce unspecific interactions through the application of stringent purification methods. Formaldehyde is a chemical compound that cross-links each type of nucleic acid and protein to other nucleic acids and amino acids. The covalent link introduced by formaldehyde requires a nucleophilic group that is readily available on biological molecules like proteins, DNA, or RNA. Depending on the concentration and the time those biological molecules are exposed to formaldehyde, the cross-linking reaction connects the entirety of all biological molecules in proximity. Depending on the question, it might be more relevant to only capture direct RBPs. For this, UV light at a length of 254 nm can be used as a cross-linking strategy to induce a covalent bond between a ribonucleic acid and an amino acid at “zero distance” [144, 145, 166].

Table 1 Benefits and limitations of most commonly used strategies to investigate RNA-protein interactions

	Strategy	Benefit	Limitation
cross-linking	254 nm UV irradiation	"zero distance" RBPs only	low efficiency
		applicable to tissue monolayer	
	365 nm UV irradiation	"zero distance" RBPs only	requires 4SU pre-treatment
		specific for 4SU incorporated RNA	
	formaldehyde	high cross-linking efficiency	unspecific cross-linking
		high tissue permeability	indirect interactors are captured
capture method	oligo(dT) capture	high probe/target affinity allows for stringent washing	detects only poly(A) RNA
	antisense capture	allows specific RNA targeting	non-target hybrids due to high-ionic strength buffers
	biotin pulldown	high affinity between streptavidin and biotin increases purification efficiency	high background based on naturally occurring substrates
	IP	target specific	requires availability of specific antibodies that are applicable to IP
	generic phenol-based enrichment	unbiased cross-linked RBP enrichment	high complexity dataset
		identification of unconventional RBPs	

The caveat of UV cross-linking is the low efficiency and the preference of pyrimidines, especially uracil, to engage in cross-linking [143]. Aside from those, multiple other factors like RNA structure, the type of engaging amino acid, as well as the contacting surface area of the protein-RNA interaction can influence the cross-linking efficiency (see 3.1). Unfortunately, the understanding of the UV cross-linking efficiencies of each nucleic acid

with the different amino acids is underreported and makes general assumptions difficult. To avoid underestimating the RBPs that interact with an RNA, much more material is used to compensate for the low efficiency of UV-crosslinking methods.

For the investigation of RBPs in the context of viral infection the usage of the uracil analogon 4-thiouridine (4SU), has been used to study newly synthesized or pre-labeled viral RNA [129, 160, 167]. Exposing cells to 4SU allows incorporation of the analogon instead of uracil into newly synthesized RNA. When RNA viruses are grown in such 4SU-exposed cells, they incorporate the analogon into their gRNA upon replication. Infecting cells with those 4SU-labeled viral particles can be used to capture RBPs relevant for viral entry or early synthesis [167, 168]. By irradiating the infected cells early in infection with 365 nm UV light, 4SU crosslinks to amino acids to investigate the interacting RBPs [169-171]. In addition, the incorporation of 4SU in newly synthesized RNA can be used to investigate the RNA synthesis of host and viral RNAs in the course of infection [172-174].

1.9. Purification of RNA-protein complexes

The target to be investigated already guides the decision on how to purify the target of interest. While protein-centric (Table 3) approaches mostly rely on immunoprecipitation (IP), purification of RNA requires enrichment through antisense oligonucleotides, oligo(dT) probes, or streptavidin pulldown of RNA-incorporated biotinylated 4SU (Table 2). Recent publications have established a generic purification strategy for the enrichment of cross-linked RNA-protein complexes [147, 175, 176]. Tiled antisense oligos can be used to enrich a specific target RNA. This allows to investigate which proteins bind directly to the specific RNA of interest [31, 177]. Oligo(dT) probes on the other hand enrich all poly(A) RNA targets, providing a general understanding of poly(A) binding proteins [135].

Supplementing the RNA-centric methods mentioned above, cross-linking immunoprecipitation (CLIP) methods have been established to investigate the RNA bound by an RBP of interest. This protein-centric approach combines IP on cross-linked RNA-protein complexes with RNA-sequencing. Different methods use different cross-linking strategies (Table 3). Cells supplemented with 4SU as used for PAR-CLIP and proximity CLIP reduces the applicability of those methods to cultivated cells but allows for the detection of the cross-linking site on a single nucleotide resolution. The

incorporated 4SU in the course of library preparation introduces a thymine (T) to cytosine (C) mutation that allows determination of the cross-linking site on a single nucleotide resolution [178]. The iCLIP and eCLIP methods maintain the information of the cross-linking site on the nucleotide level while avoiding the necessity to use 4SU [179, 180]. A big benefit especially of methods like eCLIP is the applicability of this assay even with *ex vivo* samples, since the method only requires UV cross-linking [180]. A general caveat of IP-based experiments is the availability of specific antibodies for endogenous antigens. Using cell culture models, this issue is solvable by the exogenous introduction of a tagged version of the protein of interest. The most recently developed antibody barcode CLIP (ABC-CLIP) uses oligonucleotide barcodes on antibodies, which allows the investigation of multiple RPBs from one sample [181]. The IP-captured RNA-protein complex in this method is ligated to the oligonucleotide barcode on the antibody and released by proteinase digest for RNA-sequencing.

Depending on the binding preference of an RBP, insights about the possible mode of functions can be concluded from a CLIP experiment. A project in which I employed eCLIP in hCMV infected cells was able to create a better understanding of the zinc finger CCCH-type antiviral protein 1 (ZC3HAV1 or ZAP), expanding the knowledge about its target specificity [34]. How RBPs bind the RNA depends on sequence motifs, secondary structures, or the type of RNA [31, 182-184]. Motif prediction tools like MEME SUITE can be used to find motifs that are enriched in binding sites of an RBP [185]. This has been performed in our work investigating the binding patterns of LARP1 in SARS-CoV-2 infected cells [31]. Using motif prediction tools on the eCLIP data revealed that LARP1 binds the SARS-CoV-2 RNA in regions rich in pyrimidine, similar to its preference as 5´ terminal oligopyrimidine (5´ TOP) binder on host mRNAs [31, 186, 187]. By combining eCLIP with the structural probing method, selective 2′-hydroxyl acylation analyzed by primer extension (SHAPE) provides additional information about the secondary structures of the RBP bound RNA [183]. With RBPs being able to bind certain types of RNA, fine-tuning the CLIP-seq library preparation or the analysis pipeline can provide answers to specific questions. As an example, the Argonaute (AGO) HITS-CLIP can be used to investigate the micro RNA (miRNA) and target mRNA pair of the RNA-induced silencing complex (RISC) by purification of AGO and subsequent ligation of the miRNA to its target

mRNA [184]. Using CLIP methods in the investigation of RNA viruses increases the complexity of the generated dataset with the introduction of neg sense RNA that can be captured by paired-end sequencing. As mentioned above, SARS-CoV-2 produces neg sense RNA to serve as a template for the synthesis of the viral mRNAs and gRNA (see. Chapter 1.5.3). Thus researching the binding profile of RBPs on neg sense RNA is especially interesting in the context of SARS-CoV-2 infection.

1.10. Aim of thesis

The recent outbreak of the COVID-19 pandemic had spurred the scientific community to provide further knowledge about coronaviruses to help understand the virus-host interplay and find new antiviral agents. Knowing that for successful infection the interaction of viral RNA with viral and host proteins is pivotal for replication, our lab has generated the first SARS-CoV-2 RNA-protein interactome [31]. This provides a list of RBPs that are direct interaction partners of the viral RNA, which potentially could act in a pro- or antiviral manner. Using the interactome datasets generated in our lab, this project uses eCLIP as a complementary method to investigate the host RBPs that bind directly to SARS-CoV-2 RNA and provide a first impression of the potential biological relevance of those interactions. This work aims to deepen our understanding of the RNA-protein interactome by validating selected RBPs and defining their RNA binding sites. Using interesting features revealed from the RBP footprint the molecular mechanism of candidate RBPs and its relevance in SARS-CoV-2 replication will be addressed.

Table 2 Examples of RNA-centric methods to investigate RNA-protein interactions

Method name	Abb.	Cross-linking strategy	4SU	Capture method	Ref.
RNA **a**ntisense **p**urification coupled with **m**ass **s**pectrometry	RAP-MS	254 nm UV	-	biotinylated tiled antisense probes and streptavidin pulldown	[177, 188]
thio**u**racil cross(**X**)-linking **m**ass **s**pectrometry	TUX-MS	365 nm UV	during infection	oligo (dT) bead capture or antisense probes	[189]
viral **c**ross-**l**inking **a**nd **s**olid-phase **p**urification	VIR-CLASP	365 nm UV	pre-labeled virus	SPRI beads	[168, 190]
viral **R**NA **i**nteractome **c**apture	vRIC	365 nm UV	during infection	oligo(dT) bead capture	[167]
compre**h**ensive **i**dentification of **R**NA-binding **p**roteins by **m**ass **s**pectrometry	ChIRP-MS	formaldehyde	-	biotinylated tiled antisense probes and streptavidin pulldown	[191, 192]
cross-**l**ink **a**ssisted **m**RNP **p**urification	CLAMP	formaldehyde	during infection	HPDP-biotin conjugate and streptavidin pulldown	[193][193]
hybridization **p**urification of **R**NA-protein complexes followed by **m**ass **s**pectrometry	HyPR-MS	formaldehyde	-	single biotinylated antisense probe, streptavidin pulldown and toe-hold mediated release	[194]

Table 3 Examples of protein-centric methods to investigate RNA-protein interactions

Method name	Abb.	Cross-linking strategy	4SU	Capture method	Ref.
enhanced **c**ross-**l**inking **IP**	eCLIP	254 nm UV	no	IP of target RBP	[180, 195]
formaldehyde **c**ross-**l**inking **IP**	fCLIP	formaldehyde	no	IP of target RBP	[196]
high-**t**hroughput **s**equencing of RNA isolated by **c**ross-**l**inking **IP**	HITS-CLIP	254 nm UV	no	IP of target RBP	[197]
individual-nucleotide resolution **c**ross-**l**inking and **IP**	iCLIP	254 nm UV	no	IP of target RBP	[179]
photo**a**ctivatable **r**ibonucleoside-enhanced **c**ross-**l**inking and **IP**	PAR-CLIP	365 nm UV	yes	IP of target RBP	[178]
proximity cross-**l**inking **IP**	proximity-CLIP	365 nm UV	yes	streptavidin affinity pulldown	[198]
antibody **b**arcode **c**ross**l**inking and **IP**	ABC	254 nm UV	no	IP of multiple RBPs in parallel	[181, 199]

2. Results

2.1. Validation of candidate SARS-CoV-2 RNA binding proteins

As a response to the Covid-19 pandemic in 2019, our lab has published the first SARS-CoV-2 RNA-protein interactome that delineates direct viral RNA binding host proteins [31]. To purify viral RNA with the directly bound cellular RBPs, biotinylated DNA oligonucleotides were designed to bind antisense to the pos sense SARS-CoV-2 RNA. SARS-CoV-2 infections for RAP-MS experiments were performed for 24 h in Huh-7 cells with a multiplicity of infection (MOI) of 10 plaque forming units per cell (PFU/cell) [31, 177, 200]. The infected cells were UV cross-linked, lysed and the coronavirus RNA enriched using biotinylated antisense oligonucleotides under stringent, protein denaturing conditions. Proteins were eluted with a nuclease cocktail and identified using tandem mass tag (TMT)-based quantitative liquid chromatography mass spectrometry (LC-MS/MS). TMT labeling prior MS allows for multiplexing and quantification of the identified peptide fragments. Out of 276 identified proteins that were enriched in the SARS-CoV-2 RAP-MS experiments, 57 displayed significantly enrichment (adjusted *P* value < 0.05, two-tailed *t*-test) and thus are defined as core interactors. Additionally, an expanded interactome was defined using a false discovery rate (FDR) of < 20 %, which adds another 47 potential SARS-CoV-2 RBPs. The expanded core interactome contains 114 proteins, of which many are known host RBPs based on previous system-wide studies (68 out of 106) [135]. The interacting host RBPs were ranked based on their significance, which was used as a basis for selecting candidate proteins for further validation (Figure 4A, from [31]). The method of choice to complement this dataset and validate the RBPs with poor statistical power is eCLIP.

The eCLIP method provides valuable information about which RNA is bound by the RBP, its footprint on the RNA, and the nucleotide that was covalently cross-linked to the RBP (Figure 4B). In brief, cells were infected with SARS-CoV-2 using an MOI of 5 PFU/cell for 24 h and irradiated with 254 nm UV light to induce a covalent link between the interacting nucleic acid and amino acid side chain. Those cells were lysed and RNase digested, to degrade RNA that is not protected by proteins. While 5 % of the lysate was stored as an

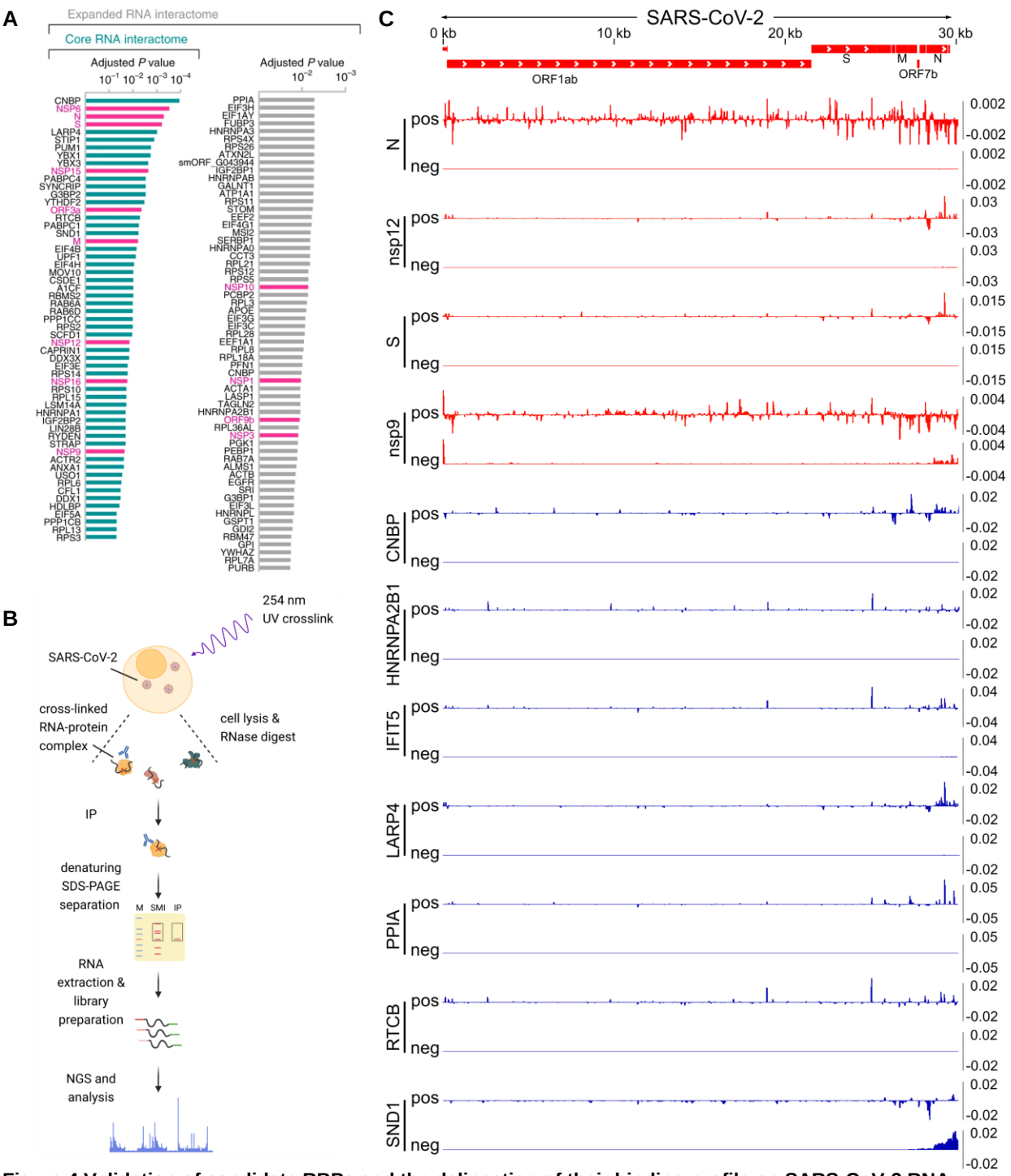

Figure 4 Validation of candidate RBPs and the delineation of their binding profile on SARS-CoV-2 RNA

A The previously published RNA interactome reveals direct SARS-CoV-2 RNA binding host proteins (from, [31]). **B** Schematic illustration of the eCLIP protocol (M = marker; SMI = size matched input control; IP = immunoprecipitated sample. Illustration created with BioRender.com. **C** Sequencing reads obtained from viral RBPs candidates are illustrated in red, and from human RBPs in blue. Obtained read probability of IP samples is subtracted from SMI read probability and aligned to the whole SARS-CoV-2 genome for positive (pos) or negative (neg) sense viral RNA. Infection for the eCLIP was performed by Dr. Nora Schmidt and data analysis by Dr. Alexander Gabel.

input control, the remaining lysate was used for an IP, to enrich for the protein of interest and it's covalently linked RNA footprint. To denature all non-covalent protein-protein and RNA-RNA interactions, a denaturing sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) was performed. This step in addition ensures to exclude all RNA fragments that are not covalently attached to a protein [201]. In parallel, the aforementioned input sample was run on the same SDS-gel. After a transfer on a nitrocellulose membrane the RNA-protein complex was excised from the expected size. Cutting out the same area in the input sample generates a size-matched input (SMI) control that provides information about all background RNA footprints running at the same height as the protein of interest. Using a Proteinase K digest the RNA footprint was released and subsequently transcribed into copy DNA (cDNA). The polymerase is not able to efficiently continue past the cross-linking site and therefore terminates the reverse transcription [202]. The resulting drop-off pattern can be used to determine the cross-linking site between the RBP and RNA at single nucleotide resolution. The cDNA was amplified with a polymerase chain reaction and after a library purification sequenced in paired-end mode using an Illumina NextSeq sequencer. Further bioinformatic processing of the obtained reads (see 4.5) is used to infer information about the site that was bound by the RBP of interest, the nucleotide engaged in the direct interaction, as well as the orientation of the originally bound RNA. The sequencing depth of the different libraries was normalized and the relative frequency of reads in an analysis window (= read probability) calculated for each sample. Afterwards, the SMI read probability was subtracted from the IP read probability. Positive values indicate a greater likelihood for a read to belong to the IP sample, while negative values indicate a greater likelihood for a read to belong to the SMI sample. This analysis method allows for a simplified visualization that shows enriched binding on the positive axis over unspecific reads in the SMI on the negative axis (Figure 4C).

2.1.1. Investigation of the binding profile of viral RBPs

The core SARS-CoV-2 RNA interactome shows enrichment of several characterized (N, nsp15, nsp12, nsp16, nsp9, nsp1) and so far undescribed (nsp6, S, ORF3a, M, nsp10, ORF9b) viral RBPs [203]. I selected several of the known RBPs (N, nsp12, and nsp9) as

well as the structural protein S to investigate their binding site on SARS-CoV-2 RNA by eCLIP (Figure 4C).

The structural protein N is reported to cover the whole gRNA of SARS-CoV-2 for packaging via the N-terminal RNA binding domain [66, 67, 204, 205]. The binding preference of N detected by eCLIP seems to be more evenly distributed but exclusive to the pos sense RNA. Additionally, the overall background signal obtained in the SMI is strong.

Nsp12 is known to be in direct contact with the SARS-CoV-2 RNA through the finger and thumb domains while it replicates and transcribes viral RNA [96, 97, 206, 207]. Unexpectedly, the nsp12 binding pattern did not show an even distribution on the pos sense RNA. Instead, strong binding is visible in the 5´ leader sequence, ORF7b, and in N. Additionally, no binding to neg sense RNA is apparent, despite the necessity of nsp12 to bind neg sense RNA in order to synthesize the viral mRNAs and gRNAs.

While N and M are structural proteins with known RNA-binding function, S has no reported RNA binding capability [66, 67, 131-133]. S binds the host cell receptor ACE2, which after furin cleavage leads to the fusion of viral and host membranes [40, 70, 72-74, 208]. Curiously, despite no reported RNA binding capability, the core SARS-CoV-2 RNA-protein interactome has revealed S as a top hit candidate protein. Such a result could be driven by abundance or specificity of this protein. Looking at the structure of S could contain a disordered region that might associate with RNA [72]. Employing eCLIP to validate the RNA binding capabilities of S shows binding to the pos sense RNA. The most prominent binding region locates in the genomic locus of N while further small binding sites map to the 5´ leader sequence, ORF1ab and S. Additionally, several of the binding sites of S overlap partially with those predicted for N.

The last investigated viral RBP is nsp9, which is known as an unspecific ssRNA and dsRNA binding protein [111, 112]. Beyond the well known RNA binding properties of nsp9 recent work suggest that NMPylated nsp9 might play a role in priming or capping of nascent RNA synthesis [95, 99-102]. While nsp9 does not bear a canonical nucleic acid binding domain or motif, there are suggestions that the N-terminus containing a small disordered stretch could contribute to RNA binding [209]. Investigation of RNA binding

profile of nsp9 reveals a strong and distinctive binding pattern in the leader sequence. In addition, smaller reads are detectable throughout the whole SARS-CoV-2 genome. Interestingly, nsp9 binding to the neg sense RNA could be detected with strong binding in the region antisense to the leader sequence, while a broader binding pattern was found spanning the region antisense to N and the 3´ UTR. This finding is particularly fascinating since the overall abundance of neg sense RNA is in estimation two orders of magnitude lower compared to pos sense RNA suggesting a strong specificity for this interaction [117]. The strong enrichment of neg sense RNA indicates a relevance of this protein in RNA synthesis supporting the literature that associates nsp9 with the RTC [95, 99-102].

2.1.2. Investigation of the binding profile of host RBPs

A multitude of general cellular RBPs as well as factors associated with innate immunity and translational regulation were enriched in our extended SARS-CoV-2 core interactome (Figure 4A) [31]. By comparing the candidates with a genome-wide CRISPR screen and other SARS-CoV-2 interactome studies, I selected candidate host RBPs for the investigation of their binding preference on the viral RNA [24, 130, 167, 210]. This list included: CCHC-Type Zinc Finger Nucleic Acid Binding Protein (CNBP), heterogeneous nuclear ribonucleoprotein A2/B1 (HNRNPA2B1), interferon-induced protein with tetratricopeptide repeats 5 (IFIT5), La-related protein 4 (LARP4), peptidylprolyl isomerase A (PPIA), RNA 2',3'-cyclic phosphate and 5'-OH ligase (RTCB) and Staphylococcal Nuclease and Tudor Domain Containing 1 (SND1).

The top candidate in our core RNA interactome as well as a strong hit in the CRISPR screen among RBPs is CNBP [24, 31]. This protein contains 7 zinc ZnF domains and an Arginine-Glycine-Glycine (RGG) motif that can facilitate RNA binding [211, 212]. CNBP has been described to bind and modulate G-quadruplex structures in RNA and DNA, therefore exerting a regulatory role at a transcriptional and translational level [211, 213, 214]. An additional relevance of CNBP as antiviral factor has been revealed for several viruses like IAV and SARS-CoV-2 [31, 214]. On SARS-CoV-2, CNBP binds in S, M, ORF3a, N, and the 3´ UTR on the pos sense viral RNA according to the obtained eCLIP data. No binding of CNBP to the neg sense RNA is detectable.

The extended core RNA interactome featured several members of the HNRNP family. This family of RBPs take part in multiple aspects of RNA biology including transcription, splicing, translation, and mRNA stability [179, 215, 216]. HNRNPA2B1 is involved in in shuttling specific RNAs between the nucleus and cytosol and was suggested as a nuclear innate immune sensor of viral DNA [216-218]. Two RRM in HNRPA2B provide the RNA binding properties, aside of those several IDRs could contribute to RNA binding [219]. Using eCLIP, I could confirm that HNRNPA2B1 binds within genes of the pos sense SARS-CoV-2 RNA of. This protein as well does not bind to neg sense RNA.

After secretion of interferons, IFIT5 is one of a multitude of genes that is induced to mount an antiviral response [220]. It is an RNA-binding protein that recognizes 5′ triphosphate (5´ppp), 5´monophosphate (5´p) or cap0 RNA and plays a role in antiviral immunity [221, 222]. A tetratricopeptide repeat (TPR) in the central cavity of this protein is the recognition domain for the 5´ ends of RNA [221-224]. Interestingly, no binding to the 5´ end of the SARS-CoV-2 RNA is detected. The strongest binding site observed for IFIT5 locates within the gene S on the pos sense viral RNA, while no binding to the neg sense RNA is detectable.

The family of La-related proteins is associated with translational regulation. LARP4 binds via a La-motif (LaM) followed by a RRM the poly(A) tail of mRNA [225]. In addition, it carries several disordered regions that could contribute additional RNA binding sites. Association of mRNA binding LARP4 with the 40S ribosomal subunit positively affects translation efficiency and mRNA stability [225]. By protecting the mRNA from deadenylation, LARP4 contributes to the stability of the bound mRNA [225, 226]. In SARS-CoV-2 infected cells, LARP4 shows an extended binding region ranging from the 3´ UTR to the gene N on the SARS-CoV-2 pos sense RNA. Additionally, a weak binding site in the 5´ leader sequence is detectable. No binding of LARP4 to the neg sense RNA was found.

Cyclophilins, like PPIA, play key roles at the host-virus interface and regulate immune signaling [227, 228]. Several reports show that it interacts either as a pro- or antiviral host factor depending on the investigated virus [228-232]. Replication of a broad range of coronaviruses can be inhibited by treatment of the infected cells with the PPIA inhibitor

cyclosporin A (CysA)[31, 233-235]. Despite most of the reports showing interactions of PPIA with viral proteins, it has appeared as a non-canonical RBP in studies investigating the protein interactome of poly(A) RNAs [135]. PPIA contains disordered residues that interact with each other via lattice contacts and thus could provide a platform for RNA binding [236]. On SARS-CoV-2, the binding profile of PPIA is exclusive to the pos sense RNA and is particularly strong in the genes encoding the structural proteins N, M, and S.

Another highly enriched factor of our core RNA interactome is RTCB, which is suggested to have antiviral properties on SARS-CoV-2 based on a CRISPR-screen [24, 237]. RTCB is reported to be an essential subunit of the RNA-splicing complex [238]. At the N-terminus, RTCB contains a RG-rich amino acid sequence close to poorly resolved secondary structure predictions [239]. Both are indications for a potential RNA binding site since poor secondary structure resolutions in crystallography can be a consequence of disordered regions while RG-rich sequences are known to be capable of RNA binding [139, 240]. Based on the eCLIP data, RTCB binds the SARS-CoV-2 RNA within genomic loci exclusively on pos sense RNA.

SND1 is a strongly enriched RBP in our core interactome containing an interesting domain architecture and multifaceted biological functions. SND1 has five staphylococcal nuclease (SN) domains as well as a Tudor domain (Figure 5A) [241]. Despite the suggestive naming, the activity of the SN domains in humans is debated. Mutations at the catalytic site in the human SN domains are considered to render the protein inactive [241]. Despite this, several studies claim a nuclease activity of SND1 relevant for the degradation of miRNAs and hyper-edited dsRNA [242-244]. The SN domains have nucleic acid binding capabilities, which are important for the interaction with RNA or DNA [241, 245-247]. In addition to the interaction with nucleic acids, the SND1 Tudor domain recognizes dimethylarginine (DMA) modifications on proteins and is relevant for protein-protein interaction [248-252]. SND1 is relevant in a multitude of RNA regulatory pathways like splicing, transcription, RNA regulation, or stress granule formation in the context of protein-protein interactions [253-259]. Therefore, it is likely that SND1 facilitates the assembly of protein complexes in proximity to RNA or DNA [260]. Despite the rather promiscuous role of SND1 in human cells, the putative nuclease domains intrigue to assume a potential antiviral role in SARS-CoV-2 infected cells [242-244]. The eCLIP

results of SARS-CoV-2 infected cells show a strong binding preference of SND1 to the viral 5´ neg sense RNA, a so far unique mode of binding that has not been described previously for human RBPs. This binding site covers a wide range starting from the sequence antisense to the 3´ UTR to the antisense gene locus of N. Binding to the neg sense RNA has so far not been observed for host RBPs in the eCLIP datasets. Combined with the low abundance of neg sense RNA this indicates a strong specificity of the RBP-RNA interaction. Interestingly, the only other RBP I could show to bind neg sense RNA, nsp9, binds the same region.

Comparative investigation of the binding profile of human and viral RBPs shows that each protein has distinct binding patterns that can be reflected by known modes of action. Therefore, using the binding profile of RBPs on an RNA can provide first insights into the potential mechanism of an RBP without a known mechanism. With the neg sense RNA serving as a template for RNA replication and transcription, the specific neg sense RNA binding patterns of SND1 and nsp9 could reflect participation of the RBPs in these events. Finding a shared sequence region bound by both proteins, could indicate that the RBPs compete or support each other in binding the RNA. With SND1 being a unique neg sense RNA binding host protein, it is intriguing to investigate the potential function of SND1 in SARS-CoV-2 replication and transcription.

2.2. SND1 is a proviral RBP and associates with viral proteins at DMVs

As described above, SND1 engages in protein-protein interactions and forms protein complexes with known relevance in RNA biology. SND1 has 5 SN domains with nuclease binding ability and a Tudor domain that recognizes and binds DMA modifications on proteins [248-252, 261] (Figure 5A). So far, the implications of SND1 in SARS-CoV-2 infection remain poorly understood. To characterize the relevance of SND1 in SARS-CoV-2 infection, our lab has generated a SND1 knockout system in $A549^{ACE2}$ cells (Figure 5B). Infecting cells for 6 or 12 hours (h) with an MOI of 3 PFU/cell, we could show a significant reduction in viral RNA levels upon deletion of SND1 indicating a proviral effect of SND1 in SARS-CoV-2 infected cells (Figure 5C). Reintroduction of SND1 into the KO cells using a

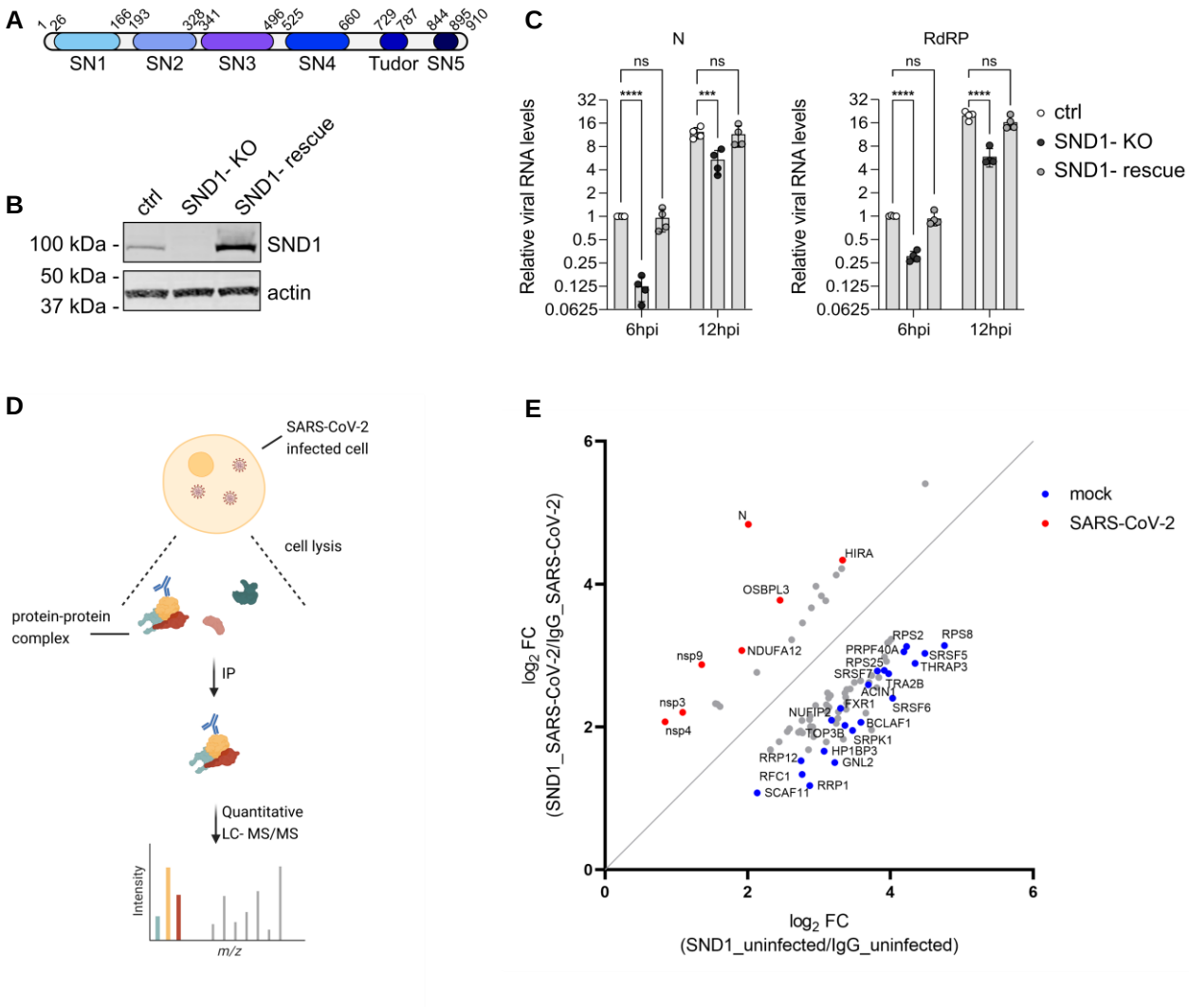

Figure 5 The protein interaction network of the proviral host factor SND1 changes upon SARS-CoV-2 infection

A Simplified visualization of the SND1 domain architecture consisting of 5 staphylococcal nuclease (SN) and a Tudor domain. **B** Western Blot validation of the A549^{ACE2} SND1 knockout (SND1-KO) and rescue (SND1-rescue) model used in this study, kindly provided by Dr. Nora Schmidt. **C** Characterization of the relevance of SND1 in SARS-CoV-2 infection by Dr. Nora Schmidt. Relative viral RNA levels were determined by infecting A549^{ACE2} control (ctrl), SND1-KO, and SND1-rescue cells with SARS-CoV-2 at a MOI of f 3 PFU/cell. After 6 or 12 hpi, RNA was extracted and a quantitative reverse transcription polymerase chain reaction (RT qPCR) with primers corresponding to N or RdRP performed. RNA levels were normalized to 18S rRNA and are shown relative to ctrl cells at 6 hpi. Mean values of *n* = 4 independent experiments with 3 technical replicates and their standard deviation are blotted. *P* values determined by two-way ANOVA with Dunnett's Test. **D** Schematic outline of the co-IP MS experiment. Illustration created with BioRender.com. **E** Differential changes of SND1 protein interaction partners in uninfected (x-axis) or SARS-CoV-2 infected A549^{ACE2}cells (y-axis) of *n* = 2 independent experiments. Proteins with a $\log_2$ FC >1.5, a *P* value of < 0.01 and an FDR < 0.2 are depicted as gray dots. In red, protein interaction partners enriched upon SARS-CoV-2 with a FDR < 0.05 interactors are depicted. Blue shows enriched interaction partners in mock cells with a FDR < 0.05. SARS-CoV-2 infection performed by Dr. Nora Schmidt. MS data analysis was performed by Dr. Alex Gabel.

lentiviral vector (SND1-rescue) restores the observed phenotype compared to SND1-KO cells that were treated with an empty vector control (SND1-KO) (Figure 5B-C). This phenotype is most prominent at early times of infection (6 hpi) but still detectable at later stages (12 hpi). Together with the preference of SND1 to bind negative sense RNA -the transcriptional template for mRNA and gRNA synthesis- it is intriguing to speculate that SND1 might support viral RNA synthesis. With SND1 as a member of a multitude of RNA/DNA-protein complexes in human cells, I aimed to identify the protein complex associated with SND1 [248-258, 262].

Using the established cell culture model, I performed a co-IP experiment followed by TMT-based quantitative mass spectrometry (co-IP MS). For this, $A549^{ACE2}$ cells were infected for 24 h with SARS-CoV-2 using an MOI of 5 PFU/cell or left uninfected (mock). The cells were lysed and SND1 purified using a specific antibody raised against the endogenous protein (Figure 5D). Proteins that unspecifically bind to SND1 are washed away and the remaining co-purified proteins were identified and quantified by mass spectrometry. As a control, unspecific antibody Immunoglobulin G (IgG) was used in an IP reaction to identify unspecific background interactions that are enriched due to the experimental conditions. Relative quantification of proteins enriched in SND1 IP over the IgG control was used to identify specific interactors of SND1. Analysis of the MS dataset revealed SND1 and known cellular binding partners, such as Metadherin (MTDH), among the highly enriched proteins, indicating a successful experimental procedure [242, 263]. Normalizing the SND1-IP to an IgG control revealed 1,732 candidate proteins with a $\log_2$ fold change (FC) ≥ 1.5, a *P* value < 0.01, and an FDR < 0.2. For a simplified visualization, only proteins with a significant change in binding preference between mock and SARS-CoV-2 infected cells are depicted as dots in the graph (Figure 5E). Interactors of SND1 that are specific in SARS-CoV-2 infected cells with an FDR < 0.05 are colored in red, while SND1-interactors diminished in uninfected cells and an FDR < 0.05 are blue. Upon infection with SARS-CoV-2, the protein-protein interactome of SND1 changed massively. Interaction partners of SND1 in mock cells are annotated as RNA splicing and processing factors according to a gene ontology (GO) enrichment analysis (Supplementary Table 1). Upon infection new viral and host protein interactors associated with cellular membranes appear (Oxysterol-binding protein-related protein 3 (OSBPL3), NADH dehydrogenase 1

alpha subcomplex subunit 12 (NDUFA12), nsp3, nsp4) [264, 265]. The viral proteins nsp3 and nsp4 form DMVs out of the ER membrane creating the site of viral replication in which the transcription machinery is located [90-92]. Intriguingly, except for nsp9, no member of the replicase complex (nsp7, -8, -10, -12, -13, -14, -16) was found to associate with SND1.

Investigating the protein-protein interactome of SND1 reveals that infection with SARS-CoV-2 dramatically changes the protein interaction partners of SND1. A shift to ER-associated factors in infection from general RNA processing and regulation could indicate that the biological role of SND1 is remodeled upon infection.

2.3. SND1 interacts with nsp3 and nsp9 at perinuclear compartments

To validate the above findings and get a better understanding of the localization of the protein-protein interactions and SND1 complexes in intact cells, I have performed confocal immunofluorescence (IF) imaging to capture the co-localization of SND1 with nsp9 or nsp3 (Figure 6). With both viral factors being associated to DMVs, microscopic imaging allows to investigate the spatial distribution of the proteins captured in the previous co-IP MS experiment. $A549^{ACE2}$ cells were infected with an MOI of 10 PFU/cell for 8 h and co-stained with antibodies targeting SND1 as well as nsp3 or nsp9. Staining of SND1 reveals an even cytosolic distribution with few high intensity foci, as well as little signal located in the nucleus (Figure 6A, B). This is supported by the literature and the described cellular functions in transcription, splicing, and RNA stability [255-257, 260]. Immunostaining of nsp3 in SARS-CoV-2 infected cells depicts a perinuclear signal that corresponds to the localization of DMVs (Figure 6A) [92]. This signal is strongest directly around the nucleus and reaches outwards with gradual reduction of the signal intensity. To understand the extent of association between those factors, I performed a Manders´ coefficient analysis of individual SARS-CoV-2 infected cells using the FIJI plugin JACoP [266, 267]. Focusing on the co-occurrence of SND1 and nsp3 revealed that 76.6 % of pixels with nsp3 signal also detect a SND1 signal (Manders´ coefficient = 0.766 ± 0.166) (Figure 6B). Nsp9, similar to nsp3, shows perinuclear staining that corresponds to the localization of DMVs (Figure 6C). Quantifying the signal co-occurrence of SND1 on nsp9 results in a Manders´ coefficient of 73.8 % (Manders´ coefficient = 0.7383 ± 0.2549) (Figure 6D). This data

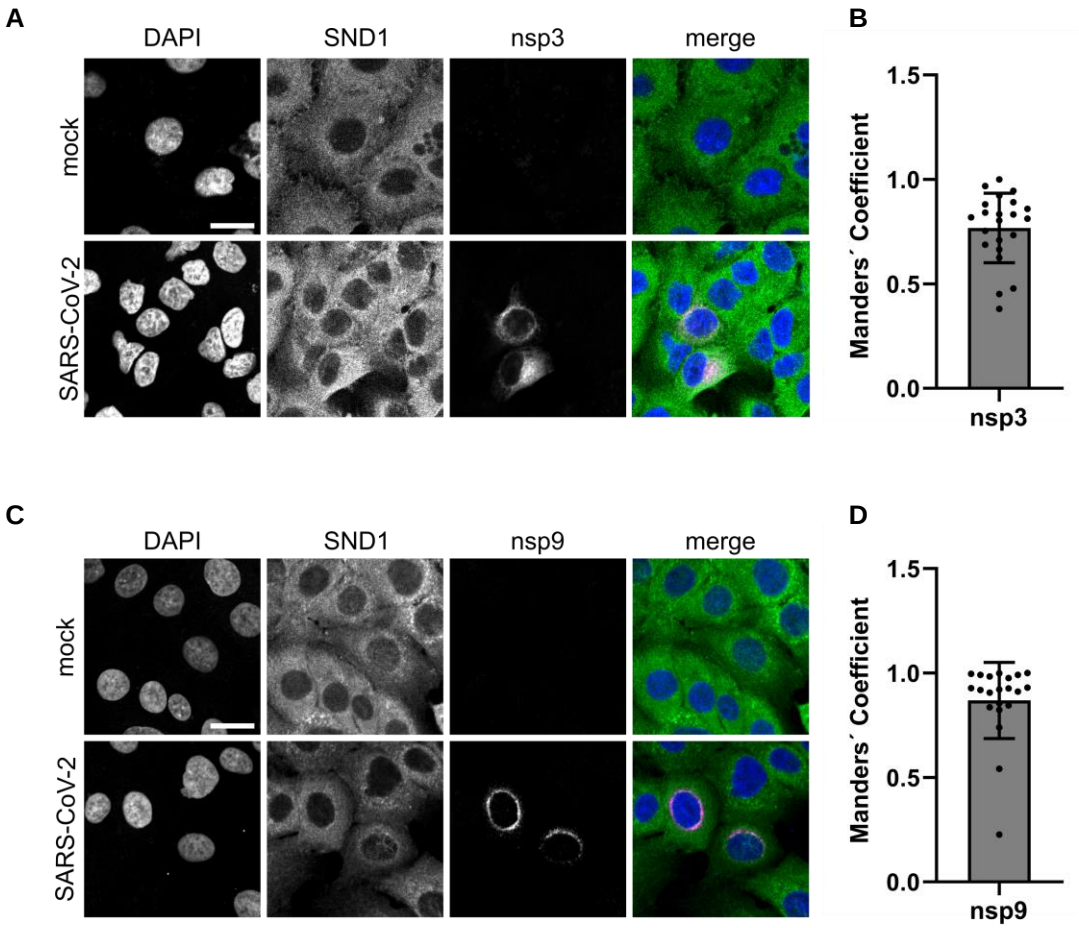

Figure 6 SND1 co-localizes with nsp3 and nsp9

A + **C** Representative co-localization (white) images of SND1 (green) and **A** nsp3 (magenta) or **C** nsp9 (magenta) with nuclear staining using DAPI (blue) in SARS-CoV-2 infected (PFU = 10 MOI; 8 hpi) or uninfected $A549^{ACE2}$ cells. Scale bar represents 20 µm. **B** + **D** Quantification of SND1 signal co-occurrence with **B** nsp3 or **D** nsp9 using the Manders´ Coefficient. **B** Representative graph of one out of two replicates with *n* = 22 quantified cells. **D** Representative graph of one out of 3 replicates with *n* = 35 quantified cells. SARS-CoV-2 infection performed by Dr. Nora Schmidt.

supports the above finding that SND1 is localized in close proximity to nsp3 and nsp9 at the site of DMVs in SARS-CoV-2 infected cells.

Since the above experiments only imply co-occurrence but not close proximity, I employed a proximity ligation assay (PLA) [268]. PLA detects protein-protein interaction at a maximum distance of 40 nm and provides further indication of close association [269]. PLA combines target specificity of IF imaging with signal enhancement through rolling circle amplification (RCA). For this, cells are fixed and permeabilized to allow access of antibodies to the target of interest. Upon binding of primary antibodies to their targets,

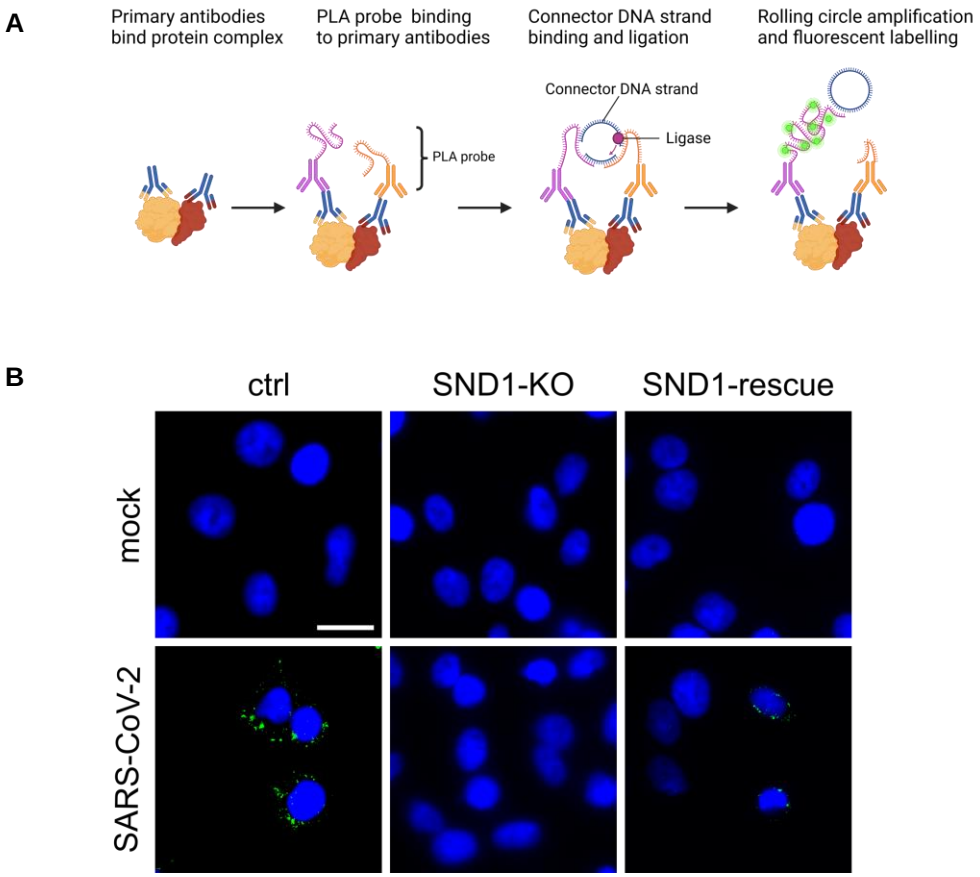

Figure 7 SND1 and nsp9 interact at foci around the nucleus

A Schematic illustration of the proximity ligation assay (PLA) to investigate the proximity of SND1 with nsp9 target proteins. Illustration performed with BioRender.com. **B** Representative images of PLA signal (green) around the nucleus stained with DAPI (blue) in SARS-CoV-2 infected (MOI =10 PFU; 8 hpi) or mock infected A549^{ACE2} ctrl, SND1-KO or rescue A549^{ACE2} cells. The scale bar represents 20 µm. SARS-CoV-2 infection performed by Dr. Nora Schmidt

secondary antibodies with covalently attached DNA oligos are added (Figure 7A). The PLA probes contain the necessary sequences for a connector DNA strand to bind, which can only occur if both PLA probes are in close proximity. Once the connector DNA strand is attached and connected by the ligase, RCA creates a DNA concatamer that binds fluorescent dyes for subsequent imaging. To investigate the association of SND1 with nsp9, A549^{ACE2} cells were infected for 8 h in the presence (ctrl or SND1-rescue) or absence of SND1 (SND1-KO) and the acquired PLA signal compared to uninfected cells. Only infected cells that express SND1 form a perinuclear spot-like PLA signal (Figure 7B lower row left and right picture), while no signal was detected in the negative control

conditions (Figure 7B upper row, middle picture lower row). This reveals the proximity of SND1 to nsp9 at perinuclear foci where DMV structures are expected.

2.4. SND1 modulates covalent binding of nsp9 at RNA ends

With the above experiments, I have established that SND1 and nsp9 are interacting in regions close to DMVs where viral RNA is synthesized. Both proteins directly bind neg sense SARS-CoV-2 RNA (Figure 4C) but the molecular consequence of this interaction remains unclear. Nsp9 is described as an unspecific ssRNA binding protein and target of nsp12 NiRAN domain mediated NMPylation [95, 99-102, 111, 112]. The *in vitro* experiments show that an NMP is covalently attached to nsp9 by the NiRAN domain. While hypothetically each nucleotide can be used for the NMPylation reaction, different preferences were reported towards either guanine (G) or uracil (U) [95, 99-102]. On nsp9, the first two asparagine (ASN) at the N-terminus of nsp9, are crucial for the NMPylation reaction [100, 102].

As eCLIP is the state-of-the-art method to map RNA fragments covalently attached to RBPs through UV induced cross-linking, we tried to use this method to investigate covalent RNA-protein linkages that occur *in cellulo*. Therefore, I have performed a modified eCLIP method, which omits the UV-crosslinking step. As cross-linking is an essential step in the eCLIP protocol, it is expected that RBPs with no natural covalent attachment of RNA (RNAylation) will not yield material to generate a library suitable for sequencing [270, 271]. The covalent RNA IP (cRIP) experiment aims to reveal the RNA fragments covalently attached to a protein of interest under biological conditions, if such linkages exist (Figure 8A). For this purpose, cells were lysed after 8 h of SARS-CoV-2 infection using an MOI of 3 PFU/cell and immediately processed according to the eCLIP method. After nsp9-IP, the crucial step of denaturing SDS-PAGE was performed to remove any non-covalently bound RNA [201]. The RNA library obtained provides information about RNA fragments that are covalently attached to nsp9 in infected human cells. The sequencing reads were separated based on their origin from the pos or neg sense viral RNA aligned to the SARS-CoV-2 genome as the read probability like described for the eCLIP data visualization (see section 2.1). Based on the nature of the experiment, significant binding sites in this experiment are called as peaks and represent the RNA sequences that are protected by

nsp9. Significant enrichment of peaks (blue tracks below the binding profile in Figure 8B-D) was determined by MACS2 with a one-sided Fischer's exact test that was corrected for multiple testing with the Benjamini-Yekutieli method [272]. As NiRAN covalently attaches the NMP with the 5' carbon of the sugar-phosphate backbone to the first ASN on nsp9 through a phosphoramidate linkage, one can assume that the predicted covalent linkage is located at the 5´ end of the detected RNA sequence [100]. These covalent linkages, are determined based on the drop-off rates using a personalized Java script and blotted below the alignments in black (Figure 8B-D) [273, 274].

In the pos sense RNA, the most striking peak with a 5´ covalent linkage locates in the leader sequence right before the TRS-L (position 7-54) (Figure 8B left, C). This region is in close proximity to the presumed initiation site of pos sense RNA synthesis [119]. Additional peaks with a 5´ covalent linkage map to the neg sense RNA in the sequence antisense to 3´ UTR (anti 3´ UTR) that corresponds to the position 29822-29879 of the SARS-CoV-2 gRNA (Figure 8B right, D). At the 3´ end, several adenines (A) are predicted as the nucleotide covalently attached to nsp9, which in reverse orientation corresponds to uracil (U). Again, this site is a presumed initiation site but in this case for neg sense RNA synthesis [118, 119]. Curiously, a significant peak with a 5´ covalent linkage is also present on the neg sense RNA antisense of the leader sequence (Figure 8B left, C). Detecting significant peaks in non-UV treated samples together with covalently linked nucleotides is the first experimental evidence showing covalently attached RNA on nsp9 in authentic viral replication.

Knowing that SND1 interacts with nsp9, we wanted to understand if the RNAs covalently bound to nsp9 are affected by the depletion of SND1. For this, cRIP experiments were carried out in SARS-CoV-2 infected SND1-KO cells and changes in nsp9 binding on pos sense or neg sense RNA, which are reflected by the peaks, compared to ctrl cells. After separate peak calling, overlapping nsp9 binding peaks of the SND1-KO and control experiment are intersected. A two-sided Fisher's exact test was performed to determine differentially binding affinities between the two conditions. The differential binding is calculated in SND1-KO cells over control cells and each intersected peak plotted as separate point (Figure 8E+F). Increased binding of nsp9 in the absence of SND1 is indicated by a $\log_2$ FC > 1, while reduced binding of nsp9 in the absence of SND1 results

in a $\log_2$ FC < -1. The absence of SND1 results in increased reads at the 5´ end of pos sense RNA (Figure 8E) and fewer reads mapping to the anti 3´ UTR (Figure 8F). This shows that the presence of SND1 influences the distribution of nsp9 on the viral RNA. All in all the results provide compelling data showing the relevance of SND1 on nsp9 RNAylation.

2.5. Summary

With this project, I validated host and viral SARS-CoV-2 RNA RBPs and delineated their binding preference on the viral RNA. This resulted in the largest collection of binding profiles of host and viral RBPs on SARS-CoV-2 RNA to date. In general, several RBPs display distinct binding preferences and mostly prefer binding to the pos sense viral RNA. This work is the first to report a host RBP binding to neg sense coronaviral RNA. SND1 uniquely binds to the 5´ end of neg sense RNA, the template for pos sense RNA synthesis. The binding region lies antisense to the 3´ UTR of the SARS-CoV-2 gRNA and sgRNAs and indicates a potential mechanism in RNA synthesis. The same region is also bound by the viral ssRNA binding protein nsp9. Using a knockout and rescue system, our lab could show that SND1 acts as a proviral host factor. This RBP is known to associate with other proteins and form protein-protein complexes that could be assembled through the Tudor domain of SND1 [248-252]. To understand the protein interactors of SND1 and how those are influenced by SARS-CoV-2 infection, I employed quantitative MS after co-IP to detect the proteins that associate with SND1 in infected versus uninfected cells. This data shows a shift from proteins associated with RNA processing, splicing, and regulation to host and viral proteins that associate with membranes. The interaction with nsp3 and nsp4 hints towards a redistribution of SND1 to the DMVs in SARS-CoV-2 infected cells. In addition, SND1 associates with nsp9, which is NMPylated by the NiRAN domain of nsp12. The NiRAN domain attaches an NMP to the first N-terminal amino acid of nsp9. By omitting the crucial UV cross-linking step in the eCLIP method, the modified cRIP experiment shows for the first time covalent attachment of RNA to nsp9 *in cellulo*. The most prominent RNA peaks detected correspond to the pos and neg sense orientation of the 5´ leader sequence and the 5´ end of the neg sense RNA. Additionally, comparing the distribution of the RNA covalently attached to nsp9 in the presence or absence of SND1 revealed that SND1 modulates the distribution of nsp9 on the coronaviral RNA.

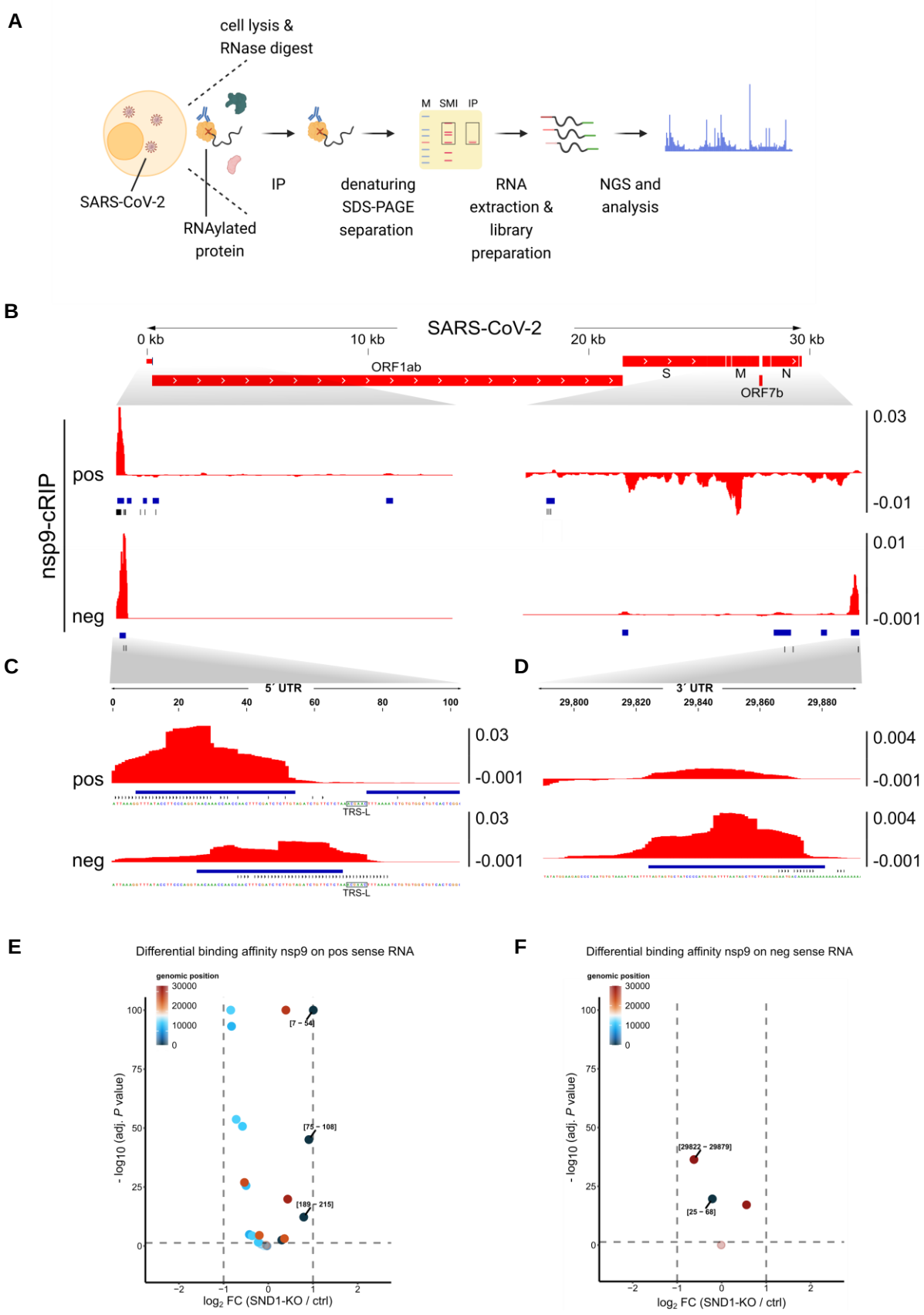
A
cell lysis &
RNase digest
SARS-CoV-2
RNAylated
protein
IP
M SMI IP
denaturing
SDS-PAGE
separation
RNA
extraction &
library
preparation
NGS and
analysis
B
SARS-CoV-2
0 kb
10 kb
20 kb
30 kb
ORF1ab
S
M
N
ORF7b
nsp9-cRIP
pos
neg
0.03
-0.01
0.01
-0.001
C
D
5´ UTR
3´ UTR
0
20
40
60
80
100
29,800
29,820
29,840
29,860
29,880
0.004
TRS-L
E
Differential binding affinity nsp9 on pos sense RNA
F
Differential binding affinity nsp9 on neg sense RNA
genomic position
30000
20000
10000
0
- log10 (adj. P value)
log2 FC (SND1-KO / ctrl)
[7 – 54]
[75 – 108]
[188 – 215]
[29822 – 29879]
[25 – 68]

Figure 8 RNAylation of nsp9 in authentic viral replication is modulated by SND1

A Experimental outline of the covalent RNA immunoprecipitation method (cRIP). Illustration created with BioRender.com **B** $A549^{ACE2}$ were infected for 12 h with SARS-CoV-2 and the lysate subjected to the cRIP experiment. Combined input subtracted read probability of IP from two independent experiments is aligned to the SARS-CoV-2 genome and zoomed into the 5' end (left) and the 3' end (right) of the viral RNA genome. Reads were separated according to their orientation to the pos sense and neg sense RNA track. Below the track, significant peaks are depicted in blue and predicted covalent linkages marked in black. **C+D** Closer zoom in on the 5´ UTR (**C**) and 3´ UTR (**D**) of alignment depicted in **B**. **E+F** Differential binding affinity analysis of significant peaks in presence or absence of SND1 as log_2 FC and the corresponding significance for the pos sense RNA (**E**) and neg sense RNA (**F**). SARS-CoV-2 infection performed by Dr. Nora Schmidt. Data processing and bioinformatical analysis performed by Dr. Alex Gabel

3. Discussion

3.1. Considerations on UV-based methodologies

To investigate RNA-protein interactions, UV-based cross-linking methods are powerful tools that increased our knowledge about the general understanding of RBPs and their influence on host and viral RNAs [31, 135, 147, 175-177, 179-181, 188, 197, 198, 275-277]. Preserving the physiological state of RNA-protein interactions at zero-distance in otherwise unmodified cells or tissues provides a great benefit for this field of research [276, 277]. However, the multitude of variables dictating cross-linking efficiency creates biases that need to be considered in the interpretation of such experiments.

The principle of photo-induced cross-linking lies in the excitation of an electron into a higher energetic state. If this excitation occurs in proximity to a reactive biomolecule, a covalent bond is formed within less than a microsecond (µs)[143, 166]. This extremely fast reaction is extremely specific since it requires direct interactions of reactive biomolecules within the time window of excitation [276, 278, 279].

Aside from the general cross-linking biochemistry, the nature of the RNA-protein interaction introduce biases in the efficiency to form covalent bonds between amino acid and nucleic acid. While all amino acids are considered to engage in cross-linking some have different efficiencies in doing so [279, 280]. An additional bias is introduced by the nucleic acid engaging in the reaction, since pyrimidines, especially uracil have a better cross-linking efficiency *in vitro* [276, 279].

Optimal conditions for UV cross-linking require close association of the protein to the nucleobase. Therefore, other types of association, like binding to the sugar residue, the phosphate backbone, as well as secondary structures (dsRNA) impact UV cross-linking efficiency negatively [276].

Furthermore, the stability of the RNA-protein interaction acts as an additional factor that dictates cross-linking efficiency. As briefly mentioned above, the reaction time of the covalent linkage is <1 µs [166]. Transient protein-RNA interactions can therefore be difficult to capture due to the short pulsed excitation for cross-linking [276, 281].

Unfortunately, UV cross-linking is extremely inefficient (~ 5 %) and requires a high amount of starting material to be employed in the methods [138, 146, 176].

While the above mentioned challenges have to be considered for all methods employing UV-crosslinking, eCLIP has additional limitations that need to be addressed when it comes to analyzing the sequencing results obtained by eCLIP [201, 282]. As for all IP-based methods, the availability of specific antibodies targeting the endogenous protein is crucial, since this might be a source of contamination within the sample [201, 283]. This is particularly challenging when good quality antibodies targeting the protein of interest are sparse. This can be bypassed through exogenous introduction of a tagged version of the target protein as shown for SARS-CoV-2 [284]. In this work, tagged versions of the viral proteins were expressed in a cell culture system and without infection context and the binding profiles on human RNA investigated. While this can provide insights into potential regulatory functions of viral proteins on host RNA metabolism, the higher complexity introduced through infection is neglected. With newly formed compartments as well as interactions between host and viral RNAs and protein-protein interactions through infection with SARS-CoV-2, potential functional mechanisms could be missed in such an experimental design.

The second consideration to be done in eCLIP experiments is for the type of RNase used, as RNases could create different footprints due to sequence biases [285]. Since RBPs can additionally bind DNA, it is important to perform DNase digestion and use specific ssRNA/DNA ligases to avoid background introduces by remaining DNA fragments.

The success of an eCLIP experiment also lies in the abundance of the investigated protein and the bound RNA. Especially when it comes to investigate RBPs of viral RNA, choosing the correct timing of the experiment is important. In addition, dynamic changes in the course of viral infection can change the availability of proteins through host shutoff and the influence the formation of protein complexes [10, 29, 286].

In order to confirm true binding sites and quantify the obtained data, choosing the right control is pivotal. While historically, IgG-IP has been used as a negative control in eCLIP experiments, these rarely create a library usable for sequencing [287]. The commonly used control that allows for normalization of eCLIP data is the SMI creating knowledge

about the non-specific background of the sample [283]. One downside of the SMI as a background control comes with extremely abundant RBPs, like the SARS-CoV-2 N protein, that can create unspecific reads thereby negatively influencing normalization. In addition, RNA-seq data can be used as negative control to correct for transcript abundance [201].

Due to technical biases, and potential differences of RBP-RNA interactions especially when it comes to comparing different RBPs or different experimental conditions, quantification can become a challenging task. Using spike-in controls can help to normalize the differences in libraries and quantify sample differences [288].

Multiple studies have investigated the protein interactome of SARS-CoV-2 RNA but lack the complementary detailed investigation of the candidate RBPs [25, 130, 167]. RBPs as central players of RNA metabolism have a key relevance in viral infection [289]. By finding the binding site of candidate SARS-CoV-2 RBPs this project has used the previously published SARS-CoV-2 RNA protein interactome to validate several candidate proteins and delineate their binding profile on the SARS-CoV-2 RNA [31]. The here generated eCLIP dataset is the so far largest collection of binding profiles targeting host and viral RBPs in the context of SARS-CoV-2 infection. Still, the binding profile alone is insufficient to elucidate the relevance of a RBP in viral replication. For each of these targets a system needs to be established that allows to study changes in infection in dependence of the target RBP [31, 174]. Combining the knowledge about binding profiles of RBPs together with the relevance in infection can help to guide an investigation about the molecular mechanisms of the target RBP in SARS-CoV-2 infection as shown in this study for SND1.

3.2. RBPs as regulators of transcription and translation

As a key member of the RTC, nsp12 provides the enzymatic core of viral RNA synthesis. One would assume that the binding coverage of nsp12 would spread evenly throughout the genome. Additionally, as nsp12 requires neg sense RNA as a template for transcription and replication binding of nsp12 to the neg sense RNA is expected [90, 93, 97, 290, 291]. Instead, small binding sites are detected at the ends of the viral genome. The replicase complex elongates the nascent RNA strand at a speed of up to 260 nucleotides per second [291]. The short dwelling time at a single location could reduce the likelihood to capture

those RNA with UV cross-linking methods. A similar phenomenon has been shown for the ribosomal protein L7 (RPL7), where despite the known relevance of this protein as an RBP of the ribosomal complex, no cross-linking to RNA could be detected [135, 281, 292]. The prominent binding site of nsp12 on the 5´ end of the pos sense RNA could results from an increased binding stability during RTC assembly, thus increasing cross-linking efficiency. Another factor that counteracts low cross-linking efficiencies is high material abundance. This could explain the nsp12 binding site detected in the N gene, since this transcript is one of the most abundant viral mRNAs [117]. The nsp12 eCLIP experiment shows that despite the caveats mentioned above RNA binding can be detected.

The main function of the viral protein N lies in packaging of gRNA for viral particle formation [66, 67]. This is reflected by the binding coverage of N over the whole gRNA in the eCLIP dataset. Additional reports claim further biological relevance of N in the course of coronaviral infection. In SARS-CoV-2 the N-terminal domain of N binding a TRS-L sequence could interact with nsp13 in the fully assembled RTC [293]. Further potential mechanisms of N in RNA synthesis have been shown in transmissible gastroenteritis coronavirus (TGEV) infection, a member of the alphacoronaviruses, stating that N associates with the RTC at the core sequence and might be involved in the regulation of template switching [294]. The binding profile provided in the eCLIP does not allow for a separate investigation of the different mechanisms, therefore the potential relevance of N in RNA synthesis can´t be distinguished from the gRNA packaging function of N in the provided eCLIP dataset. This might be resolved by choosing an earlier time point than 24 hpi in which viral particle formation is not overshadowing other potential mechanistic roles of N.

Next, I investigated nsp9, which through *in vitro* experiments has been reported as an unspecific ssRNA binding protein [111, 112]. Curiously, despite the report of no binding preference, the nsp9 binding profile shows a stronger binding preferences to the 5´ end of pos sense SARS-CoV-2 RNA and both ends of the neg sense RNA. An additional interesting finding is that despite the low abundance of nsp9 and neg sense RNA strong binding profiles could be obtained indicating a strong specificity of this interaction [117, 295]. With the association of nsp9 with the neg sense RNA, one can assume that nsp9 might be relevant in viral RNA synthesis. *In vitro* studies revealing nsp9 as a target of

NiRAN mediated NMPylation suggest potential relevance of nsp9 in priming and capping [95, 99-102].

The strongest host RBP in our core interactome dataset, CNBP, is a G-quadruplex binding protein with antiviral properties in SARS-CoV-2 infected cells [24, 31, 211, 213, 214]. As a transcriptional co-regulator, it activates and maintains the innate immune response against viral infections [214]. While some show a regulatory role of CNBP on gene expression in the immune response to viral infection, others report a regulatory effect of CNBP on translation [211, 213, 296]. CNBP contacts the SARS-CoV-2 RNA exclusively on the pos sense RNA, which could hint to translational regulation of CNBP. A recent study has investigated the presence of G-quadruplex structures in SARS-CoV-2 RNAs [297]. The predicted G-quadruplex structures of the pos sense viral RNA are located in close proximity to peaks detected in the eCLIP dataset. Integrating the here provided eCLIP dataset with bioinformatical prediction algorithms, could help to find the binding motif of CNBP on SARS-CoV-2 RNA, as it has been recently been performed for the prediction of RBP binding sites [182, 298]. G-quadruplex structures are known to inhibit reverse transcriptases and influence translation efficiency negatively [299, 300]. The consequence of CNBP binding to G-quadruplex structures is debated, as some report unfolding of the structure G-quadruplex structures in translation and others claim stabilization in response to gene regulation [211, 213, 297]. To investigate the relevance of CNBP in SARS-CoV-2 infections, methods that investigate the transcriptional and translational changes upon infection could be employed [172, 301]. Transcriptional changes can be measured by thiol(SH)-linked alkylation for the metabolic sequencing of RNA (SLAM-seq). Newly synthesized RNA is detected through mutational changes in the sequencing data that are introduces through the incorporated 4SU [172]. This allows to investigate changes in expression dynamics that could be compared to cells containing or lacking CNBP. The same cell model system could be used to employ ribosome profiling. This method enriches for ribosomes containing protected RNA footprints. The footprints are afterwards recovered and sequenced to gain information about the ribosome bound mRNA [301]. Pretreating the cells with chemical compounds targeting ribosomes helps to gain a detailed understanding of the translation status of a certain mRNA. As an example, cycloheximide is used to halt elongation or harringtonine to block ribosomal initiation [301,

302]. Therefore, changes in translational initiation or active translation on an mRNA can be detected.

Several members of the HNRNP family can be found within the extended core RNA interactome. HNRNPA2B1 is an abundant cellular protein that plays a role in multiple steps of the RNA metabolism like splicing, translation and RNA trafficking [179, 215, 216]. In addition, a study investigating HIV transcription could find a proviral effect of HNRNPA2B1 by unfolding G-quadruplex structures in the LTR promotor thus enhancing viral transcription [303]. Recent work in SARS-CoV-2 infected cells has shown, that successful knock-down of HNRNPA2B1 results in reduced SARS-CoV-2 RNA levels and virus production revealing it as a proviral host factor [210]. The eCLIP dataset reveals exclusive pos sense RNA binding in intergenic regions. Comparing the binding sites of a known G-quadruplex binding protein CNBP and HNRNPA2B1 shows a partial overlap close to regions predicted to contain G-quadruplex structures (Figure 4B). It is possible that HNRNPA2B1 could bind and destabilize the G-quadruplex on SARS-CoV-2 and thereby enhance translation and replication efficiency. Additionally, the small molecule pyridostatin (PDS) stabilizes G-quadruplex formation and inhibits SARS-CoV-2 replication further emphasizing the necessity of G-quadruplex formation in viral replication [304, 305]. To support the descriptive nature of the binding profiles generated with eCLIP, further experimental evidence for the binding of CNBP and HNRNPA2B1 at G-quadruplex structures is needed. While there are methods to investigate G-quadruplex structures with their corresponding RBP using fluorine-19 nuclear magnetic resonance spectroscopy, this method relies on the microinjection of fluorine-19 labeled RNA [306]. A second possibility would be to use the fluorescent probe V-P1 that supposedly binds exclusively RNA G-quadruplex structures [307]. Using this probe in cell-based systems would allow the investigation of CNBP and HNRNPA2B1 binding to such structures with imaging-based techniques. As this would constitute only indirect evidence, one could complement these findings *in vitro* with the optical tweezer assay [308]. This single molecule technique can be used to traps RNA fragments known to carry G-quarduplex structures and investigate changes in secondary structures in presence of RBPs. Understanding the binding motifs of CNBP and HNRNPA2B1 on the SARS-CoV-2 RNA could reveal a potential interesting antagonism of a proviral and antiviral host factor in SARS-CoV-2 infection.

Another strong factor directly interacting with SARS-CoV-2 RNA is LARP4, a poly(A) binding RBP that protects mRNA from deadenylation, and degradation, thereby regulating its translation [225, 226]. In agreement with published work, we can see LARP4 binding in the 3´ UTR together with an additional binding site in the 5´ UTR. The spliced transcripts alignment to a reference (STAR) algorithm used to map sequencing reads allows to detect splice junction reads but does not allow for multi-mapping [309]. Thus if sequences map to multiple locations on the SARS-CoV-2 genome, which would be the case for poly(A) binding proteins, the reads would be discarded. Therefore the absent signal in the poly(A) tail is most likely a consequence of data processing and not of biological activity. Manual curation and adaptation to a more lenient multi-mapping parameter might allow for the detection of reads mapping to the poly(A) tail. An actively translated host mRNA through association with RBPs is thought to form a circular structure that facilitates faster re-association of the ribosomal complexes to the mRNA [310-312]. Commonly, the circularization is mitigated via eIF4E-eIF4G-PABP interaction, of which the latter two are also detected in our SARS-CoV-2 RAP-MS dataset [31]. LARP4 binding to the ends of the SARS-CoV-2 genome could hint at the likely case that the same mechanisms of host mRNA translation occurs for SARS-CoV-2 mRNAs. To confirm that LARP4 is relevant for the translation of SARS-CoV-2 RNA, ribosome profiling could be employed in a cell culture model that depicts changes of viral replication induced by LARP4.

The relevance of SND1 in RNA metabolism is multifaceted, ranging from splicing, transcription, RNA stability to stress granule formation [253-259]. Different studies show SND1 as a member of different RNA-protein or DNA-protein complexes as co-transcriptional activator, co-regulator of splicing or potential endonuclease of miRNAs in RISC [242, 243, 255, 256, 313]. While analysis of the eCLIP experiment revealed few binding sites in the pos sense RNA, the majority of reads correspond to the 5´ end of the neg sense RNA, a region, which was also covered by nsp9. This, to my knowledge, is the first report of a host RBP binding to neg sense viral RNAs. Since negative sense viral RNA serves as the template during SARS-CoV-2 RNA replication and transcription, it is intriguing to hypothesize a potential relevance of SND1 for viral RNA synthesis. In addition, it has been reported that SND1 recognizes and binds m^6A containing mRNAs as well as hyper-edited dsRNAs [244, 261]. However, if the binding profile detected is determined by

posttranscriptional modifications requires closer investigation of the epitranscriptome, meaning the sum of all posttranscriptional modifications, on neg sense SARS-CoV-2 RNA.

Overall, several candidate host RBPs might have a regulatory influence on coronaviral RNA transcription or translation. Despite the large amount of literature available for a multitude of the candidate RBPs, it is relevant to keep in mind that upon viral infections the usual role of an RBP might change. Revealing the relevance of this RNA-protein interaction in infection as well as the molecular mechanism behind it remains necessary for a thorough investigation of the RNA-RBP interplay.

3.3. Host RBPs with relevance in innate immunity

Successful viral replication if not contained by the immune system poses a threat to the individual's survival. Once a cell is infected, it can use genetically encoded innate immune factors to eliminate or suppress a viral infection. Such factors of innate immunity can act as pattern recognition receptor (PRR), sensing pathogen-associated microbial patterns (PAMP) and inducing the expression of genes that help in the combat against infection. The interferon-stimulated gene (ISG) IFIT5 is one example of a PRR. It has an adaptable recognition specificity for 5´ppp, 5´p or cap-0 on the 5´end of RNA and modulates nuclear factor kappa-light-chain-enhancer of activated B cells (NF-κB) as well as interferon regulatory factor (IRF) 3 signaling positively [223, 314, 315]. The here provided eCLIP data shows no binding of IFIT5 to the 5´ end of SARS-CoV-2 RNA. The most prominent binding site in the genomic locus of S does not align with the known 5´ end binding of IFIT5. One possibility to investigate the relevance of IFIT5 in SARS-CoV-2 infection could be to investigate the signaling pathways downstream of IFIT5 activation. In addition, it has been shown, that IFIT5 associates with the cytosolic dsRNA sensing PRR retinoic acid-inducible gene I (RIG-I) [315]. It would be interesting to see, if SARS-CoV-2 infection is recognized by IFIT5 and induces antiviral signaling upstream of RIG-I. Fittingly, co-localization of RIG-I with IFIT5 has already been shown [222]. A study in MHV could show that active viral replication activates NF-κB and IRF3 translocation through the dsRNA recognizing PRRs RIG-I and melanoma differentiation-associated protein 5 (MDA5) [316]. Therefore it might be possible that as viral replication progresses, dsRNA

that is sheltered in DMVs could be recognized by IFIT5. Activated IFIT5 in combination with RIG-I could induce antiviral signaling.

PPIA has a multitude of different biological functions with respect to innate immunity, ranging from signal transduction to protein shuttling between the nucleus and cytosol [317-319]. In IAV infection, PPIA interacts with the M1 protein and impairs the early stages of replication [230]. Opposing to that, a proviral effect of PPIA was reported through the interaction of the capsid protein in HIV infection [231]. Similarly, others report the conflicting role of PPIA in viral infections [233-235, 320-322]. While the published data focus on protein mediated mechanisms of PPIA in viral infection, it appears as a non-canonical RBP in an mRNA-bound proteome study, as well as our SARS-CoV-2 RNA–protein interactome dataset [31, 135]. The here provided eCLIP data validates the RNA binding capabilities of PPIA and depicts strong binding in genes located at the 3´ end of the genome. PPIA is a proviral factor in SARS-CoV-2 infection but the precise mechanism in which PPIA supports viral replication remains to be elucidated [31, 235]. It is especially interesting to see that the immunosuppressive drug cyclosporin A (CysA) inhibits coronaviral replication [31, 235]. The PPIA-CysA complex binds and inhibits calcineurin, a phosphatase that activates the nuclear factor of activated T cell cytoplasmic (NFATc) through dephosphorylation [227, 323]. It would be interesting to investigate, if the interaction of PPIA with CysA influence the SARS-CoV-2 RNA-binding ability of PPIA. Additionally, the influence of calcineurin as a potential antiviral factor is an interesting target for further research.

Mounting an antiviral immune response by the host cell poses evolutionary pressure on a replicating virus to avoid its elimination. This arms race between the host immune response and viral infection can pressure a virus to silence antiviral pathways or modulate the activity of antiviral RBPs. One might be inclined to interpret the RBP-RNA interaction as an antiviral response to the infection but the opposite might be the case, as seen for the opposing role of PPIA depending on the type of viral infection. It is crucial to investigate such interactions in close detail to fully understand the nature of these interactions.

3.4. Unexpected RBPs

Conventional methods using domain architectures to predict RBPs have drastically underestimated the true extent of RBPs, thus UV cross-linking methods elevate our current understanding of true RBPs and the protein domains relevant for RNA interaction [135, 324]. Using the same strategy of UV cross-linking, our core interactome dataset has revealed unexpected RBPs, out of which the binding sites of S and RTCB on SARS-CoV-2 RNA I have validated in this work.

S is one of the most studied proteins in SARS-CoV-2 since it is the key viral protein for viral entry. In addition, in context of viral particles S is exposed to antibody mediated clearance [325]. This exposure provides a selective pressure for the accumulation of mutations to avoid antibody recognition, while at the same time maintaining its functionality [326]. Intriguingly, in addition to these well-established functions, we found S to be one of the strongest core interactors in our SARS-CoV-2 RNA interactome. Several publications predict that the C-terminal domain of S reaches into the viral particle where it could interact with the gRNA [208, 327]. Using eCLIP, I could detect several binding sites and thus support the provided experimental evidence of an RNA binding capability of S [31]. This binding might be detected as an artifact of high protein abundance combined with close proximity of S and the genomic RNA in viral particles and might not necessary relate to a functional relevance of S as a RBP. To gain further insights into this interaction it would be necessary to find the RNA interacting peptide fragment of S, which can be performed by mass spectrometric experiments like RNA binding domain mapping (RBDmap) and validate if absence or mutations of the interacting peptide influence RNA binding of S [324].

RtcB is a potential antiviral factor and a strong candidate in our core SARS-CoV-2 RNA interactome [24, 31]. As the ligase subunit of the RNA-splicing complex, it joins spliced RNA fragments with broad substrate specificity [238, 328]. RtcB binds exclusively to the pos sense SARS-CoV-2 RNA and does not show any specific binding preference, leaving its relevance for the viral replication an open question. Mutation of the active site H428 might give evidence if the enzymatic ligase activity of RTCB affects SARS-CoV-2 replication [329].

In summary, the systematic validation of viral RBPs reveals specificity of the RBP binding patterns. The eCLIP data reveals that the majority of our candidate RBPs bind pos sense RNA, with the exception of SND1 and nsp9 that show binding to neg sense RNA, a key intermediate for RNA synthesis and replication. It is important to consider that more neg sense RNA binding proteins could be present in SARS-CoV-2 infected cells. Separate investigation of the neg sense RNA interactome is crucial to enhance our knowledge about RBPs regulating SARS-CoV-2 RNA synthesis. The finding of SND1 as a host RBP to bind neg sense RNA is a novelty and implicates a potential role in viral RNA synthesis. Revealing this unique binding pattern prompted further investigation to validate the relevance of SND1 in SARS-CoV-2 infection.

3.5. Involvement of SND1 in RNA synthesis of SARS-CoV-2 replication

The neg sense SARS-CoV-2 RNA serves as a template for RNA synthesis and is required for the replication of gRNA and the transcription of mRNA. Synthesis of neg sense RNA starts at the 3´ end of the template RNA with a so far unknown priming mechanism. As it is known, that the replicase complex requires a primer for RNA synthesis, it has been debated if a multimeric structure of nsp7 and nsp8 might act as a primase and initiate priming *de novo* (see section 1.5.3) [96, 97, 106-109]. Therefore, one can assume that first a RNP complex required for priming and initiation needs to be assembled at the correct site of the viral RNA. After successful priming, RNA synthesis would require assembly of the RTC at the 3´ end of the template RNA resulting in the synthesis nascent RNA in 5´ to 3´ directionality by RdRp [96]. The UTRs of SARS-CoV-2 are highly structured and engage in short and long-distance RNA-RNA interactions that are relevant for replication, which might be supported by RBPs [110, 330, 331]. With the proviral host factor SND1 interacting with proteins as well as its binding specificity to neg sense viral RNA could therefore play a supportive role for the assembly of RPCs needed for RNA synthesis. SND1 has the strongest phenotype early in infection where inefficient RNA synthesis could result in reduced RNA levels throughout infection. While others have suggested antiviral properties of SND1 in a knock-down assay, we prove that complete ablation of SND1 on protein level leads to reduced amounts of SARS-CoV-2 RNA, which is rescued upon reintroduction of SND1 [130, 174]. In addition, we show in so far unpublished work that early viral RNA synthesis is impaired in SND1-KO cells, by

comparing the levels of the total RNA level with the newly synthesized RNA using SLAM-seq [174]. While the total level of viral RNA early upon infection is not affected by absence of SND1, the amount of newly synthesized RNA is reduced. Thus, SND1 plays an important role in RNA synthesis early in infection.

SND1 contains multiple nucleic acid binding domains (SN1-SN5) and a Tudor domain. While one tandem repeat of SN domains mediates the nucleic acid binding ability of SND1, the Tudor domain binds dimethylated arginine (DMA) and is relevant for the assembly of RNP complexes [248, 250-252, 261]. The Tudor domain of SND1 binds preferentially symmetric dimethylarginine (SDMA), a posttranscriptional modification (PMT) that is not reported for SARS-CoV-2 proteins [255, 313]. Curiously, MTDH interacts with SND1 via the SN1-2 domains [332]. Therefore, it is possible that other domains of SND1 engage in protein-protein interactions. Investigating the protein complex surrounding SND1, this project reveals, that the interaction partners of SND1 change drastically upon infection. While in naïve cells, factors associated with pre-mRNA splicing, RNA stability, and ribosomal proteins interact with SND1, SARS-CoV-2 infection changes the interaction partners of SND1 to membrane-associated proteins and viral replicase proteins. Curiously, the nuclear host protein Histone Cell Cycle Regulator (HIRA) appears as a novel protein interactor upon SARS-CoV-2 infection, a factor that in genome-wide CRISPR screens is revealed as a strong antiviral factor [24, 25]. HIRA is responsible for the regulation of gene expression and acts as an antiviral factor against herpes simplex virus 1 (HSV-1) [333]. While HIRA was not detected as a direct RBP in published interactomes, it interacts with SND1 and exhibits antiviral properties on SARS-CoV-2 [24, 25, 31]. Since SND1 interacts with SARS-CoV-2 RNA as well as HIRA it would be interesting to validate HIRA as an SARS-CoV-2 RNA associated protein. By combining UV cross-linking with chemical protein-protein cross-linkers like disuccinimidyl suberate (DSS) further knowledge about indirect RNA binding proteins or larger RPCs could be obtained [334]. Additionally it would be interesting to compare if HIRA is recruited by SND1 or if HIRA recognizes the viral initiation RPC. A deeper understanding of the protein-protein interaction on a molecular level might explain the role of HIRA in SARS-CoV-2 infection.

In infected cells, SND1 interacts with the viral membrane proteins nsp3 and nsp4, which are key components for the formation of replication compartments. By rearranging the

ER-membrane, they form DMVs and convoluted membranes [90-92]. These segregated compartments are the key location of viral RNA replication and connect the DMV lumen with the cytoplasm through a pore formed by nsp3 and nsp4 [90, 92]. This pore allows the passage of newly synthesized viral mRNA into the cytosol for translation. By interacting with nsp3 and nsp4 a fraction of SND1 molecules is closely associated with the DMVs and might play a relevance for the localization of the RTC at the pore complex allowing an efficient release of mRNA into the cytosol. Another factor that associates with SND1 and could hint toward replication is N. As already mentioned above, the primary function is binding of gRNA for viral particle formation though it could have additional relevance in the regulation of template switching [293, 294].

Interestingly, SND1 also interacts with nsp9, a factor associated with priming and capping that binds the same stretch of neg sense RNA as SND1.

3.6. The role of nsp9 as protein primer in SARS-CoV-2 RNA synthesis

The NiRAN domain of nsp12 covalently attaches a NMP to the first amino acid ASN of nsp9. This reaction leads to a covalent bond between the NMP to the first amino acid ASN of nsp9 through a phosphoramidate linkage with the 5' carbon of the nucleic acid [95, 98-102]. Several different studies investigating the NMPylation reaction *in vitro* have reported different biases on the nucleotide engaging in this reaction. While the majority of these studies report a preference to incorporate U and assume a relevance of nsp9 in protein-priming, two others that additionally report de-NMPylation of nsp9 suggest additional function in capping [95, 98-100]. A large portion of the work investigating viral replication is performed in *in vitro* studies, trying to delineate molecular mechanisms in a simplified and controllable reaction condition. While it allows focusing on single molecular reactions, the biggest drawback of *in vitro* studies is the neglect of the higher complexity structures and dynamic changes occurring in authentic infection. Changing the availability of protein interaction partners or separating them physically through compartmentalization may additionally allow viral proteins to exert multiple functions. Employing an adapted eCLIP protocol that allows the investigation of naturally occurring covalent linkages, this project shows covalently attached SARS-CoV-2 RNA on nsp9 in authentic viral infection. Looking at the covalently linked nucleotides confirms that each type of ribonucleotide (U,

A, C, G) is predicted to be attached to nsp9, like reported in the *in vitro* studies [95, 99, 100, 102]. The strongest enriched RNA sequence in the cRIP experiment corresponds to the leader sequence on the pos sense RNA located at the 5´ end of the SARS-CoV-2 RNA. This sequence of RNA is present on each sgRNA as a result of template switching. The TRS plays an important role in the synthesis of sgRNA through template switching (Figure 3B) and could act as a driving force of recombination [41, 42]. Curiously, the read probability distribution in the cRIP dataset reveals an abrupt drop shortly before the leader-TRS. Seeing these drops around the TRS could indicate junction-spanning reads. One would expect to see a mirrored pattern of these reads at the beginning of each gene, which is only detectable for the loci S and ORF3a. No other genomic locus depicts significant peaks at the beginning of the gene, which could be an artifact resulting of insufficient enrichment over the input control or low abundance of junction-spanning reads.

The second intriguing peak is found on the neg sense RNA aligning to the 3´ UTR which reaches into the poly(A) tail. This region corresponds to the 5´ end of the neg sense RNA and contains the suspected initiation site of neg sense viral RNA synthesis. The sequencing results reveal a multitude of cross-linking sites corresponding to the poly(U). Corresponding to the preference for UTP as the substrate for NMPylation reported in the *in vitro* studies this could support a potential relevance of UMPylted nsp9 in neg sense RNA priming [95, 99]. With the covalent attachment of nsp9 to both, the pos and neg sense 5´ ends of SARS-CoV-2 RNA could hint to a relevance of nsp9 as a protein-primer for RNA synthesis. The biases depicted in the *in vitro* studies (preferred covalent attachment of U > A > G > C) give an additional hint towards a potential relevance in priming, since the first nucleotide on all SARS-CoV-2 pos sense RNAs is an A and the first nucleotide on the neg sense RNA an U. Therefore, the NMPylated nsp9 could base pair through the covalently attached nucleotide with the corresponding template RNA and act as a protein-primer for the elongation of RNA synthesis by the SARS-CoV-2 replicase. A similar mechanism of protein priming has already been described for the poliovirus VPg. UMPylated VPg acts as a protein-primer for poliovirus RNA synthesis but remains attached at the 5´ end of the poliovirus gRNA. Since the poliovirus gRNA contains an internal ribosome entry site (IRES), it does not require a cap structure for the initiation of translation [103-105]. The high accumulation of reads corresponding to the pos sense

leader also hints at the potential relevance of RNAylated nsp9 as a substrate for the capping machinery. The current research allows to interpret that nsp9 maintains multiple functions in the initial phase of RNA synthesis. As suggested by *Wang et al*. 2022 nsp9 might be involved in priming as well as in capping of new viral mRNA [95]. As coronaviral mRNAs contain a *N*7- and 2′-*O*-methylated cap it suggests itself that the covalently attached nsp9 has to be removed prior to translation [95, 100]. We hypothesize, that after successful priming of viral RNA replication, the replicase could use the nucleotide covalently attached to nsp9 to start with the elongation of the nascent RNA. Once the RNA is elongated by the assembled RTC, a free nsp12 could remove the covalently attached elongating RNA strand from nsp9 and attach a GDP to create the GpppA cap structure needed to initiate the capping reaction as demonstrated *in vitro* in the studies of Park *et al.*, Wang *et al.* and Walker *et. al.* [95, 100, 126]. A second possibility of nsp9 removal could be the endonucleolytic activity of nsp15 on poly(U) [335]. In this scenario, nsp15 would cleave nsp9 from the nascent RNA leaving it in an oligo-UMPylated form. The poly(U)-nsp9 could then either prime successive rounds of RNA synthesis or be recycled through degradation of the oligo(U) attached to nsp9 by nsp15 and subsequent de-NMPylation by NiRAN [95, 100, 335]. Whether the neg sense RNA can be capped remains an open question, as no proof for or against has been provided yet. It is possible, that for neg sense RNA, nsp9 could stay attached to act as a 5´ end protector to avoid recognition by PRRs like IFIT5.

A third interesting binding site of nsp9 in the eCLIP and cRIP experiments map to the locus 25-68 indicating the synthesis of an “anti leader” RNA. The synthesis of an anti leader sequence has to my knowledge not been reported. Based on the lack of any translational start sites in the pos sense orientation of this short RNA stretch, it is unlikely that any protein-coding relevance can be attributed. On the other hand, the presence of an antisense RNA that base pairs to the leader sequence could indicate a regulatory function that would require further investigation. Since the sequencing results obtained only preserve a short fraction that aligns to parts of the anti leader the existence this RNA would need experimental validation.

Recently, Slanina *et al.* 2023 have compared the preservation of nsp12 NiRAN mediated nsp9 NMPylation across the family of *Coronaviridae* [102]. This work revealed that the

NMPylation of nsp9 by nsp12 NiRAN is conserved in *Orthocoronavirinae* and that the homologous nsp12 proteins in HCoV-229E, SARS-CoV-1 and TGEV can NMPylate the noncognate nsp9 of the other viruses. This finding emphasizes the fundamental relevance of nsp9 NMPylation. The conserved nature of nsp9 NMPylation could imply that the protein-priming mechanism of nsp9 discussed here might have a universal relevance in viral replication of Coronaviruses.

The here provided cRIP data demonstrates that SARS-CoV-2 RNA is covalently linked through the 5´end to nsp9 in authentic viral replication. The publically available and herewith provided data about coronaviral RNA synthesis suggest that nsp9 could act as a protein-primer to initiate RNA synthesis while simultaneously provide the substrate for the first steps in cap synthesis.

3.7. Modulation of protein priming by SND1

The proviral host factor SND1 investigated in this study interacts with nsp9 presumably close to the replication organelles. In this study, I show that absence of SND1 leads to a shift in the distribution of covalently attached nsp9 at the 5' ends of pos and neg sense viral RNA. Thus, SND1 could modulate the correct distribution of nsp9 on the viral RNA and increase the efficiency of RNA synthesis (Figure 9). Since priming and capping are especially relevant at the beginning of the viral replication cycle, this could explain why we see the strongest phenotype of SND1 depletion early in infection before RNA synthesis reaches exponential growth (Figure 5C). One possible explanation of how SND1 could modulate the redistribution is its binding preference on the SARS-CoV-2 RNA. SND1 binds preferably to the 5´ end of neg sense viral RNA and could provide a sequence specificity for nsp9, which binds RNA irrespective of the sequence [111, 112]. Studies have shown, that one tandem repeat of SN domains is sufficient for RNA binding and that the SN domains in addition can engage in protein-protein interactions [261, 332]. In fact, a domain deletion experiment in our current manuscript in preparation reveals that the SN3 domain of SND1 is required for the interaction with nsp9 [174]. Thus, the SN1-SN2 domains could be required for the binding specificity of SND1 to the neg sense viral RNA, while SN3 recruits the UMPylated nsp9 to increase the concentration of nsp9 for efficient pos sense RNA synthesis (Figure 9). It is possible, that the other domains could recruit further

proteins to the 5´ end of the negative strand RNA. This could be mediated through the recognition of certain structures or motifs, like DMA modifications for Tudor [248, 250-252, 255, 261, 313]

My thesis provides foundational work for the functional interrogation of a multitude of SARS-CoV-2 RBPs. The revelation of a unique neg sense RNA binding specificity of the proviral host factor SND1 led to the discovery of an infection induced remodeling of protein interaction partners. Upon infection, SND1 interacts with proteins of the coronaviral RTC, including the ssRNA binding protein nsp9. Additionally, this work shows nsp9 RNAylation *in cellulo* and that SND1 modulates the occupancy of nsp9 on the 5´ ends of SARS-CoV-2 RNA. We hypothesize that after synthesis of neg sense RNA, SND1 binds at its 5´ end recruiting nsp9 to ensure correct accumulation of nsp9 at sites relevant for initiation of RNA synthesis (Figure 9). Absence of SND1 results in an imbalance of nsp9 availability at the correct template and thus decreases efficiency of successive replication rounds. As an unspecific ssRNA binding protein, a guided accumulation of nsp9 by SND1 at the correct sites would be crucial for efficient priming of viral RNA synthesis. This project is the first to provide compelling *in cellulo* data for the involvement of protein-priming by nsp9 as the primary step of RNA replication in coronaviruses.

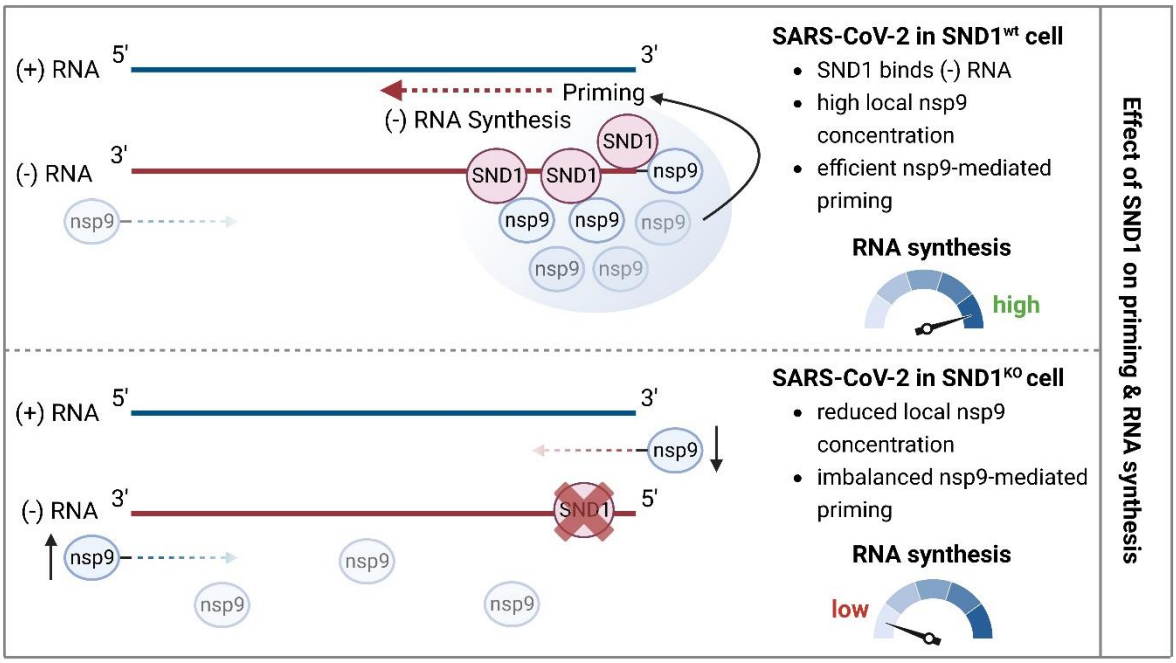

Figure 9 Graphical summary of the working hypothesis on nsp9 mediated protein-priming supported by SND1

Our hypothesis is that NMPylated nsp9 initiates the first round of replication by association with the 3´ ends of the template RNAs. The availability of neg sense RNA allows specific binding of SND1 to the 5´ end of neg sense (-) RNA and increases the local concentration of nsp9 at the 5´ end. This increases the efficiency of nsp9 mediated priming for the synthesis of viral pos sense (+) RNA. Absence of SND1 leads to a reduced local concentration of nsp9, an imbalanced covalent nsp9-RNA binding and subsequent reduced RNA synthesis. Created by Dr. Nora Schmidt and Prof. Mathias Munschauer in BioRender.com [174].

4. Material & Methods

4.1. Material

Table 4 List of antibodies and the used dilutions

Antibody	RRID	Working condition	Provider
actin antibody	Cat# A2103, RRID:AB_476694	1:500 (WB)	Abcam
CNBP antibody	Cat# 67109-1-Ig, RRID:AB_2882413	5 µg per mg total protein (IP)	Proteintech
HIST1H1C antibody	Cat# 19649-1-AP, RRID:AB_10694432	5 µg per mg total protein (IP)	Proteintech
HNRNPA2B1 antibody	Cat# 14813-1-AP, RRID:AB_2279638	5 µg per mg total protein (IP)	Proteintech
IFIT5 antibody	Cat# 13378-1-AP, RRID:AB_10640300	5 µg per mg total protein (IP)	Proteintech
IRDye® 680RD Goat anti-Mouse IgG Secondary antibody	926-68070	1:2000 (WB)	LI-Cor (Freitext)
IRDye® 680RD Goat anti-Rabbit IgG Secondary antibody	926-68071	1:2000 (WB)	LI-Cor (Freitext)
IRDye® 800CW Goat anti-Mouse IgG Secondary antibody	926-32210	1:2000 (WB)	LI-Cor (Freitext)
IRDye® 800CW Goat anti-Rabbit IgG Secondary antibody	926-32211	1:2000 (WB)	LI-Cor (Freitext)

LARP4 antibody	Cat# 16529-1-AP, RRID:AB_2296671	5 µg per mg total protein (IP)	Proteintech
Normal Rabbit IgG antibody	Cat# 2729, RRID:AB_1031062	5 µg per mg total protein	Cell Signaling Technology
PPIA antibody	Cat# ab58144, RRID:AB_941220	5 µg per mg total protein (IP)	Abcam
RTCB antibody	Cat# 19809-1-AP, RRID:AB_10695047	5 µg per mg total protein (IP)	Proteintech
SARS-CoV-2 (COVID-19) nsp12 antibody	Cat# GTX135467, RRID:AB_2887496	5 µg per mg total protein (IP)	GeneTex
SARS-CoV-2 (COVID-19) nsp9 antibody	Cat# GTX135467, RRID:AB_2887496	5 µg per mg total protein (IP); 1:2000 (IF)	GeneTex
SARS-CoV-2 (COVID-19) nucleocapsid antibody	Cat# GTX135357, RRID:AB_2868464	5 µg per mg total protein (IP)	GeneTex
SARS-CoV-2 (COVID-19) S antibody	Cat# ab272504, RRID:AB_2847845	5 µg per mg total protein (IP)	Abcam
SND1 antibody	Cat# A302-883A, RRID:AB_10631268	5 µg per mg total protein (IP)	Bethyl Laboratories
SND1 antibody	Cat# 60265-1-Ig, RRID:AB_2881386	1:500 (IF); 1:300 (WB)	Proteintech

Table 5 List of consumables

Consumables	Identifier	Provider
2x Q5 Hot start master mix	M0494L	New England Biolabs

96 Fast PCR-Platte Vollrand	72.1980.202	Sarstedt
AmpureXP bead suspension	A63881	Beckman Coulter
BD Micro-Fine + 324892 0.5ml Insulin Syringe with 29G x 12.7mm Needle	324892	BD
Chamber slide with a removable 12 well c	81201	ibidi
Direct-zol™ RNA MiniPrep	R2053	Zymo Research
Dynabeads Protein G for Immunoprecipitation, 5 ml	10003D	Invitrogen
Dynabeads™ MyOne™ Silane, 5 ml	37002D	LIFE Technologies
Ex-gel 2% SYBR gold	G402002	LIFE Technologies
Fluoromount-G Mounting Medium-25 mL	00-4958-02	LIFE Technologies
Hard-Shell(R) 384-Well PCR Plates, thin wall, skirted, clear/clear	HSP3801	Bio-Rad Laboratories
HyClone water	10307052	Cytiva
iBind Cards-40 cards	SLF1010X4	LIFE Technologies
iBind Solution Kit-1 kit	SLF1020	LIFE Technologies
NEBNext® rRNA Depletion Kit	E7400X	New England Biolabs
Novex™ iBlot™ 2 Transfer Stacks, nitrocellulose, mini	IB23002	ThermoFisher Scientific
NuPAGE™ 3 bis 8 %, Tris-Acetat	EA0378BOX	Invitrogen

NuPAGE™ 4 to 12%, Bis-Tris	NW04120BOX	Invitrogen
Phase Lock Gel™	733-2478	VWR International
Ponceau solution	33427.01	Serva Electrophoresis
Qubit dsDNA HS Assay Kit	Q32854	LIFE Technologies

Table 6 List of chemicals and reagents

Chemicals and Reagents	Identifier	Provider
0.1 M DTT	3860-5GM	Merck Millipore
0.5M EDTA	A4892.1000	VWR International
0.5M TCEP	646547-10X1ML	Sigma-Aldrich
100 mM ATP	R0441	ThermoFisher Scientific
100 mM dNTPs (25 mM each)	200415	Agilent
100% DMSO	41639-100ML	Sigma-Aldrich
10x PBS	AM9624	LIFE Technologies
1M MgCl2	11693289	Th. Geyer
1M Tris-HCl pH 7.5	11823244	Th. Geyer
1M Tris-HCl pH 8.5	11823240	Th. Geyer
20X Bolt MES SDS Running Buffer-500 mL	B0002	LIFE Technologies
20X Bolt MOPS SDS Running Buffer-500 mL	B0001	LIFE Technologies

4X Bolt LDS Sample Buffer	B0007	LIFE Technologies
5M NaCl	A7006.1000	VWR International
Buffer RLT (220 ml)	79216	Qiagen
CaCl2	21114-1L	Th. Geyer
cOmplete™, Mini, EDTA-free Protease Inhibitor Cocktail	15672129	Fisher Scientific
DAPI	D1306	LIFE Technologies
DMEM, high glucose, GlutaMAX Supplement, pyruvate-10 x 500 mL	31966047	LIFE Technologies
Fetal Bovine Serum, qualified, heat inactivated, E.U.-approved, South America Origin-500 mL	10500064	LIFE Technologies
IGEPAL CA-630	I8896-100ML	Sigma-Aldrich
KCl	6781.3	Roth - Carl Roth
NaDeoxycholate	30970-25G	Sigma-Aldrich
N-lauroylsarcosine	SAFSL7414-50ML	VWR International
NuPAGE Tris-Acetate SDS Running Buffer (20X)-500 mL	LA0041	LIFE Technologies
Phenol – chloroform – isoamyl alcohol mixture	77619	Sigma-Aldrich
Pierce Bovine Serum Albumin Standard Ampules, 2 mg/mL	23209	LIFE Technologies
Trichlormethan/Chloroform ROTIPURAN®, min. 99 %, p.a.	3313.4	Roth - Carl Roth

Triton X-100	T9284-100ML	Sigma-Aldrich
Trypsin-EDTA (0.25%), phenol red-100 mL	25200056	LIFE Technologies
Tween-20	P9416-50ML	Sigma-Aldrich

Table 7 List of enzymes and kits

Enzymes and Kits	Identifier	Provider
Affinity Script Enzyme	M0494S	Agilent
Duolink®proximity ligation assay(PLA®)	DUO92014-30RXN	Sigma Aldrich
FastAP Thermosensitive Alkaline Phosphatase 1 U/µL	EF0651	Thermo Scientific
Murine RNase inhibitor	600109	New England Biolabs
Pierce™ BCA™ Protein-Assay	23225	Thermo Scientific
Proteinase K	P8107S	New England Biolabs
Q5 Hot Start High-Fidelity 2X Master Mix	M0494L	New England Biolabs
Qubit dsDNA HS Kit	Q32851	Invitrogen
RNA Clean & Concentrator-5	R1016	Zymo Research
RNase I 100 U/µl	AM2295	Invitrogen
T4 PNK enzyme	M0201L	New England Biolabs
T4 RNA ligase high concentration	M0437M	New England Biolabs

TURBO DNase 2 U/µL	AM2239	Invitrogen
Zymoclean Gel DNA Recovery Kit	D4007	Zymo Research

Table 8 List of buffer and media

Buffer and Media	Ingredients
10× RNA Ligase Buffer without DTT	500 mM Tris-HCl,pH 7.5, 100 mM MgCl2
10x FastAP buffer	100 mM Tris-HCl, 50 mM MgCl2, 1 M KCl, 0.2% Triton X-100
10x RNA ligase Buffer with DTT (NEB)	500 mM Tris-HCl pH7.5, 100 mM MgCl2, 100 mM DTT
200x salt mix	0.5 M MgCl2, 0.1 M CaCl2
2x Co-IP lysis buffer	100 mM Tris-HCl pH 7.5, 300 mM NaCl, 2% Igepal-CE630, 1% NaDeoxycholate, 0.5 mM TCEP, Proteinase K inhibitor
2x Laemmli Buffer	4% SDS, 20% Glycerol, 120 mM Tris-HCl pH 6.8, 0.02% w/v Bromphenolblue, 40 mM DTT
4x LiDS sample buffer	4x LDS sample buffer, 50 mM TCEP
Co-IP wash buffer	150 mM NaCl, 50 mM Tris pH 7.5
Co-IP wash buffer WB	150 mM NaCl, 50 mM Tris pH 7.5, 0.1% Sod Deoxycholate, 0.1% IGEPAL
complemented DMEM	DMEM + 10 % FBS, 1% P/S
eCLIP lysis buffer	50 mM Tris-HCl pH 7.5, 150 mM NaCl, 1 mM EDTA, 1% (v/v) IGEPAL CA-630, 0.5% NaDeoxycholate, 0.25 mM TCEP, complete EDTA-free protease inhibitor cocktail
IP no salt buffer	50 mM Tris-HCl pH 7.5, 1 mM EDTA, 0.5% (v/v) IGEPAL CA-630
IP wash buffer	50 mM Tris-HCl pH 7.5, 300 mM NaCl, 1 mM EDTA, 1% (v/v) IGEPAL CA-630, 0.5% NaDeoxycholate, 0.25 mM TCEP

NLS elution buffer	20 mM Tris-HCl pH 7.5, 10 mM EDTA, 2% N-lauroylsarcosine, 2.5 mM TCEP
PBS 0.02 %Tween-20	0.02 % v/v Tween-20 in PBS

Table 9 List of oligonucleotides

Oligos	Sequence	Stock [80]	Provider
18S_fwd	5'-ATGGCCGTTCTTAGTTGGTG-3'	100	Microsynth
18S_rev	5'-GAACGCCACTTGTCCCTCTA-3'	100	Microsynth
2P_Universal	5´-AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCT-3´	25	Schmidt et al. 2020
AR17	CAGACGTGTGCTCTTCCGA	20	IDT
Barcoded Primer	5´-CAAGCAGAAGACGGCATACGAGATNNNNNNNNGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT-3´	25	Schmidt et al. 2020
N_fwd	5'-CACATTGGCACCCGCAATC-3'	100	Microsynth
N_rev	5'-GAGGAACGAGAAGAGGCTTG-3'	100	Microsynth
Rand103Tr	/5Phos/NNNNNNNNNNAGATCGGAAGAGCGTCGTGT/3SpC3/	80	Eurofins Genomic
random 9mer primer		100	Microsynth
RdRp_fwd	5'-GTGARATGGTCATGTGTGGCGG-3´	100	Microsynth

RdRp_rev	5´-CARATGTTAAASACACTATTAGCATA-3'	100	Microsynth
RiL-19	/5Phos/rArGrArUrCrGrGrArArGrArGrCrArCrArCrGrUrC/3SpC3/	40	IDT

Table 10 List of Cell lines

Cell line	Provider
A549^{ACE2}	kindly provided by Prof. Andreas Pichlmair
A549^{ACE2} SND1-ctrl	kindly provided by Yuanjie Wei and Dr. Nora Schmidt
A549^{ACE2} SND1-KO	kindly provided by Yuanjie Wei and Dr. Nora Schmidt
A549^{ACE2} SND1-rescue	kindly provided by Yuanjie Wei and Dr. Nora Schmidt
Vero-E6-TMPRSS2	Prof. Stefan Pöhlmann

Table 11 List of Instruments and Software

Instruments and Software	Provider
Centrifuge 5425	Eppendorf
Centrifuge 5430 R	Eppendorf
GS Gene Linker	Bio-Rad Laboratories
E-Gel Power Snap	ThermoFisher Scientific
FIJI	Schindelin et al. 2012
HulaMixer	ThermoFisher Scientific

iBind	ThermoFisher Scientific
iBlot2	ThermoFisher Scientific
Integrative Genomics Viewer (V.2.3.8)	Robinson et al. 2011
JACoP	Bolte, Cordeliéres et al. 2006
Leica TCS SP5	Leica
MiniAmp Thermocycler	ThermoFisher Scientific
MiniSeq	Illina
Odyssey DLx	Licor
Pipet-lite XLS	Rainin
QuantStudio 5	ThermoFisher Scientific
QuantStudio TM Design & Analysis Software v1.5.1	ThermoFisher Scientific
Qubit 4 Fluorometer	ThermoFisher Scientific
ThermoMixer C	Eppendorf
Bandelin Sonopuls HD 3100 Ultraschall-Homogenisator	Sonoplus
Thunder Imager Live Cell 3D Assay	Leica

4.2. Methods

4.3. Cell culture

The human $A549^{ACE2}$ and Huh-7 cell lines were cultivated in DMEM (Thermo Fisher Scientific) supplemented with 10 % heat-inactivated fetal bovine serum (FBS), 100 U/ml streptomycin, and 100 mg/ml penicillin. Cultivation was performed at 37 °C at an atmosphere containing 5 % CO_2. The $A549^{ACE2}$ cells are a kind gift from Prof. Pichlmair. For virus stock generation, Vero-E6-TMPRSS2 (African green monkey) cells that were kindly gifted by Prof. Pöhlmann were cultured as described above.

4.4. Virus stock preparation

BSL-3 work was performed by Dr. Nora Schmidt and Sebastian Zielinski. To generate the SARS-CoV-2 stock first, a patient isolated was passaged in Vero-E6-TMPRSS2 three times. Next, Vero-E6-TMPRSS2 cells were infected at an MOI 0.01 PFU/cell with an inoculum incubation of 2 h at 37 °C. The inoculum was removed and cells were cultured in OptiMem (Gibco) supplemented with 1 % FBS for 48-72 h. After infection, the supernatant was harvested and cleared by centrifugation for 10 min, 500 g at 4 °C, aliquoted, and stored at -80°C. To determine the viral titers, a plaque assay was performed on Vero-E6-TMPRSS2 cells as described previously [31].

4.5. Enhanced crosslinking and immunoprecipitation- sequencing (eCLIP-seq) and covalent RNA immunoprecipitation- sequencing (cRIP-seq)

Infection and lysate preparation for eCLIP-, and cRIP-seq

To delineate the protein-binding site on RNA at various conditions, eCLIP was performed with slight modifications from the van Nostrand et al. 2016 protocol [180]. In addition we have developed a method based on eCLIP to investigate natural covalent RNA-protein modifications (cRIP-seq). For both, 24 million cells per immunoprecipitation (IP) were infected with SARS-CoV-2 at MOI 3 PFU/ml for the time period indicated in the experiment. Subsequently for eCLIP, the media was removed, washed with ice-cold 1x PBS and the cells UV-crosslinked (254nm, 0.8 J/cm^2 for BSL-3 infection) and scraped in ice-cold 1x PBS. For both experiments, cells were pelleted by centrifugation (200 g, 8 min, 4 °C), the supernatant discarded and the pellet resuspended in 2x eCLIP lysis buffer (100 mM Tris-HCl pH 7.5, 300 mM NaCl, 2 mM EDTA, 2 % IGEPAL, 1 % sodium

deoxycholate, 0.5 mM TCEP, EDTA-free Protease Inhibitor Cocktail) for eCLIP-seq experiments, or 2x Co-IP buffer (100 mM Tris-HCl pH 7.4, 300 mM NaCl, 2 % IGEPAL, 1 % sodium deoxycholate, 0.5 mM TCEP, EDTA-free Protease Inhibitor Cocktail) for cRIP-seq. After 30 min incubation at room temperature, the sample was extracted from the BSL-3 and immediately processed for IP.

Lysate processing and RNA extraction

After extraction from the BSL-3 the samples were diluted to a final concentration of 1x eCLIP lysis buffer (50 mM Tris-HCl pH 7.5, 150 mM NaCl, 1 mM EDTA, 1 % IGEPAL, 0.5 % sodium deoxycholate, 0.25 mM TCEP, EDTA-free Protease Inhibitor Cocktail) by addition of one volume RNAse free H_2O. The non crosslinked sampels were supplemented with 0.5 M EDTA to a final concentration of 1 mM. Subsequently the lysates were sonicated (total 2 kJ, 10 % amplitude, 0.7 s on, 2.3 s off) and a limited digest performed with RNase I (1:100 of 1:25 Stock dilution) for 10 min at 22 °C. Next, DNA was digested with TURBO™ DNase (1:250) in 2.5 mM $MgCl_2$ and 0.5 mM $CaCl_2$ for 20 min at 37 °C. Remaining debris was removed by centrifugation at 5,000 g for 15 min at 4 °C. With a BCA assay the total protein concentration was determined and per mg total protein 30 µl of magnetic Protein G beads (Invitrogen) used for pre-clear. The Protein G beads were washed twice in PBS 0.02 % Tween-20 before transferring them to the lysate for 45 min at 4 °C for pre-clear. After incubation 5 % of the total lysate was saved for a size-matched Input (SMI) control prior to IP. The full list of antibodies used for IP is provided in Table 4. For every mg of total protein, 5 µg antibody were coupled to 30 µl magnetic Protein G beads for 1 h at room temperature. To the pre-cleared lysate, the antibody-coupled beads were transferred for overnight incubation at 4 °C rotating. After IP, the beads were washed twice with eCLIP lysis buffer without Proteinase Inhibitor (50 mM Tris-HCl pH 7.5, 150 mM NaCl, 1 mM EDTA, 1 % IGEPAL, 0.5 % sodium deoxycholate, 0.25 mM TCEP), IP wash buffer (50 mM Tris-HCl pH 7.4, 300 mM NaCl, 1 mM EDTA, 1 % IGEPAL CA-630, 0.5 % NaDeoxycholate, 0.25 mM TCEP) and IP no salt buffer (50 mM Tris-HCl pH 7.4, 1 mM EDTA, 0.5 % IGEPAL CA-630) each. The beads were rinsed once with 1x FastAP buffer (100 mM Tris-HCl p.H 8.0 at 37 °C, 50 mM $MgCl_2$, 1 M KCl, 0.2 % Triton X-100) and dephosphorylated with FastAP enzyme (39 µl RNase-free H_2O, 5 µl 10x FastAP buffer, 1 µl murine RNase inhibitor, 5 µl FastAP enzyme) for 20 min at 37 °C, 800 rpm on the

ThermoMixer C. Afterwards, to ensure complete dephosphorylation, a treatment with T4 PNK (120 µl RNase-free H_2O, 20 µl 10x PNK buffer, 2 µl murine RNase inhibitor, 7 µl T4 PNK enzyme, 1 µl TURBO™DNase) was performed for 20 min at 37 °C. Following end-repair, 1 ml cold IP no salt buffer was added to the beads and incubated for 5 min on ice, allowing stabilization of the antibody-antigen interaction. Then, the beads were washed once with 1 ml 1x RNA ligase buffer (50 mM Tris-HCl pH 7.5, 10 mM $MgCl_2$) and then resuspended with the RNA ligation enzyme mix (13 µl RNase-free H_2O, 3 µl 10x RNA ligase buffer without DTT, 0.3 µl 0.1 M ATP, 0.8 µl 100 % DMSO, 9 µl 50 % PEG8000, 0.4 µl murine RNase inhibitor, 2.5 µl T4 RNA ligase high concentration). For the first 3´ adapter ligation 1 µl 40 µM RiL-19 was added to the beads and incubated for 1 h 15 min at 23 °C with intermitted shaking for 15 s at 1800 rpm on the ThermoMixer C. Then, 1 ml cold IP no salt buffer was added to the beads, which was incubated for 5 min on ice before denaturing the proteins in 40 µl 1x LiDS sample buffer (1x Bolt™ LDS sample buffer, 12.5 mM TCEP). The SMI sample was supplemented 1:4 with 4x LiDS sample buffer (4x Bolt™ LDS sample buffer, 50 mM TCEP). All samples were boiled at 80 °C for 10 min. The samples were loaded without beads for size separation by SDS-PAGE (depending on the target size, gradient 4-12 % Bis-Tris gel with MOPS buffer or MES buffer or a 3-8 % Tris-acetate gel with Tris-acetate buffer) side by side with the SMI. After a transfer on a nitrocellulose membrane (iBlot2, Program 0), the protein- RNA complex was excised from the expected height up to 75 kDa above. The same area was cut from the SMI, which provides information about all protein-protected RNA fragments, running at this height. The RNA was eluted from the excised membrane fragments with a Proteinase K digest (25 µl RNase-free H_2O, 12.5 µl 5 M NaCl, 200 µl NLS elution buffer (20 mM Tris-HCl pH 7.4, 10 mM EDTA, 2 % N-lauroylsarcosine, 2.5 mM TCEP), 12.5 µl Proteinase K), which was subsequently isolated with 250 µl acid-phenol:chloroform isoamyl mix and the help of Phase Lock Gel™ tubes. After rigorous vortexing and centrifugation for 5 min at 13,000 g, the aqueous phase is transferred into a new Phase Lock Gel™ tube and again vortexed with 250 µl chloroform. The last step was repeated to ensure removal of any remaining protein fragments. Subsequently, the RNA was cleaned up using the ZYMO RNA clean and concentrator kit. For this the samples were mixed with 500 µl RNA binding buffer and 1,125 µl 100 % EtOH. The whole

volume was loaded three times on the columns, washed once with 400 µl RNA prep buffer, 700 µl RNA wash buffer and 400 µl RNA wash buffer, each time centrifuging for 30 s at 16,000 g. The columns were then dried for 2 min at 21,000 g and eluted in 11 µl RNase-free H_2O, centrifuging at 16,000 g for 30 s and reloading the volume for a total of three times. Next, the SMI samples were end-repaired with FastAP and T4 PNK. As above the FastAP reaction (10 µl RNase-free H2O, 2.5 µl 10x FastAP buffer, 0.5 µl murine RNase inhibitor, 2.5 µl FastAP enzyme) was performed for 20 min at 37 °C, shaking at 800 rpm on the ThermoMixer C. Then another incubation for 20 min at 37 °C were performed for complete dephosphorylation with T4 PNK (55 µl RNase-free H_2O, 10 µl 10x PNK buffer, 1 µl 0.1 M DTT, 1 µl murine RNase inhibitor, 7 µl T4 PNK enzyme, 1 µl TURBO^{TM}DNase). The reaction was then cleaned with 20 µl silane beads (Life Technologies), which were washed twice with RLT buffer and resuspended in a total volume of 300 µl RLT buffer to be added to the sample. Together with 10 µl 5 M NaCl and 615 µl 100 % EtOH, the samples were incubated for 15 min at room temperature, rotating slowly. The samples were washed three times with 75 % EtOH before drying at 37 °C. The beads were resuspended in 11 µl RNase-free H_2O for 5 min at room temperature. 5 µl of the eluted SMI RNA was used for the RiL-19 adapter ligation. The RNa was mixed with 1.5 µl 100 % DMSO and 0.5 µl 40 µM RiL-19 and denatured at 65 °C for 2 min. The RNA was cooled immediately on ice for 1 min and incubated for adapter ligation (1.5 µl RNase-free H_2O, 2 µl 10x RNA ligase buffer with DTT, 0.3 µl 100 % DMSO, 0.2 µl 0.1 M ATP, 8 µl 50 % PEG8000, 0.2 µl murine RNase inhibitor, 1.3 µl T4 RNA ligase high concentration) for 1 h 15 min at 23 °C with intermitted shaking for 15 s at 1800 rpm on the ThermoMixer C. Another silane cleanup was performed, with 20 µl beads, which were washed twice in RLT and resuspended in 61.6 µl RLT buffer. After addition of 61.6 µl 100 % EtOH, RNA was bound for 15 min at room temperature. The beads were then washed three times with 75 % EtOH prior to drying at 37 °C and elution with 11 µl RNase-free H_2O for 5 min at room temperature.

cDNA synthesis and library preparation

For cDNA synthesis, IP and SMI samples were denatured at 65 °C for 2 min and immediately cooled on ice for 1 min. The reverse transcription (4 µl RNase-free H_2O, 2 µl 10x Affinity Script buffer, 2 µl 0.1 M DTT, 0.8 µl 100 mM dNTPs (25 mM each), 0.3 µl murine RNase inhibitor, 0.9 µl Affinity Script enzyme) was performed for 45 min at 55 °C. To remove leftover primer, 3.5 µl ExoSAP-IT was added to the sample and incubated for 13 min at 37 °C. To stop the reaction, 1 µl 0.5 M EDTA. For further inactivation of ExoSAP-IT and degradation of RNA , the samples were heated for 12 min at 70 °C containing 3 µl 1 M NaOH. To adjust the pH, 3 µl 1 M HCl were added after the reaction. Afterwards another silane clean-up was performed. For this, 10 µl beads were washed twice with 1 ml RLT buffer and in 93 µl RLT buffer added to the sample. After addition of 111.6 µl 100 % EtOH, the cDNA was bound to the beads for 5 min at room temperature. Subsequently, the beads were washed three times with 80 % EtOH and dried at 37 °C for 5 min. The 3´ adapter ligation was performed on the silane beads, therefore elution took place with the Rand103Tr3 adapter (5 µl RNase-free water, 0.8 µl Rand103Tr3 adapter, 1 µl 100 % DMSO) for 5 min at room temperature. Prior to enzyme addition, the samples were heated for 2 min at 75 °C and immediately put on ice. Afterwards, the enzyme mix (1.1 µl RNase-free water, 2 µl 10x RNA ligase buffer with DTT, 0.2 µl 0.1 M ATP, 9 µl 50 % PEG8000, 1.5 µl high concentration T4 RNA ligase) was added with stirring motions the beads were incubated over night at 22 °C with intermitted shaking for 15 s at 1800 rpm on the ThermoMixer C. The next day, 5 µl of washed silane beads (Life Technologies) were resuspended in 60 µl RLT buffer and added to the sample. After addition of 60 µl 100 % EtOH, cDNA was bound for 5 min at room temperature. The beads were washed three times in 75 % EtOH and dried at 37 °C for 5 min. Then, beads were resuspended in 45 µl RNase-free water for 5 min. As a next step, 21 µl of the eluted cDNA was amplified and barcoded with PCR (25 µl 2x Q5 Hot start master mix, 2 µl 2P Universal Primer) with a 30 s 98°C denaturation step, 6 cycles of amplification with low temperature annealing (15 s 98 °C, 30 s 68 °C, 40 s 72 °C) and 9 cycles with high temperature annealing (15 s 98 °C, 60 s 72 °C), following with a final extension for 5 min 72 °C. Immediately afterwards the samples were cooled at 4 °C and 2 µl samples were used for a QC gel. Depending on the library quality, a maximum of 5 further high temperature

annealing cycles were added to a total maximum of 20 amplification cycles. For a final purification step, the libraries were run on a gel and the regions between 175 bp and 300 bp excised. The libraries were recovered using the Zymo DNA gel recovery kit. For this the gel fragments were dissolved in 3x the weight as volume agarose dissolving buffer (ADB) (e.g. 300 µl ADB for 100 mg gel fragments) for 10 min at 55 °C and the whole volume loaded thrice onto the columns, each time spinning for 30 s at 13,000 g. The column was washed twice with 200 µl DNA wash buffer and dried for 2 min at 21,000 g in a fresh tube to reduce buffer contaminations. Then, the libraries were eluted in 11 µl RNase-free water and reloaded twice, such that the whole volume passes the column three times. Lastly, the DNA concentration was determined by the Qubit HS DNA kit, with 1 µl of the library. If applicable, the libraries were diluted to a 4 nM pool for Illumina sequencing.

eCLIP and cRIP data analysis

Basic procedures for read preprocessing and read alignment SARS-CoV-2 eCLIP are based on the work of Nostrand et al. [180]. Paired-end sequencing reads are adapter- and quality trimmed using cutadapt (v1.18) [341]. Reads with a total length less than 18 nt are discarded. Unique molecular identifiers (UMI) are removed from the paired-end reads and stored for later de-duplication. Afterwards, paired reads are aligned by the splicing aware alignment program STAR [309]. All reads are strand specifically aligned to the human genome (hg38) and the SARS-CoV-2 genome to delineate RNA fragments derived from the pos. or neg. sense RNA strand. Based on the paired-read alignment and the UMI information, reads were de-duplicated to remove artificial amplification introduced by PCR.

The detection of binding regions, peaks, was performed by MACS2 [273]. To confirm the detected peaks a one-sided Fisher´s exact test (H1: reads in IPpeak > greater than expected) was used. The resulting *p values* were corrected for multiple testing by using the Benjamini-Yekutieli method [272]. Only peaks with a corrected p value below 0.05 were considered as IP-enriched. Visualizing the overall eCLIP signal with respect to highlight enriched regions in IP over SMI, the relative information content between IP and SMI was computed for each sample [342-344]. The relative information content was calculated as $p_i \times log_2(p_i/q_i)$, where p_i and q_i are defined as the fraction of total aligned

reads in IP and SMI at genome position *i* of the SARS-CoV-2 genome. Regions showing a positive signal are considered as IP-enriched, while negative signals correspond to SMI-enriched regions. A cross-linking site was defined as the first nucleotide before the 5'-end of a R2 read that overlaps a significantly enriched peak. The coverage of each crosslinking-site in IP and SMI was calculated as the number of R2 reads sharing the same 5'-end. Statistically enriched crosslinking sites in IP were calculated by using the Fisher's exact test in the same manner as described above. The alignments were visualized using the Integrative Genomics Viewer (V.2.3.8) [345].

Differential binding analysis of cRIP datasets

To statistically quantify the differences of covalently bound RNA fragments, in SND1-KO and control cells, a differential binding analysis (DBA) was conducted. Therefore, peak calling was performed separately for control and SND1-KO samples by MACS2 [273]. *Bedtools intersect* was used to identify overlapping peak intervals between control and SND1-KO [346]. The common peak intervals were confirmed to be enriched over SMI by the above-described one-sided Fisher's exact test in combination with the Benjamini-Yekutieli *p value* adjustment. After merging the confirmed peak intervals of the control and SND1-KO samples, a two-sided Fisher's exact test was performed to determine differentially binding affinities between the two conditions. Peak intervals showing an adjusted Benjamini-Yekutieli *p value* below 0.05 were considered to show differential binding.

4.6. Western Blot

For Westernblot cells were seeded in a 24-well plate, treated as indicated and harvested upon confluency in 200 µl 1x Laemmli Buffer which was freshly supplemented with DTT. The sample was then transferred into reaction tubes and with the help of the Micro-Fine syringe (BD) DNA sheared. Afterwards, the sample was boiled for 5 min at 90 °C and then run on a SDS gradient gel (Life Technologies) in fresh buffer (MOPS, MES or Tris-acetate (Life Technologies) according to gel type and target size) at 180 V for 30-50 min. The proteins were transfer was perform using the iBlot™2 system (ThermoFisher Scientific) according to the manufacturer protocol using Program 0. Western Blot staining was performed using the iBind System (ThermoFisher Scientific) according to the manufacturer

protocol. The secondary antibody staining was performed with fluorescently labeled secondary antibodies (Li-Cor) according to the iBind manufacturer protocol. As imaging platform, the Odyssey DLx (Li-Cor) was used and imaged with a 169 µm resolution and analyzed using the Software Image Studio Lite version 5.2 (Li-Cor).

4.7. Reverse transcription quantitative Polymerase Chain Reaction (RT qPCR)

RNA obtained from cells after treatment was purified using the Direct-zol™ RNA MiniPrep Kit and eluted in 18 µl RNase-free water. For the reverse transcription, 8 µl of RNA was denatured at 70 °C for 3 min, immediately cooled on ice. Afterwards, 1 µl of random 9mer primer was added, carefully mixed by pipetting and incubated for 10 min at room temperature. 11 µl of the reverse transcription mix (5.2 µl RNase-free water, 2 µl 10x AffinityScript buffer, 0.8 µl 25 mM dNTPs, 2 µl 100 mM DTT, 1 µl AffinityScript RT enzyme) was added to the sample and incubated for 1 h at 55 °C, then 70 °C for 15 min. After the incubation was done, the cDNA was diluted with 100 µl RNase-free water and a separate 1:100 dilution of the cDNA performed for the 18S control. For each target a primer mix was prepared containing equal amounts of forward and reverse primer to be at a final concentration of 25 µM. For the qPCR, per target a PCR mixture was prepared consisting of 0.75 µl 25 µM primer mix and 13.5 µl Luna® Universal qPCR. 14.25 µl of PCR mixture was distributed into a Multiply® PCR Plate (Sarstedt) to which 12.75 µl cDNA was added and mixed by pipetting. 5 µl of the PCR mix was transferred as quadruplicates into a Hard-Shell® 384-Well PCR Plate (Bio-Rad Laboratories) and run on the QuantStudio™ 5 (Thermo Fisher Scientific). Analysis was performed using the QuantStudio Design & Analysis Software version 1.5.1. Normalization was performed according to the $\Delta\Delta C_t$ method [347].

4.8. Co-IP for mass spectrometry (MS)

Sample preparation

The protein-protein interactomes was investigated by IP followed by mass spectrometry. Per condition, 20 mio cells were infected with SARS-CoV-2 at MOI 3 PFU/ml for the duration indicated in the experiment. The cells were washed once with cold PBS and scraped in cold PBS. Afterwards, the cells were spin for 10 min at 500 g for 4 °C and lysed

in 1 ml 2x Co-IP lysis buffer (100 mM Tris-HCl pH 7.4, 300 mM NaCl, 2 % IGEPAL, 1 % NaDeoxycholate, 0.5 mM TCEP). After 30 min of incubation at room temperature, the lysate was diluted to 1x and centrifuged at 14,000 g for 10 min to remove cellular debris. Afterwards, the supernatant was extracted from BSL-3 and the protein concentration determined by BCA assay. Per mg protein, 50 µl magnetic protein G beads were washed twice with PBS 0.02 % Tween-20 for pre-clear. After pre-clear, 1 % of the lysate was saved for QC as Input sample. For antibody coupling, for each mg of total protein, 50 µl washed magnetic protein G beads were diluted to 2x of the original volume and incubated with 5 µg of antibody for 30 min at room temperature. After the pre-clear, the magnetic beads were discarded and the antibody coupled beads transferred for IP at 4 °C over night. The beads were washed twice with Co-IP wash buffer (150 mM NaCl, 50 mM Tris pH 7.5) containing 0.05% Igepal and twice with Co-IP wash buffer without any detergent. The flow through, as well as each wash step was saved for a final QC WB. During the final wash step, 5 % of the bead slurry was taken as eluate for the QC WB. Finally, the beads were resuspended in 50 µl 1x LDS Bolt sample buffer for mass spectrometry by the EMBL Proteomic Core Facility.

Protein digestion and TMT labeling

To avoid run-to-run variation, samples were multiplexed by tandem mass tag (TMT) labelling. For this, disulfide bonds were reduced with 10 mM DTT at 56°C for 20 min in 50 mM HEPES pH 8.5 and alkylated with 20 mM 2-chloroacetamide at 24°C for 20 min in 50 mM HEPES pH 8.5. Subsequently, the samples were cleaned up according to the single-pot solid-phase-enhanced sample preparations (SP3) protocol [348, 349]. Sequencing grade Trypsin (Promega) was added in an enzyme to protein ratio of 1:50 for an overnight reaction at 37°C. Peptides were labelled with TMT10-plex (Co-IP samples) or TMT6-plex (input lysates) Isobaric Label Reagent (ThermoFisher) according to the manufacturer's instructions [350, 351]. Briefly, 0.8 mg of reagent was dissolved in 42 µl 100% acetonitrile to which 8 µl of stock was added and incubated for 1 h room temperature. Quenching was performed for 15 min with 5% hydroxylamine. For the run, samples were combined and desalted on an OASIS® HLB µElution Plate (Waters). Using an Agilent 1200 Infinity high-performance liquid chromatography system equipped with a Gemini C18 column (3 µm, 110 Å, 100 x 1.0 mm, Phenomenex), offline high pH reverse

phase fractionation was performed [352]. Thirty-two fractions were collected and pooled into 6 fractions (Co-IP samples) or 12 fractions (input lysates), dried under vacuum centrifugation, and reconstituted in 10 µl 1% formic acid, 4% acetonitrile for LC-MS analysis.

Quantification and identification of peptides and proteins

To process the raw data, IsobarQuant and Mascot (v2.2.07) were used, which was searched against a combined Uniprot proteome database of *Homo sapiens* (UP000005640) and *SARS-CoV-2* (UP000464024) containing common contaminants and reversed sequences [353]. Carbamidomethyl on cystein and TMT6 on lysine (lysates, for Co-IP: TMT 0) were set as fixed modifications and acetyl (Protein N-term), Oxidation (M) and TMT6 on N-termini (lysates, for Co-IP TMT 10) set as variable modifications. A mass error tolerance of 10 ppm for the full scan (MS1) and a spectra of 0.02 Da for MS/MS (MS2) was allowed. To define the cleavage sites, trypsin was chosen as protease with a maximum of two missed cleavages. A minimum peptide length of seven amino acids and a false discovery rate on peptide and protein level of 0.01 were set as additional parameters.

Statistical data analysis

To process the raw output files of IsobarQuant (protein.txt – files) R was used as programming language. For the analysis, only proteins with at least two unique peptides were considered for the quantitative analysis. Raw TMT reporter ion signals (signal_sum columns) were cleaned for batch effects using limma and further normalized using vsn (variance stabilization normalization) [354, 355]. To maintain the abundance difference, different normalization coefficients were estimated for the IgG and SND1 conditions during the normalization of the IP experiment. The limma package was used to test for differential expression. The replicate information was added as an argument for the ‘lmFit’ function of limma as a factor in the design matrix.

GO-term analysis

After filtering of SND1 interacting proteins, gene IDs of the candidate proteins were submitted to a GO-term analysis using the online tool provided by the Open Biological

Ontologies Foundry (http://geneontology.org/). The analysis tool is based on the PANTHER Classification System [356]. GO biological process was selected as annotation data set and tested with the Fisher´s Exact test and corrected using a FDR cutoff < 0.05. The resulting Go-terms were filtered for an FDR < 0.05 and the highest branch belonging to a group of biological processes selected for display. The number of genes from the implemented list grouping in the same GO biological process were plotted in a bar graph.

4.9. Indirect Immunofluorescence for co-localization analysis

Staining

Approximately $3x10^4$ cells were seeded per well of a 12-well chamber slide (IBIDI, 81201) and infected the next day with SARS-CoV-2 at MOI 3 PFU/cell. At 8 hpi, the cells were washed 1x with PBS and fixed in 4% Paraformaldehyde (Thermo Fisher Scientific, 28908) for 20 min at room temperature. The slides were washed three times with 1x PBS and permeabilized with 0.5 % Triton-X-100 in 1x PBS for 5 min at room temperature. After three more washes with 1x PBS, the cells were blocked for 1 h at room temperature with 5 % FBS in 1x PBS. For each well, 100 µl of primary antibody mixture was prepared in 5 % FBS in 1x PBS and incubated at room temperature for 1 h. The antibody dilutions used are listed in . After three washes, the secondary antibodies labeled with Alexa-488 or Alexa-568 that target the primary antibodies were diluted 1: 1,000 in 5 % FBS in 1x PBS and incubated for 1 h at room temperature in darkness. The slides were washed three times after antibody incubation and stained for 2 min with DAPI (1: 2000 in 1x PBS). The slides were washed three times and once rinsed with ddH_2O. The silicon walls were removed and any access liquid removed prior to mounting with Fluoromount G (Life Technologies). The slides were left to dry in darkness overnight prior imaging and for long-term storage kept at 4 °C.

Microscopy and image processing

The images were obtained on the confocal microscope (Leica TCS SP5) using a 63.0x1.4 oil objective, an argon laser at 30 % power and a scanning speed of 200 Hz. For the nsp3 assay, laser intensities of 405 nm (DAPI), 488 nm (nsp3) and 561 nm (SND1) were set to 15 %, 20 %, and 10 % respectively. The nsp9 assay settings were 30 % 405 nm (DAPI), 20 % 488 nm (nsp9) and 20 % for 561 nm (SND1). For better

visualization, raw images were processed using FIJI [266], and a window corresponding 80x80 µm cropped. Co-localization analysis was performed on raw images with the JACoP plugin with threshold values of 49 for SND1 and 30 for nsp3, and 60 for SND1 and 74 for nsp9 [267]. For the nsp9 experiment, two biological replicates were performed with each two technical replicates. The nsp3 experiment was performed with one biological replicate and two technical replicates.

4.10. Proximity Ligation Assay

Staining

The Duolink®proximity ligation assay (PLA®) (Sigma Aldrich) was performed with $A549^{ACE2}$ cells seeded in 12- well removable slides (ibidi) and infected at various time points. Fixation was performed in 4 % PFA for a total of 20 min at room temperature. The cells were washed twice in 1x PBS and permeabilized in 0.5 % Triton-x-100 in PBS for 5 min at room temperature. Subsequently the cells were washed three times in 1x PBS and blocked in 100 µl Duolink® Blocking Solution per well for 1 h at 37°C in a humidified chamber. The antibody was diluted in 100 µl Duolink® Antibody Diluent. A summary of all used antibody dilutions for the applications are listed in table 4. The staining was performed over night at 4 °C in a humidified chamber. The PLUS and MINUS PLA probes were diluted separately 1:5 in Duolink® Antibody Diluent at a total volume of 100 µl per well. The primary antibody solution was aspirated and subsequently washed three times in 1x Wash Buffer A, incubating each time for 5 min at room temperature. Any excess wash buffer was aspirated and the PLA probe solution added to each well for 1 h at 37 °C. After incubation, the samples were washed thrice in 1x Wash Buffer A, each time incubating 5 min at room temperature. The PLA assay ligase was diluted 1:40 in 1x Ligation buffer. Any excess buffer was removed and 100 µl of ligase mix added. The reaction was incubated for 30 min at 37 °C and subsequently washed three times with 1x Wash Buffer A. In parallel, the 1x Amplification buffer was prepared and kept in darkness. The Polymerase mix was prepared by diluting the PLA polymerase 1:80 in 1x Amplification Buffer and 100 µl added to the aspirated wells for 1 h 40 min at 37 °C in darkness to develop the signal. Afterwards, the slides were washed three times for 10 min at room temperature in 1x Wash Buffer B and once with 0.01x Wash Buffer B for 1 min.

Any excess wash buffer was removed and the slide mounted with the Duolink® In Situ Mounting Medium with DAPI. For this, 5 µl of the mounting medium was added on each cell patch and the coverslip carefully affixed. After 15 min of resting time in darkness at room temperature, the coverslip was sealed with clear nail polish and dried for 15 min at room temperature in darkness. The slides were stored at -20 °C in darkness for up to 6 months and imaged the first time within a week.

Microscopy and image processing

For imaging, the Thunder Imager Live Cell 3D Assay was used with the 40x oil objective. The laser intensity of laser intensities of 435 nm (DAPI) and 488 nm (PLA signal) were set to 5 % and 10 % respectively. Images were acquired with an exposure time of 100 ms for DAPI and 80 ms for the PLA signal. For better visualization, raw images were processed using FIJI [266]

5. References

1. Rivers, T.M., *Viruses and Koch's Postulates.* J Bacteriol, 1937. **33**(1): p. 1-12.

2. Alt, K., *Behandlungsversuche mit Arsenophenylglyzin bei Paralytikern.* Muenchener Medizinische Wochenschrift 1909. **56**: p. 1457-1459.

3. Kaufman, H.E., *Clinical cure of herpes simplex keratitis by 5-iodo-2-deoxyuridine.* Proc Soc Exp Biol Med, 1962. **109**: p. 251-2.

4. Lee, W.W., et al., *POTENTIAL ANTICANCER AGENTS.1 XL. SYNTHESIS OF THE β-ANOMER OF 9-(D-ARABINOFURANOSYL)-ADENINE.* Journal of the American Chemical Society, 1960. **82**(10): p. 2648-2649.

5. Agliano, F., et al., *Long Noncoding RNAs in Host–Pathogen Interactions.* Trends in Immunology, 2019. **40**(6): p. 492-510.

6. Thacore, H. and J.S. Youngner, *Cells persistently infected with Newcastle disease virus. II. Ribonucleic acid and protein synthesis in cells infected with mutants isolated from persistently infected L cells.* J Virol, 1970. **6**(1): p. 42-8.

7. Su, S., et al., *RNA Polymerase and the Shut-off of Host RNA and Protein Synthesis in T4 Phage Infection.* Nature, 1970. **225**(5227): p. 62-63.

8. Snyder, L., *Change in RNA Polymerase associated with the Shutoff of Host Transcription by T4.* Nature New Biology, 1973. **243**(126): p. 131-134.

9. Zuo, J., et al., *The DNase of gammaherpesviruses impairs recognition by virus-specific CD8+ T cells through an additional host shutoff function.* J Virol, 2008. **82**(5): p. 2385-93.

10. Bujanic, L., et al., *The key features of SARS-CoV-2 leader and NSP1 required for viral escape of NSP1-mediated repression.* Rna, 2022. **28**(5): p. 766-779.

11. Lu, X., et al., *SARS-CoV nucleocapsid protein antagonizes IFN-β response by targeting initial step of IFN-β induction pathway, and its C-terminal region is critical for the antagonism.* Virus Genes, 2011. **42**(1): p. 37-45.

12. Kopecky-Bromberg, S.A., et al., *Severe acute respiratory syndrome coronavirus open reading frame (ORF) 3b, ORF 6, and nucleocapsid proteins function as interferon antagonists.* J Virol, 2007. **81**(2): p. 548-57.

13. Hu, Y., et al., *The Severe Acute Respiratory Syndrome Coronavirus Nucleocapsid Inhibits Type I Interferon Production by Interfering with TRIM25-Mediated RIG-I Ubiquitination.* J Virol, 2017. **91**(8).

14. Flint, J., et al., *Principles of Virology, Volume 1: Molecular Biology*. 2020: Wiley.

15. Baltimore, D., *Expression of animal virus genomes.* 1971. **35**(3): p. 235-241.

16. Gorbalenya, A.E., et al., *The new scope of virus taxonomy: partitioning the virosphere into 15 hierarchical ranks.* Nature Microbiology, 2020. **5**(5): p. 668-674.

17. Song, Y., et al., *A genome-wide CRISPR/Cas9 gene knockout screen identifies immunoglobulin superfamily DCC subclass member 4 as a key host factor that promotes influenza virus endocytosis.* PLoS Pathog, 2021. **17**(12): p. e1010141.

18. Yi, C., et al., *Genome-wide CRISPR-Cas9 screening identifies the CYTH2 host gene as a potential therapeutic target of influenza viral infection.* Cell Rep, 2022. **38**(13): p. 110559.

19. OhAinle, M., et al., *A virus-packageable CRISPR screen identifies host factors mediating interferon inhibition of HIV.* Elife, 2018. **7**.

20. Zhang, Q., et al., *Genome-wide CRISPR/Cas9 transcriptional activation screen identifies a histone acetyltransferase inhibitor complex as a regulator of HIV-1 integration.* Nucleic Acids Res, 2022. **50**(12): p. 6687-6701.

21. Zhu, S., et al., *Genome-wide CRISPR activation screen identifies candidate receptors for SARS-CoV-2 entry.* Sci China Life Sci, 2022. **65**(4): p. 701-717.

22. Schneider, W.M., et al., *Genome-Scale Identification of SARS-CoV-2 and Pan-coronavirus Host Factor Networks.* Cell, 2021. **184**(1): p. 120-132.e14.

23. Zhu, Y., et al., *A genome-wide CRISPR screen identifies host factors that regulate SARS-CoV-2 entry.* Nat Commun, 2021. **12**(1): p. 961.

24. Wei, J., et al., *Genome-wide CRISPR Screens Reveal Host Factors Critical for SARS-CoV-2 Infection.* Cell, 2020.

25. Flynn, R.A., et al., *Discovery and functional interrogation of SARS-CoV-2 RNA-host protein interactions.* Cell, 2021. **184**(9): p. 2394-2411.e16.

26. Kuo, R.-L., et al., *Interactome Analysis of NS1 Protein Encoded by Influenza A H7N9 Virus Reveals an Inhibitory Role of NS1 in Host mRNA Maturation.* Journal of Proteome Research, 2018. **17**(4): p. 1474-1484.

27. Mindaye, S.T., et al., *Impact of Influenza A Virus Infection on the Proteomes of Human Bronchoepithelial Cells from Different Donors.* J Proteome Res, 2017. **16**(9): p. 3287-3297.

28. Bogdanow, B., et al., *The dynamic proteome of influenza A virus infection identifies M segment splicing as a host range determinant.* Nature Communications, 2019. **10**(1): p. 5518.

29. Bojkova, D., et al., *Proteomics of SARS-CoV-2-infected host cells reveals therapy targets.* Nature, 2020. **583**(7816): p. 469-472.

30. Li, N., et al., *METTL3 regulates viral m6A RNA modification and host cell innate immune responses during SARS-CoV-2 infection.* Cell Rep, 2021. **35**(6): p. 109091.

31. Schmidt, N., et al., *The SARS-CoV-2 RNA-protein interactome in infected human cells.* Nat Microbiol, 2020.

32. Hornung, V., et al., *5'-Triphosphate RNA is the ligand for RIG-I.* Science, 2006. **314**(5801): p. 994-7.

33. Züst, R., et al., *Ribose 2'-O-methylation provides a molecular signature for the distinction of self and non-self mRNA dependent on the RNA sensor Mda5.* Nat Immunol, 2011. **12**(2): p. 137-43.

34. Gonzalez-Perez, A.C., et al., *The Zinc Finger Antiviral Protein ZAP Restricts Human Cytomegalovirus and Selectively Binds and Destabilizes Viral UL4/UL5 Transcripts.* 2021. **12**(3): p. e02683-20.

35. Hage, A., et al., *The RNA helicase DHX16 recognizes specific viral RNA to trigger RIG-I-dependent innate antiviral immunity.* Cell Reports, 2022. **38**(10): p. 110434.

36. Thorne, L.G., et al., *Evolution of enhanced innate immune evasion by SARS-CoV-2.* Nature, 2022. **602**(7897): p. 487-495.

37. Zhang, F., et al., *Global discovery of human-infective RNA viruses: A modelling analysis.* PLOS Pathogens, 2020. **16**(11): p. e1009079.

38. Al-Tawfiq, J.A., et al., *Surveillance for emerging respiratory viruses.* Lancet Infect Dis, 2014. **14**(10): p. 992-1000.

39. Huang, C., et al., *Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China.* Lancet, 2020. **395**(10223): p. 497-506.

40. Zhou, P., et al., *A pneumonia outbreak associated with a new coronavirus of probable bat origin.* Nature, 2020. **579**(7798): p. 270-273.

41. Chrisman, B.S., et al., *Indels in SARS-CoV-2 occur at template-switching hotspots.* BioData Min, 2021. **14**(1): p. 20.

42. Yang, Y., et al., *Characterizing Transcriptional Regulatory Sequences in Coronaviruses and Their Role in Recombination.* Mol Biol Evol, 2021. **38**(4): p. 1241-1248.

43. *Virology: Coronaviruses.* Nature, 1968. **220**(5168): p. 650-650.

44. Yao, H., et al., *Molecular Architecture of the SARS-CoV-2 Virus.* Cell, 2020. **183**(3): p. 730-738.e13.

45. Woo, P.C., et al., *Discovery of seven novel Mammalian and avian coronaviruses in the genus deltacoronavirus supports bat coronaviruses as the gene source of*

alphacoronavirus and betacoronavirus and avian coronaviruses as the gene source of gammacoronavirus and deltacoronavirus. J Virol, 2012. **86**(7): p. 3995-4008.

46. Christl A Donnelly, A.C.G., Gabriel M Leung, Anthony J Hedley, Christophe Fraser, Steven Riley,, L.-M.H. Laith J Abu-Raddad, Thuan-Quoc Thach, Patsy Chau, King-Pan Chan, Tai-Hing Lam, Lai-Yin Tse, Thomas Tsang,, and J.H.B.K. Shao-Haei Liu, Edith M C Lau, Neil M Ferguson, Roy M Anderson, *Epidemiological determinants of spread of causal agent of*

severe acute respiratory syndrome in Hong Kong. The Lancet, 2003.

47. World Health Organization, W. *SARS case fatality ratio, incubation period*. 2003; Available from: https://www.who.int/emergencies/disease-outbreak-news/item/2003_05_07a-en.

48. Fouchier, R.A.M., et al., *Koch's postulates fulfilled for SARS virus.* Nature, 2003. **423**(6937): p. 240-240.

49. Drosten, C., et al., *Identification of a Novel Coronavirus in Patients with Severe Acute Respiratory Syndrome.* 2003. **348**(20): p. 1967-1976.

50. Kuiken, T., et al., *Newly discovered coronavirus as the primary cause of severe acute respiratory syndrome.* The Lancet, 2003. **362**(9380): p. 263-270.

51. Ksiazek, T.G., et al., *A Novel Coronavirus Associated with Severe Acute Respiratory Syndrome.* 2003. **348**(20): p. 1953-1966.

52. Lau, S.K., et al., *Severe acute respiratory syndrome coronavirus-like virus in Chinese horseshoe bats.* Proc Natl Acad Sci U S A, 2005. **102**(39): p. 14040-5.

53. Li, W., et al., *Bats are natural reservoirs of SARS-like coronaviruses.* Science, 2005. **310**(5748): p. 676-9.

54. Kan, B., et al., *Molecular evolution analysis and geographic investigation of severe acute respiratory syndrome coronavirus-like virus in palm civets at an animal market and on farms.* J Virol, 2005. **79**(18): p. 11892-900.

55. World Health Organization, W. *Middle East respiratory syndrome coronavirus (MERS-CoV)*. 2022; Available from: https://www.who.int/news-room/fact-sheets/detail/middle-east-respiratory-syndrome-coronavirus-(mers-cov).

56. Zaki, A.M., et al., *Isolation of a Novel Coronavirus from a Man with Pneumonia in Saudi Arabia.* 2012. **367**(19): p. 1814-1820.

57. Azhar, E.I., et al., *Evidence for camel-to-human transmission of MERS coronavirus.* N Engl J Med, 2014. **370**(26): p. 2499-505.

58. Biancolella, M., et al., *COVID-19 2022 update: transition of the pandemic to the endemic phase.* Hum Genomics, 2022. **16**(1): p. 19.

59. Antia, R. and M.E. Halloran, *Transition to endemicity: Understanding COVID-19.* Immunity, 2021. **54**(10): p. 2172-2176.

60. Holmes, E.C., et al., *The origins of SARS-CoV-2: A critical review.* Cell, 2021. **184**(19): p. 4848-4856.

61. Lytras, S., et al., *Exploring the Natural Origins of SARS-CoV-2 in the Light of Recombination.* Genome Biology and Evolution, 2022. **14**(2).

62. Afelt, A., et al., *Distribution of bat-borne viruses and environment patterns.* Infect Genet Evol, 2018. **58**: p. 181-191.

63. Smith, E.C., et al., *Coronaviruses lacking exoribonuclease activity are susceptible to lethal mutagenesis: evidence for proofreading and potential therapeutics.* PLoS Pathog, 2013. **9**(8): p. e1003565.

64. Bouvet, M., et al., *RNA 3'-end mismatch excision by the severe acute respiratory syndrome coronavirus nonstructural protein nsp10/nsp14 exoribonuclease complex.* Proc Natl Acad Sci U S A, 2012. **109**(24): p. 9372-7.

65. Bhatt, P.R., et al., *Structural basis of ribosomal frameshifting during translation of the SARS-CoV-2 RNA genome.* Science, 2021. **372**(6548): p. 1306-1313.

66. Klein, S., et al., *SARS-CoV-2 structure and replication characterized by in situ cryo-electron tomography.* Nature Communications, 2020. **11**(1): p. 5885.

67. Tseng, Y.T., et al., *Identifying SARS-CoV membrane protein amino acid residues linked to virus-like particle assembly.* PLoS One, 2013. **8**(5): p. e64013.

68. Boson, B., et al., *The SARS-CoV-2 envelope and membrane proteins modulate maturation and retention of the spike protein, allowing assembly of virus-like particles.* J Biol Chem, 2021. **296**: p. 100111.

69. Fund, I.M., *World Economic Outlook:*

The Great Lockdown. 2020, Washington, DC.

70. Harrison, D., et al., *Impact of the SARS-CoV-2 pandemic on health-care workers.* Hospital Practice, 2020. **48**(4): p. 161-164.

71. Sender, R., et al., *The total number and mass of SARS-CoV-2 virions.* Proc Natl Acad Sci U S A, 2021. **118**(25).

72. Walls, A.C., et al., *Structure, Function, and Antigenicity of the SARS-CoV-2 Spike Glycoprotein.* Cell, 2020. **181**(2): p. 281-292.e6.

73. Wan, Y., et al., *Receptor Recognition by the Novel Coronavirus from Wuhan: an Analysis Based on Decade-Long Structural Studies of SARS Coronavirus.* J Virol, 2020. **94**(7).

74. Letko, M., A. Marzi, and V. Munster, *Functional assessment of cell entry and receptor usage for SARS-CoV-2 and other lineage B betacoronaviruses.* Nature Microbiology, 2020. **5**(4): p. 562-569.

75. Braun, M., et al., *ACE2 Co-evolutionary Pattern Suggests Targets for Pharmaceutical Intervention in the COVID-19 Pandemic.* iScience, 2020. **23**(8): p. 101384.

76. Fu, J., et al., *Expressions and significances of the angiotensin-converting enzyme 2 gene, the receptor of SARS-CoV-2 for COVID-19.* Molecular Biology Reports, 2020. **47**(6): p. 4383-4392.

77. Kam, Y.W., et al., *Cleavage of the SARS coronavirus spike glycoprotein by airway proteases enhances virus entry into human bronchial epithelial cells in vitro.* PLoS One, 2009. **4**(11): p. e7870.

78. Matsuyama, S., et al., *Protease-mediated enhancement of severe acute respiratory syndrome coronavirus infection.* Proc Natl Acad Sci U S A, 2005. **102**(35): p. 12543-7.

79. Lan, J., et al., *Structure of the SARS-CoV-2 spike receptor-binding domain bound to the ACE2 receptor.* Nature, 2020. **581**(7807): p. 215-220.

80. Xia, S., et al., *Peptide-Based Membrane Fusion Inhibitors Targeting HCoV-229E Spike Protein HR1 and HR2 Domains.* Int J Mol Sci, 2018. **19**(2).

81. Bertram, S., et al., *TMPRSS2 activates the human coronavirus 229E for cathepsin-independent host cell entry and is expressed in viral target cells in the respiratory epithelium.* J Virol, 2013. **87**(11): p. 6150-60.

82. Hoffmann, M., et al., *SARS-CoV-2 Cell Entry Depends on ACE2 and TMPRSS2 and Is Blocked by a Clinically Proven Protease Inhibitor.* Cell, 2020. **181**(2): p. 271-280.e8.

83. Bayati, A., et al., *SARS-CoV-2 infects cells after viral entry via clathrin-mediated endocytosis.* J Biol Chem, 2021. **296**: p. 100306.

84. Inoue, Y., et al., *Clathrin-dependent entry of severe acute respiratory syndrome coronavirus into target cells expressing ACE2 with the cytoplasmic tail deleted.* J Virol, 2007. **81**(16): p. 8722-9.

85. Harcourt, B.H., et al., *Identification of severe acute respiratory syndrome coronavirus replicase products and characterization of papain-like protease activity.* J Virol, 2004. **78**(24): p. 13600-12.

86. Ziebuhr, J., J. Herold, and S.G. Siddell, *Characterization of a human coronavirus (strain 229E) 3C-like proteinase activity.* 1995. **69**(7): p. 4331-4338.

87. Ziebuhr, J., V. Thiel, and A.E. Gorbalenya, *The Autocatalytic Release of a Putative RNA Virus Transcription Factor from Its Polyprotein Precursor Involves Two Paralogous Papain-like Proteases That Cleave the Same Peptide Bond *.* Journal of Biological Chemistry, 2001. **276**(35): p. 33220-33232.

88. Baranov, P.V., et al., *Programmed ribosomal frameshifting in decoding the SARS-CoV genome.* Virology, 2005. **332**(2): p. 498-510.

89. Su, M.C., et al., *An atypical RNA pseudoknot stimulator and an upstream attenuation signal for -1 ribosomal frameshifting of SARS coronavirus.* Nucleic Acids Res, 2005. **33**(13): p. 4265-75.

90. Wolff, G., et al., *A molecular pore spans the double membrane of the coronavirus replication organelle.* Science, 2020. **369**(6509): p. 1395-1398.

91. Knoops, K., et al., *SARS-coronavirus replication is supported by a reticulovesicular network of modified endoplasmic reticulum.* PLoS Biol, 2008. **6**(9): p. e226.

92. Zimmermann, L., et al., *SARS-CoV-2 nsp3-4 suffice to form a pore shaping replication organelles.* 2022: p. 2022.10.21.513196.

93. Snijder, E.J., et al., *A unifying structural and functional model of the coronavirus replication organelle: Tracking down RNA synthesis.* PLoS Biol, 2020. **18**(6): p. e3000715.

94. von Brunn, A., et al., *Analysis of intraviral protein-protein interactions of the SARS coronavirus ORFeome.* PLoS One, 2007. **2**(5): p. e459.

95. Wang, B., D. Svetlov, and I. Artsimovitch, *NMPylation and de-NMPylation of SARS-CoV-2 nsp9 by the NiRAN domain.* Nucleic Acids Res, 2021. **49**(15): p. 8822-8835.

96. Hillen, H.S., et al., *Structure of replicating SARS-CoV-2 polymerase.* Nature, 2020. **584**(7819): p. 154-156.

97. Kirchdoerfer, R.N. and A.B. Ward, *Structure of the SARS-CoV nsp12 polymerase bound to nsp7 and nsp8 co-factors.* Nat Commun, 2019. **10**(1): p. 2342.

98. Lehmann, K.C., et al., *Discovery of an essential nucleotidylating activity associated with a newly delineated conserved domain in the RNA polymerase-containing protein of all nidoviruses.* Nucleic Acids Res, 2015. **43**(17): p. 8416-34.

99. Slanina, H., et al., *Coronavirus replication–transcription complex: Vital and selective NMPylation of a conserved site in nsp9 by the NiRAN-RdRp subunit.* 2021. **118**(6): p. e2022310118.

100. Park, G.J., et al., *The mechanism of RNA capping by SARS-CoV-2.* Nature, 2022. **609**(7928): p. 793-800.

101. Yan, L., et al., *A mechanism for SARS-CoV-2 RNA capping and its inhibition by nucleotide analog inhibitors.* Cell, 2022. **185**(23): p. 4347-4360.e17.

102. Slanina, H., et al., *Conserved Characteristics of NMPylation Activities of Alpha- and Betacoronavirus NiRAN Domains.* 2023. **0**(0): p. e00465-23.

103. Lee, Y.F., et al., *A protein covalently linked to poliovirus genome RNA.* Proc Natl Acad Sci U S A, 1977. **74**(1): p. 59-63.

104. Nomoto, A., et al., *The location of the polio genome protein in viral RNAs and its implication for RNA synthesis.* Nature, 1977. **268**(5617): p. 208-213.

105. Paul, A.V., et al., *Identification of an RNA hairpin in poliovirus RNA that serves as the primary template in the in vitro uridylylation of VPg.* J Virol, 2000. **74**(22): p. 10359-70.

106. Imbert, I., et al., *A second, non-canonical RNA-dependent RNA polymerase in SARS coronavirus.* The EMBO journal, 2006. **25**(20): p. 4933-4942.

107. te Velthuis, A.J., S.H. van den Worm, and E.J. Snijder, *The SARS-coronavirus nsp7+nsp8 complex is a unique multimeric RNA polymerase capable of both de novo initiation and primer extension.* Nucleic Acids Res, 2012. **40**(4): p. 1737-47.

108. Zhai, Y., et al., *Insights into SARS-CoV transcription and replication from the structure of the nsp7–nsp8 hexadecamer.* Nature Structural & Molecular Biology, 2005. **12**(11): p. 980-986.

109. Konkolova, E., et al., *Structural analysis of the putative SARS-CoV-2 primase complex.* J Struct Biol, 2020. **211**(2): p. 107548.

110. Züst, R., et al., *Genetic interactions between an essential 3' cis-acting RNA pseudoknot, replicase gene products, and the extreme 3' end of the mouse coronavirus genome.* J Virol, 2008. **82**(3): p. 1214-28.

111. Egloff, M.-P., et al., *The severe acute respiratory syndrome-coronavirus replicative protein nsp9 is a single-stranded RNA-binding subunit unique in the RNA virus world.* 2004. **101**(11): p. 3792-3796.

112. Sutton, G., et al., *The nsp9 replicase protein of SARS-coronavirus, structure and functional insights.* Structure, 2004. **12**(2): p. 341-53.

113. Tagliabracci, V., et al., *The mechanism of RNA capping by SARS-CoV-2.* Res Sq, 2022.

114. Wang, D., et al., *The SARS-CoV-2 subgenome landscape and its novel regulatory features.* Molecular Cell, 2021.

115. Ugolini, C., et al., *Nanopore ReCappable sequencing maps SARS-CoV-2 5' capping sites and provides new insights into the structure of sgRNAs.* Nucleic Acids Res, 2022. **50**(6): p. 3475-3489.

116. Zhang, Y., et al., *In vivo structure and dynamics of the SARS-CoV-2 RNA genome.* Nat Commun, 2021. **12**(1): p. 5695.

117. Kim, D., et al., *The Architecture of SARS-CoV-2 Transcriptome.* Cell, 2020. **181**(4): p. 914-921.e10.

118. Sola, I., et al., *Continuous and Discontinuous RNA Synthesis in Coronaviruses.* Annu Rev Virol, 2015. **2**(1): p. 265-88.

119. Malone, B., et al., *Structures and functions of coronavirus replication-transcription complexes and their relevance for SARS-CoV-2 drug design.* Nat Rev Mol Cell Biol, 2022. **23**(1): p. 21-39.

120. Burgess, H.M., et al., *Targeting the m(6)A RNA modification pathway blocks SARS-CoV-2 and HCoV-OC43 replication.* Genes Dev, 2021. **35**(13-14): p. 1005-1019.

121. Finkel, Y., et al., *The coding capacity of SARS-CoV-2.* Nature, 2021. **589**(7840): p. 125-130.

122. Mateos-Gomez, P.A., et al., *Long-Distance RNA-RNA Interactions in the Coronavirus Genome Form High-Order Structures Promoting Discontinuous RNA Synthesis during Transcription.* 2013. **87**(1): p. 177-186.

123. Ma, Y., et al., *Structural basis and functional analysis of the SARS coronavirus nsp14-nsp10 complex.* Proc Natl Acad Sci U S A, 2015. **112**(30): p. 9436-41.

124. Aouadi, W., et al., *Binding of the Methyl Donor S-Adenosyl-l-Methionine to Middle East Respiratory Syndrome Coronavirus 2'-O-Methyltransferase nsp16 Promotes Recruitment of the Allosteric Activator nsp10.* J Virol, 2017. **91**(5).

125. Chen, Y., et al., *Biochemical and Structural Insights into the Mechanisms of SARS Coronavirus RNA Ribose 2'-O-Methylation by nsp16/nsp10 Protein Complex.* PLOS Pathogens, 2011. **7**(10): p. e1002294.

126. Walker, A.P., et al., *The SARS-CoV-2 RNA polymerase is a viral RNA capping enzyme.* Nucleic Acids Res, 2021. **49**(22): p. 13019-13030.

127. Wu, H.Y., et al., *Regulation of coronaviral poly(A) tail length during infection.* PLoS One, 2013. **8**(7): p. e70548.

128. Peng, Y.H., et al., *Characterization of the Role of Hexamer AGUAAA and Poly(A) Tail in Coronavirus Polyadenylation.* PLoS One, 2016. **11**(10): p. e0165077.

129. Gordon, D.E., et al., *A SARS-CoV-2 protein interaction map reveals targets for drug repurposing.* Nature, 2020. **583**(7816): p. 459-468.

130. Lee, S., et al., *The SARS-CoV-2 RNA interactome.* Mol Cell, 2021. **81**(13): p. 2838-2850.e6.

131. Bos, E.C., et al., *The production of recombinant infectious DI-particles of a murine coronavirus in the absence of helper virus.* Virology, 1996. **218**(1): p. 52-60.

132. Neuman, B.W., et al., *A structural analysis of M protein in coronavirus assembly and morphology.* Journal of Structural Biology, 2011. **174**(1): p. 11-22.

133. Xu, R., et al., *Construction of SARS-CoV-2 Virus-Like Particles by Mammalian Expression System.* 2020. **8**.

134. Athmer, J., et al., *Selective Packaging in Murine Coronavirus Promotes Virulence by Limiting Type I Interferon Responses.* mBio, 2018. **9**(3).

135. Baltz, Alexander G., et al., *The mRNA-Bound Proteome and Its Global Occupancy Profile on Protein-Coding Transcripts.* Molecular Cell, 2012. **46**(5): p. 674-690.

136. Liao, J.-Y., et al., *EuRBPDB: a comprehensive resource for annotation, functional and oncological investigation of eukaryotic RNA binding proteins (RBPs).* Nucleic Acids Research, 2019. **48**(D1): p. D307-D313.

137. Castello, A., et al., *Comprehensive Identification of RNA-Binding Domains in Human Cells.* Molecular Cell, 2016. **63**(4): p. 696-710.

138. Beckmann, B.M., et al., *The RNA-binding proteomes from yeast to man harbour conserved enigmRBPs.* Nat Commun, 2015. **6**: p. 10127.

139. Zeke, A., et al., *Deep structural insights into RNA-binding disordered protein regions.* Wiley Interdiscip Rev RNA, 2022. **13**(5): p. e1714.

140. Hentze, M.W., et al., *A brave new world of RNA-binding proteins.* Nature Reviews Molecular Cell Biology, 2018. **19**(5): p. 327-341.

141. Iselin, L., et al., *Uncovering viral RNA–host cell interactions on a proteome-wide scale.* Trends in Biochemical Sciences, 2022. **47**(1): p. 23-38.

142. Tanner, J.A., et al., *The severe acute respiratory syndrome (SARS) coronavirus NTPase/helicase belongs to a distinct class of 5' to 3' viral helicases.* J Biol Chem, 2003. **278**(41): p. 39578-82.

143. Urdaneta, E.C. and B.M. Beckmann, *Fast and unbiased purification of RNA-protein complexes after UV cross-linking.* Methods, 2020. **178**: p. 72-82.

144. Wagenmakers, A.J.M., R.J. Reinders, and W.J. Van Venrooij, *Cross-linking of mRNA to Proteins by Irradiation of Intact Cells with Ultraviolet Light.* 1980. **112**(2): p. 323-330.

145. Budowsky, E.I., et al., *Induction of polynucleotide-protein cross-linkages by ultraviolet irradiation.* 1986. **159**(1): p. 95-101.

146. Castello, A., et al., *Insights into RNA Biology from an Atlas of Mammalian mRNA-Binding Proteins.* Cell, 2012. **149**(6): p. 1393-1406.

147. Queiroz, R.M.L., et al., *Comprehensive identification of RNA-protein interactions in any organism using orthogonal organic phase separation (OOPS).* Nat Biotechnol, 2019. **37**(2): p. 169-178.

148. Garcia, J.A., et al., *Human immunodeficiency virus type 1 LTR TATA and TAR region sequences required for transcriptional regulation.* 1989. **8**(3): p. 765-778.

149. Plotch, S.J., et al., *A unique cap(m7GpppXm)-dependent influenza virion endonuclease cleaves capped RNAs to generate the primers that initiate viral RNA transcription.* Cell, 1981. **23**(3): p. 847-58.

150. Slobodin, B., et al., *Transcription Impacts the Efficiency of mRNA Translation via Co-transcriptional N6-adenosine Methylation.* Cell, 2017. **169**(2): p. 326-337.e12.

151. Zhang, X., et al., *An Atomic Structure of the Human Spliceosome.* Cell, 2017. **169**(5): p. 918-929.e14.

152. Bono, F., et al., *The Crystal Structure of the Exon Junction Complex Reveals How It Maintains a Stable Grip on mRNA.* Cell, 2006. **126**(4): p. 713-725.

153. Brownawell, A.M. and I.G. Macara, *Exportin-5, a novel karyopherin, mediates nuclear export of double-stranded RNA binding proteins.* J Cell Biol, 2002. **156**(1): p. 53-64.

154. Read, E.K. and P. Digard, *Individual influenza A virus mRNAs show differential dependence on cellular NXF1/TAP for their nuclear export.* J Gen Virol, 2010. **91**(Pt 5): p. 1290-301.

155. Amorim, M.J., et al., *Nuclear export of influenza A virus mRNAs requires ongoing RNA polymerase II activity.* Traffic, 2007. **8**(1): p. 1-11.

156. Winkler, R., et al., *m6A modification controls the innate immune response to infection by targeting type I interferons.* Nature Immunology, 2019. **20**(2): p. 173-182.

157. Zaccara, S. and S.R. Jaffrey, *A Unified Model for the Function of YTHDF Proteins in Regulating m6A-Modified mRNA.* Cell, 2020. **181**(7): p. 1582-1595.e18.

158. Fechter, P. and G.G. Brownlee, *Recognition of mRNA cap structures by viral and cellular proteins.* 2005. **86**(5): p. 1239-1249.

159. Mangus, D.A., M.C. Evans, and A. Jacobson, *Poly(A)-binding proteins: multifunctional scaffolds for the post-transcriptional control of gene expression.* Genome Biology, 2003. **4**(7): p. 223.

160. Ricardo-Lax, I., et al., *Replication and single-cycle delivery of SARS-CoV-2 replicons.* Science, 2021. **374**(6571): p. 1099-1106.

161. Jackson, R.J., C.U.T. Hellen, and T.V. Pestova, *The mechanism of eukaryotic translation initiation and principles of its regulation.* Nature Reviews Molecular Cell Biology, 2010. **11**(2): p. 113-127.

162. Bish, R., et al., *Comprehensive Protein Interactome Analysis of a Key RNA Helicase: Detection of Novel Stress Granule Proteins.* Biomolecules, 2015. **5**(3): p. 1441-66.

163. Kamenska, A., et al., *The DDX6-4E-T interaction mediates translational repression and P-body assembly.* Nucleic Acids Res, 2016. **44**(13): p. 6318-34.

164. Sheth, U. and R. Parker, *Decapping and Decay of Messenger RNA Occur in Cytoplasmic Processing Bodies.* Science, 2003. **300**(5620): p. 805-808.

165. Lejeune, F., X. Li, and L.E. Maquat, *Nonsense-Mediated mRNA Decay in Mammalian Cells Involves Decapping, Deadenylating, and Exonucleolytic Activities.* Molecular Cell, 2003. **12**(3): p. 675-687.

166. Hockensmith, J.W., et al., *Laser cross-linking of nucleic acids to proteins. Methodology and first applications to the phage T4 DNA replication system.* J Biol Chem, 1986. **261**(8): p. 3512-8.

167. Kamel, W., et al., *Global analysis of protein-RNA interactions in SARS-CoV-2-infected cells reveals key regulators of infection.* Molecular Cell, 2021. **81**(13): p. 2851-2867.e7.

168. Kim, B., et al., *Discovery of Widespread Host Protein Interactions with the Pre-replicated Genome of CHIKV Using VIR-CLASP.* Mol Cell, 2020. **78**(4): p. 624-640.e7.

169. Favre, A., et al., *Substitution of uridine in vivo by the intrinsic photoactivable probe 4-thiouridine in Escherichia coli RNA. Its use for E. coli ribosome structural analysis.* Eur J Biochem, 1986. **160**(3): p. 441-9.

170. Melvin, W.T., et al., *Incorporation of 6-thioguanosine and 4-thiouridine into RNA. Application to isolation of newly synthesised RNA by affinity chromatography.* Eur J Biochem, 1978. **92**(2): p. 373-9.

171. Lozzio, C.B. and P.W. Wigler, *Cytotoxic effects of thiopyrimidines.* J Cell Physiol, 1971. **78**(1): p. 25-32.

172. Herzog, V.A., et al., *Thiol-linked alkylation of RNA to assess expression dynamics.* Nat Methods, 2017. **14**(12): p. 1198-1204.

173. Jürges, C., L. Dölken, and F. Erhard, *Dissecting newly transcribed and old RNA using GRAND-SLAM.* Bioinformatics, 2018. **34**(13): p. i218-i226.

174. Nora Schmidt, S.G., Yuanjie Wei, Alexander Gabel, Sebastian Zielinski, Hasmik Keshishian, Caleb A. Lareau, Liv Zimmermann, Jana Makroczyova, Cadence Pearce, Karsten Krey, Thomas Hennig, Sebastian Stegmaier, Lambert Moyon, Marc Horlacher,Simone Werner, Jens Aydin, Marco Olguin-Nava, Ramya Potabattula, Anuja Kibe, Lars Dölken, Redmond P. Smyth, Neva Caliskan, Annalisa Marsico, Christine Krempl, Jochen Bodem, Andreas Pichlmair, Steven A. Carr, Petr Chlanda, Florian Erhard, and Mathias Munschauer, *The host protein SND1 binds SARS-CoV-2 negative-sense RNA and promotes viral RNA synthesis through NSP9.* in revision, 2023.

175. Trendel, J., et al., *The Human RNA-Binding Proteome and Its Dynamics during Translational Arrest.* Cell, 2019. **176**(1): p. 391-403.e19.

176. Urdaneta, E.C., et al., *Purification of cross-linked RNA-protein complexes by phenol-toluol extraction.* Nat Commun, 2019. **10**(1): p. 990.

177. McHugh, C.A. and M. Guttman, *RAP-MS: A Method to Identify Proteins that Interact Directly with a Specific RNA Molecule in Cells.* Methods in molecular biology (Clifton, N.J.), 2018. **1649**: p. 473-488.

178. Hafner, M., et al., *PAR-CliP--a method to identify transcriptome-wide the binding sites of RNA binding proteins.* J Vis Exp, 2010(41).

179. König, J., et al., *iCLIP reveals the function of hnRNP particles in splicing at individual nucleotide resolution.* Nat Struct Mol Biol, 2010. **17**(7): p. 909-15.

180. Van Nostrand, E.L., et al., *Robust transcriptome-wide discovery of RNA-binding protein binding sites with enhanced CLIP (eCLIP).* Nature methods, 2016. **13**(6): p. 508-514.

181. Lorenz, D.A., et al., *Multiplexed transcriptome discovery of RNA-binding protein binding sites by antibody-barcode eCLIP.* Nature Methods, 2023. **20**(1): p. 65-69.

182. Ghanbari, M. and U. Ohler, *Deep neural networks for interpreting RNA-binding protein target preferences.* Genome Res, 2020. **30**(2): p. 214-226.

183. Corley, M., et al., *Footprinting SHAPE-eCLIP Reveals Transcriptome-wide Hydrogen Bonds at RNA-Protein Interfaces.* Molecular Cell, 2020. **80**(5): p. 903-914.e8.

184. Chi, S.W., et al., *Argonaute HITS-CLIP decodes microRNA-mRNA interaction maps.* Nature, 2009. **460**(7254): p. 479-86.

185. Bailey, T.L., et al., *MEME SUITE: tools for motif discovery and searching.* Nucleic Acids Res, 2009. **37**(Web Server issue): p. W202-8.

186. Lahr, R.M., et al., *La-related protein 1 (LARP1) binds the mRNA cap, blocking eIF4F assembly on TOP mRNAs.* Elife, 2017. **6**.

187. Philippe, L., et al., *Global analysis of LARP1 translation targets reveals tunable and dynamic features of 5' TOP motifs.* Proc Natl Acad Sci U S A, 2020. **117**(10): p. 5319-5328.

188. Engreitz, J.M., et al., *The Xist lncRNA exploits three-dimensional genome architecture to spread across the X chromosome.* Science, 2013. **341**(6147): p. 1237973.

189. Lenarcic, E.M., et al., *Thiouracil cross-linking mass spectrometry: a cell-based method to identify host factors involved in viral amplification.* J Virol, 2013. **87**(15): p. 8697-712.

190. Kim, B., et al., *Viral crosslinking and solid-phase purification enables discovery of ribonucleoprotein complexes on incoming RNA virus genomes.* Nature Protocols, 2021. **16**(1): p. 516-531.

191. Chu, C., et al., *Systematic Discovery of Xist RNA Binding Proteins.* Cell, 2015. **161**(2): p. 404-416.

192. Chu, C. and H.Y. Chang, *ChIRP-MS: RNA-Directed Proteomic Discovery*, in *X-Chromosome Inactivation: Methods and Protocols*, T. Sado, Editor. 2018, Springer New York: New York, NY. p. 37-45.

193. LaPointe, A.T., et al., *Identification and characterization of Sindbis virus RNA-host protein interactions.* Journal of Virology, 2018. **92**(7).

194. Spiniello, M., et al., *HyPR-MS for Multiplexed Discovery of MALAT1, NEAT1, and NORAD lncRNA Protein Interactomes.* Journal of Proteome Research, 2018. **17**(9): p. 3022-3038.

195. Van Nostrand, E.L., et al., *Robust, Cost-Effective Profiling of RNA Binding Protein Targets with Single-end Enhanced Crosslinking and Immunoprecipitation (seCLIP).* Methods Mol Biol, 2017. **1648**: p. 177-200.

196. Kim, B. and V.N. Kim, *fCLIP-seq for transcriptomic footprinting of dsRNA-binding proteins: Lessons from DROSHA.* Methods, 2019. **152**: p. 3-11.

197. Licatalosi, D.D., et al., *HITS-CLIP yields genome-wide insights into brain alternative RNA processing.* Nature, 2008. **456**(7221): p. 464-9.

198. Benhalevy, D., D.G. Anastasakis, and M. Hafner, *Proximity-CLIP provides a snapshot of protein-occupied RNA elements in subcellular compartments.* Nature Methods, 2018. **15**(12): p. 1074-1082.

199. Lorenz, D.A., et al., *Multiplexed transcriptome discovery of RNA-binding protein binding sites by antibody-barcode eCLIP.* Nat Methods, 2023. **20**(1): p. 65-69.

200. McHugh, C.A., et al., *The Xist lncRNA interacts directly with SHARP to silence transcription through HDAC3.* Nature, 2015. **521**(7551): p. 232-236.

201. Chen, X., et al., *Practical considerations on performing and analyzing CLIP-seq experiments to identify transcriptomic-wide RNA-protein interactions.* Methods, 2019. **155**: p. 49-57.

202. Urlaub, H., K. Hartmuth, and R. Lührmann, *A two-tracked approach to analyze RNA–protein crosslinking sites in native, nonlabeled small nuclear ribonucleoprotein particles.* Methods, 2002. **26**(2): p. 170-181.

203. Wang, Y., M. Grunewald, and S. Perlman, *Coronaviruses: An Updated Overview of Their Replication and Pathogenesis.* Methods Mol Biol, 2020. **2203**: p. 1-29.

204. Tan, Y.W., et al., *Amino acid residues critical for RNA-binding in the N-terminal domain of the nucleocapsid protein are essential determinants for the infectivity of coronavirus in cultured cells.* Nucleic Acids Research, 2006. **34**(17): p. 4816-4825.

205. Kang, S., et al., *Crystal structure of SARS-CoV-2 nucleocapsid protein RNA binding domain reveals potential unique drug targeting sites.* Acta Pharm Sin B, 2020. **10**(7): p. 1228-1238.

206. Wang, Q., et al., *Structural Basis for RNA Replication by the SARS-CoV-2 Polymerase.* Cell, 2020. **182**(2): p. 417-428.e13.

207. Yin, W., et al., *Structural basis for inhibition of the RNA-dependent RNA polymerase from SARS-CoV-2 by remdesivir.* Science, 2020. **368**(6498): p. 1499-1504.

208. Shang, J., et al., *Structural basis of receptor recognition by SARS-CoV-2.* Nature, 2020. **581**(7807): p. 221-224.

209. Littler, D.R., et al., *A natural product compound inhibits coronaviral replication in vitro by binding to the conserved Nsp9 SARS-CoV-2 protein.* Journal of Biological Chemistry, 2021. **297**(6): p. 101362.

210. Labeau, A., et al., *Characterization and functional interrogation of the SARS-CoV-2 RNA interactome.* Cell Reports, 2022. **39**(4): p. 110744.

211. Benhalevy, D., et al., *The Human CCHC-type Zinc Finger Nucleic Acid-Binding Protein Binds G-Rich Elements in Target mRNA Coding Sequences and Promotes Translation.* Cell Reports, 2017. **18**(12): p. 2979-2990.

212. de Vries, T., et al., *Sequence-specific RNA recognition by an RGG motif connects U1 and U2 snRNP for spliceosome assembly.* 2022. **119**(6): p. e2114092119.

213. Chen, S., et al., *Mechanistic studies for the role of cellular nucleic-acid-binding protein (CNBP) in regulation of c-myc transcription.* Biochimica et Biophysica Acta (BBA) - General Subjects, 2013. **1830**(10): p. 4769-4777.

214. Lee, E., et al., *CNBP acts as a key transcriptional regulator of sustained expression of interleukin-6.* Nucleic Acids Res, 2017. **45**(6): p. 3280-3296.

215. Geuens, T., D. Bouhy, and V. Timmerman, *The hnRNP family: insights into their role in health and disease.* Human Genetics, 2016. **135**(8): p. 851-867.

216. Villarroya-Beltri, C., et al., *Sumoylated hnRNPA2B1 controls the sorting of miRNAs into exosomes through binding to specific motifs.* Nat Commun, 2013. **4**: p. 2980.

217. Munro, T.P., et al., *Mutational Analysis of a Heterogeneous Nuclear Ribonucleoprotein A2 Response Element for RNA Trafficking*.* Journal of Biological Chemistry, 1999. **274**(48): p. 34389-34395.

218. Wang, L., M. Wen, and X. Cao, *Nuclear hnRNPA2B1 initiates and amplifies the innate immune response to DNA viruses.* 2019. **365**(6454): p. eaav0758.

219. Wu, B., et al., *Molecular basis for the specific and multivariant recognitions of RNA substrates by human hnRNP A2/B1.* Nat Commun, 2018. **9**(1): p. 420.

220. Niikura, T., R. Hirata, and S.C. Weil, *A Novel Interferon-Inducible Gene Expressed during Myeloid Differentiation.* Blood Cells, Molecules, and Diseases, 1997. **23**(3): p. 337-349.

221. Abbas, Y.M., et al., *Structural basis for viral 5′-PPP-RNA recognition by human IFIT proteins.* Nature, 2013. **494**(7435): p. 60-64.

222. Katibah, G.E., et al., *tRNA binding, structure, and localization of the human interferon-induced protein IFIT5.* Mol Cell, 2013. **49**(4): p. 743-50.

223. Katibah, G.E., et al., *Broad and adaptable RNA structure recognition by the human interferon-induced tetratricopeptide repeat protein IFIT5.* Proc Natl Acad Sci U S A, 2014. **111**(33): p. 12025-30.

224. Feng, F., et al., *Crystal structure and nucleotide selectivity of human IFIT5/ISG58.* Cell Research, 2013. **23**(8): p. 1055-1058.

225. Yang, R., et al., *La-Related Protein 4 Binds Poly(A), Interacts with the Poly(A)-Binding Protein MLLE Domain via a Variant PAM2w Motif, and Can Promote mRNA Stability.* Molecular and Cellular Biology, 2011. **31**(3): p. 542-556.

226. Mattijssen, S., et al., *Single molecule poly(A) tail-seq shows LARP4 opposes deadenylation throughout mRNA lifespan with most impact on short tails.* Elife, 2020. **9**.

227. Liu, J., et al., *Calcineurin is a common target of cyclophilin-cyclosporin A and FKBP-FK506 complexes.* Cell, 1991. **66**(4): p. 807-815.

228. Liu, W., et al., *Cyclophilin A-regulated ubiquitination is critical for RIG-I-mediated antiviral immune responses.* Elife, 2017. **6**.

229. Braaten, D. and J. Luban, *Cyclophilin A regulates HIV-1 infectivity, as demonstrated by gene targeting in human T cells.* Embo j, 2001. **20**(6): p. 1300-9.

230. Liu, X., et al., *Cyclophilin A interacts with influenza A virus M1 protein and impairs the early stage of the viral replication.* 2009. **11**(5): p. 730-741.

231. Zhong, Z., et al., *Cytoplasmic CPSF6 Regulates HIV-1 Capsid Trafficking and Infection in a Cyclophilin A-Dependent Manner.* mBio, 2021. **12**(2).

232. Pavitrakar, D.V., et al., *Cyclophilin A: a possible host modulator in Chandipura virus infection.* Arch Virol, 2021. **166**(11): p. 3143-3150.

233. de Wilde, A.H., et al., *Cyclosporin A inhibits the replication of diverse coronaviruses.* J Gen Virol, 2011. **92**(Pt 11): p. 2542-2548.

234. de Wilde, A.H., et al., *MERS-coronavirus replication induces severe in vitro cytopathology and is strongly inhibited by cyclosporin A or interferon-α treatment.* 2013. **94**(8): p. 1749-1760.

235. Fenizia, C., et al., *Cyclosporine A Inhibits Viral Infection and Release as Well as Cytokine Production in Lung Cells by Three SARS-CoV-2 Variants.* Microbiol Spectr, 2022. **10**(1): p. e0150421.

236. Keedy, D.A., et al., *Mapping the conformational landscape of a dynamic enzyme by multitemperature and XFEL crystallography.* Elife, 2015. **4**.

237. Schmidt, N., et al., *The SARS-CoV-2 RNA-protein interactome in infected human cells.* Nat Microbiol, 2021. **6**(3): p. 339-353.

238. Popow, J., et al., *HSPC117 Is the Essential Subunit of a Human tRNA Splicing Ligase Complex.* 2011. **331**(6018): p. 760-764.

239. Kroupova, A., et al., *Molecular architecture of the human tRNA ligase complex.* Elife, 2021. **10**.

240. Keen, D.A. and A.L. Goodwin, *The crystallography of correlated disorder.* Nature, 2015. **521**(7552): p. 303-309.

241. Ponting, C.P., *P100, a transcriptional coactivator, is a human homologue of staphylococcal nuclease.* Protein Sci, 1997. **6**(2): p. 459-63.

242. Elbarbary, R.A., et al., *Tudor-SN-mediated endonucleolytic decay of human cell microRNAs promotes G(1)/S phase transition.* Science, 2017. **356**(6340): p. 859-862.

243. Yoo, B.K., et al., *Increased RNA-induced silencing complex (RISC) activity contributes to hepatocellular carcinoma.* Hepatology, 2011. **53**(5): p. 1538-48.

244. Scadden, A.D.J., *The RISC subunit Tudor-SN binds to hyper-edited double-stranded RNA and promotes its cleavage.* Nature Structural & Molecular Biology, 2005. **12**(6): p. 489-496.

245. Meister, G., et al., *Human Argonaute2 mediates RNA cleavage targeted by miRNAs and siRNAs.* Mol Cell, 2004. **15**(2): p. 185-97.

246. Rand, T.A., et al., *Biochemical identification of Argonaute 2 as the sole protein required for RNA-induced silencing complex activity.* 2004. **101**(40): p. 14385-14389.

247. Liu, J., et al., *Argonaute2 is the catalytic engine of mammalian RNAi.* Science, 2004. **305**(5689): p. 1437-41.

248. Friberg, A., et al., *Structure and Ligand Binding of the Extended Tudor Domain of D. melanogaster Tudor-SN.* Journal of Molecular Biology, 2009. **387**(4): p. 921-934.

249. Liu, H., et al., *Structural basis for methylarginine-dependent recognition of Aubergine by Tudor.* Genes Dev, 2010. **24**(17): p. 1876-81.

250. Liu, K., et al., *Structural basis for recognition of arginine methylated Piwi proteins by the extended Tudor domain.* Proc Natl Acad Sci U S A, 2010. **107**(43): p. 18398-403.

251. Tripsianes, K., et al., *Structural basis for dimethylarginine recognition by the Tudor domains of human SMN and SPF30 proteins.* Nature Structural & Molecular Biology, 2011. **18**(12): p. 1414-1420.

252. Courchaine, E.M., et al., *DMA-tudor interaction modules control the specificity of in vivo condensates.* Cell, 2021. **184**(14): p. 3612-3625.e17.

253. Tong, X., et al., *The Epstein-Barr virus nuclear protein 2 acidic domain forms a complex with a novel cellular coactivator that can interact with TFIIE.* Mol Cell Biol, 1995. **15**(9): p. 4735-44.

254. Paukku, K., J. Yang, and O. Silvennoinen, *Tudor and nuclease-like domains containing protein p100 function as coactivators for signal transducer and activator of transcription 5.* Mol Endocrinol, 2003. **17**(9): p. 1805-14.

255. Gao, X., et al., *Tudor Staphylococcal Nuclease (Tudor-SN) Participates in Small Ribonucleoprotein (snRNP) Assembly via Interacting with Symmetrically Dimethylated Sm Proteins*.* Journal of Biological Chemistry, 2012. **287**(22): p. 18130-18141.

256. Cappellari, M., et al., *The transcriptional co-activator SND1 is a novel regulator of alternative splicing in prostate cancer cells.* Oncogene, 2014. **33**(29): p. 3794-3802.

257. Yang, J., et al., *Transcriptional co-activator protein p100 interacts with snRNP proteins and facilitates the assembly of the spliceosome.* Nucleic Acids Res, 2007. **35**(13): p. 4485-94.

258. Weissbach, R. and A.D. Scadden, *Tudor-SN and ADAR1 are components of cytoplasmic stress granules.* Rna, 2012. **18**(3): p. 462-71.

259. Gutierrez-Beltran, E., et al., *Tudor staphylococcal nuclease: biochemistry and functions.* Cell Death & Differentiation, 2016. **23**(11): p. 1739-1748.

260. Caudy, A.A., et al., *A micrococcal nuclease homologue in RNAi effector complexes.* Nature, 2003. **425**(6956): p. 411-414.

261. Li, C.-L., et al., *Structural and functional insights into human Tudor-SN, a key component linking RNA interference and editing.* Nucleic acids research, 2008. **36**(11): p. 3579-3589.

262. Tsuchiya, N., et al., *SND1, a component of RNA-induced silencing complex, is up-regulated in human colon cancers and implicated in early stage colon carcinogenesis.* Cancer Res, 2007. **67**(19): p. 9568-76.

263. Blanco, M.A., et al., *Identification of Staphylococcal Nuclease Domain-containing 1 (SND1) as a Metadherin-interacting Protein with Metastasis-promoting Functions*.* Journal of Biological Chemistry, 2011. **286**(22): p. 19982-19992.

264. Lehto, M., et al., *Targeting of OSBP-related protein 3 (ORP3) to endoplasmic reticulum and plasma membrane is controlled by multiple determinants.* Experimental Cell Research, 2005. **310**(2): p. 445-462.

265. Murray, J., et al., *The subunit composition of the human NADH dehydrogenase obtained by rapid one-step immunopurification.* J Biol Chem, 2003. **278**(16): p. 13619-22.

266. Schindelin, J., et al., *Fiji: an open-source platform for biological-image analysis.* Nature Methods, 2012. **9**(7): p. 676-682.

267. Bolte, S. and F.P. Cordelières, *A guided tour into subcellular colocalization analysis in light microscopy.* 2006. **224**(3): p. 213-232.

268. Fredriksson, S., et al., *Protein detection using proximity-dependent DNA ligation assays.* Nature Biotechnology, 2002. **20**(5): p. 473-477.

269. López-Cano, M., V. Fernández-Dueñas, and F. Ciruela, *Proximity Ligation Assay Image Analysis Protocol: Addressing Receptor-Receptor Interactions*, in *Computer Optimized Microscopy: Methods and Protocols*, E. Rebollo and M. Bosch, Editors. 2019, Springer New York: New York, NY. p. 41-50.

270. Ule, J., et al., *CLIP identifies Nova-regulated RNA networks in the brain.* Science, 2003. **302**(5648): p. 1212-5.

271. Ule, J., et al., *CLIP: a method for identifying protein-RNA interaction sites in living cells.* Methods, 2005. **37**(4): p. 376-86.

272. Benjamini, Y. and D. Yekutieli, *The control of the false discovery rate in multiple testing under dependency.* 2001. **29 %J The Annals of Statistics**(4): p. 1165-1188, 24.

273. Zhang, Y., et al., *Model-based analysis of ChIP-Seq (MACS).* Genome Biol, 2008. **9**(9): p. R137.

274. Feng, J., et al., *Identifying ChIP-seq enrichment using MACS.* Nature Protocols, 2012. **7**(9): p. 1728-1740.

275. Hafner, M., et al., *Transcriptome-wide identification of RNA-binding protein and microRNA target sites by PAR-CLIP.* Cell, 2010. **141**(1): p. 129-41.

276. Meisenheimer, K.M. and T.H. Koch, *Photocross-Linking of Nucleic Acids to Associated Proteins.* Critical Reviews in Biochemistry and Molecular Biology, 1997. **32**(2): p. 101-140.

277. Ramanathan, M., D.F. Porter, and P.A. Khavari, *Methods to study RNA-protein interactions.* Nat Methods, 2019. **16**(3): p. 225-234.

278. Watkins, C.P. and R.A. Flynn, *A (cross)link in the chains.* Nature Chemistry, 2023. **15**(1): p. 5-6.

279. Knörlein, A., et al., *Nucleotide-amino acid π-stacking interactions initiate photo cross-linking in RNA-protein complexes.* Nature Communications, 2022. **13**(1): p. 2719.

280. Kramer, K., et al., *Photo-cross-linking and high-resolution mass spectrometry for assignment of RNA-binding sites in RNA-binding proteins.* Nature Methods, 2014. **11**(10): p. 1064-1070.

281. Munschauer, M., *High-Resolution Profiling of Protein-RNA Interactions*. 2015: Springer.

282. Vieira-Vieira, C.H. and M. Selbach, *Opportunities and Challenges in Global Quantification of RNA-Protein Interaction via UV Cross-Linking.* Front Mol Biosci, 2021. **8**: p. 669939.

283. Hafner, M., et al., *CLIP and complementary methods.* Nature Reviews Methods Primers, 2021. **1**(1): p. 20.

284. Banerjee, A.K., et al., *SARS-CoV-2 Disrupts Splicing, Translation, and Protein Trafficking to Suppress Host Defenses.* Cell, 2020. **183**(5): p. 1325-1339.e21.

285. Gerashchenko, M.V. and V.N. Gladyshev, *Ribonuclease selection for ribosome profiling.* Nucleic Acids Research, 2016. **45**(2): p. e6-e6.

286. Hashimoto, Y., et al., *Temporal dynamics of protein complex formation and dissociation during human cytomegalovirus infection.* Nature Communications, 2020. **11**(1): p. 806.

287. Krakau, S., H. Richard, and A. Marsico, *PureCLIP: capturing target-specific protein–RNA interaction footprints from single-nucleotide CLIP-seq data.* Genome Biology, 2017. **18**(1): p. 240.

288. Blackburn, J., et al., *Use of synthetic DNA spike-in controls (sequins) for human genome sequencing.* Nature Protocols, 2019. **14**(7): p. 2119-2151.

289. Embarc-Buh, A., R. Francisco-Velilla, and E. Martinez-Salas, *RNA-Binding Proteins at the Host-Pathogen Interface Targeting Viral Regulatory Elements.* Viruses, 2021. **13**(6).

290. Wu, H.Y. and D.A. Brian, *Subgenomic messenger RNA amplification in coronaviruses.* Proc Natl Acad Sci U S A, 2010. **107**(27): p. 12257-62.

291. Campagnola, G., et al., *The SARS-CoV nsp12 Polymerase Active Site Is Tuned for Large-Genome Replication.* 2022. **96**(16): p. e00671-22.

292. Hemmerich, P., et al., *Structural and functional properties of ribosomal protein L7 from humans and rodents.* Nucleic Acids Res, 1993. **21**(2): p. 223-31.

293. Perry, J.K., et al., *An atomistic model of the coronavirus replication-transcription complex as a hexamer assembled around nsp15.* J Biol Chem, 2021. **297**(4): p. 101218.

294. Zúñiga, S., et al., *Coronavirus Nucleocapsid Protein Facilitates Template Switching and Is Required for Efficient Transcription.* 2010. **84**(4): p. 2169-2175.

295. Stukalov, A., et al., *Multilevel proteomics reveals host perturbations by SARS-CoV-2 and SARS-CoV.* Nature, 2021. **594**(7862): p. 246-252.

296. Chen, Y., et al., *CNBP controls IL-12 gene transcription and Th1 immunity.* Journal of Experimental Medicine, 2018. **215**(12): p. 3136-3150.

297. Bezzi, G., et al., *CNBP Binds and Unfolds In Vitro G-Quadruplexes Formed in the SARS-CoV-2 Positive and Negative Genome Strands.* 2021. **22**(5): p. 2614.

298. Horlacher, M., et al., *A computational map of the human-SARS-CoV-2 protein-RNA interactome predicted at single-nucleotide resolution.* NAR Genom Bioinform, 2023. **5**(1): p. lqad010.

299. Hagihara, M., et al., *A reverse transcriptase stop assay revealed diverse quadruplex formations in UTRs in mRNA.* Bioorganic & Medicinal Chemistry Letters, 2010. **20**(7): p. 2350-2353.

300. Varshney, D., et al., *RNA G-quadruplex structures control ribosomal protein production.* Scientific Reports, 2021. **11**(1): p. 22735.

301. Ingolia, N.T., et al., *Genome-wide analysis in vivo of translation with nucleotide resolution using ribosome profiling.* Science, 2009. **324**(5924): p. 218-23.

302. Ingolia, N.T., L.F. Lareau, and J.S. Weissman, *Ribosome profiling of mouse embryonic stem cells reveals the complexity and dynamics of mammalian proteomes.* Cell, 2011. **147**(4): p. 789-802.

303. Scalabrin, M., et al., *The cellular protein hnRNP A2/B1 enhances HIV-1 transcription by unfolding LTR promoter G-quadruplexes.* Sci Rep, 2017. **7**: p. 45244.

304. Rodriguez, R., et al., *A novel small molecule that alters shelterin integrity and triggers a DNA-damage response at telomeres.* J Am Chem Soc, 2008. **130**(47): p. 15758-9.

305. Liu, G., et al., *RNA G-quadruplex in TMPRSS2 reduces SARS-CoV-2 infection.* Nature Communications, 2022. **13**(1): p. 1444.

306. Bao, H.-L. and Y. Xu, *Investigation of higher-order RNA G-quadruplex structures in vitro and in living cells by 19F NMR spectroscopy.* Nature Protocols, 2018. **13**(4): p. 652-665.

307. Yu, L., et al., *Fluorescent Visualization of Nucleolar G-Quadruplex RNA and Dynamics of Cytoplasm and Intranuclear Viscosity.* 2021. **3**(11): p. 2725-2739.

308. Pekarek, L., S. Buck, and N. Caliskan, *Optical Tweezers to Study RNA-Protein Interactions in Translation Regulation.* J Vis Exp, 2022(180).

309. Dobin, A., et al., *STAR: ultrafast universal RNA-seq aligner.* Bioinformatics, 2012. **29**(1): p. 15-21.

310. Piron, M., et al., *Rotavirus RNA-binding protein NSP3 interacts with eIF4GI and evicts the poly(A) binding protein from eIF4F.* Embo j, 1998. **17**(19): p. 5811-21.

311. Wells, S.E., et al., *Circularization of mRNA by eukaryotic translation initiation factors.* Mol Cell, 1998. **2**(1): p. 135-40.

312. Kim, B., et al., *Single-molecule visualization of mRNA circularization during translation.* Experimental & Molecular Medicine, 2023. **55**(2): p. 283-289.

313. Shaw, N., et al., *The multifunctional human p100 protein 'hooks' methylated ligands.* Nature Structural & Molecular Biology, 2007. **14**(8): p. 779-784.

314. Zheng, C., et al., *IFIT5 positively regulates NF-κB signaling through synergizing the recruitment of IκB kinase (IKK) to TGF-β-activated kinase 1 (TAK1).* Cell Signal, 2015. **27**(12): p. 2343-54.

315. Zhang, B., et al., *IFIT5 potentiates anti-viral response through enhancing innate immune signaling pathways.* 2013. **45**(10): p. 867-874.

316. Li, J., Y. Liu, and X. Zhang, *Murine coronavirus induces type I interferon in oligodendrocytes through recognition by RIG-I and MDA5.* J Virol, 2010. **84**(13): p. 6472-82.

317. Zhu, C., et al., *Cyclophilin A participates in the nuclear translocation of apoptosis-inducing factor in neurons after cerebral hypoxia-ischemia.* J Exp Med, 2007. **204**(8): p. 1741-8.

318. Pan, H., et al., *Cyclophilin A is required for CXCR4-mediated nuclear export of heterogeneous nuclear ribonucleoprotein A2, activation and nuclear translocation of ERK1/2, and chemotactic cell migration.* J Biol Chem, 2008. **283**(1): p. 623-637.

319. Colgan, J., et al., *Cyclophilin A regulates TCR signal strength in CD4+ T cells via a proline-directed conformational switch in Itk.* Immunity, 2004. **21**(2): p. 189-201.

320. Castro, A.P., et al., *Redistribution of cyclophilin A to viral factories during vaccinia virus infection and its incorporation into mature particles.* J Virol, 2003. **77**(16): p. 9052-68.

321. Shen, Z., et al., *Cyclosporin a inhibits rotavirus replication and restores interferon-beta signaling pathway in vitro and in vivo.* PLoS One, 2013. **8**(8): p. e71815.

322. Madan, V., et al., *Inhibition of HCV replication by cyclophilin antagonists is linked to replication fitness and occurs by inhibition of membranous web formation.* Gastroenterology, 2014. **146**(5): p. 1361-72.e1-9.

323. Kilka, S., et al., *The proline-rich N-terminal sequence of calcineurin Abeta determines substrate binding.* Biochemistry, 2009. **48**(9): p. 1900-10.

324. Castello, A., et al., *Identification of RNA-binding domains of RNA-binding proteins in cultured cells on a system-wide scale with RBDmap.* Nature Protocols, 2017. **12**(12): p. 2447-2464.

325. Greaney, A.J., et al., *Mapping mutations to the SARS-CoV-2 RBD that escape binding by different classes of antibodies.* Nature Communications, 2021. **12**(1): p. 4196.

326. Errico, J.M., L.J. Adams, and D.H. Fremont, *Antibody-mediated immunity to SARS-CoV-2 spike.* Adv Immunol, 2022. **154**: p. 1-69.

327. Kim, J., et al., *Targeting the Interaction Between Spike Protein and Nucleocapsid Protein for Suppression and Detection of Human Coronavirus OC43.* 2022. **13**.

328. Chakravarty, A.K., et al., *RNA ligase RtcB splices 3'-phosphate and 5'-OH ends via covalent RtcB-(histidinyl)-GMP and polynucleotide-(3')pp(5')G intermediates.* Proc Natl Acad Sci U S A, 2012. **109**(16): p. 6072-7.

329. Desai, K.K., et al., *Structures of the noncanonical RNA ligase RtcB reveal the mechanism of histidine guanylylation.* Biochemistry, 2013. **52**(15): p. 2518-25.

330. Ziv, O., et al., *The Short- and Long-Range RNA-RNA Interactome of SARS-CoV-2.* Mol Cell, 2020. **80**(6): p. 1067-1077.e5.

331. Wacker, A., et al., *Secondary structure determination of conserved SARS-CoV-2 RNA elements by NMR spectroscopy.* Nucleic Acids Res, 2020. **48**(22): p. 12415-12435.

332. Guo, F., et al., *Structural insights into the tumor-promoting function of the MTDH-SND1 complex.* Cell Rep, 2014. **8**(6): p. 1704-1713.

333. Rai, T.S., et al., *Histone chaperone HIRA deposits histone H3.3 onto foreign viral DNA and contributes to anti-viral intrinsic immunity.* Nucleic Acids Res, 2017. **45**(20): p. 11673-11683.

334. Sinz, A., *Cross-Linking/Mass Spectrometry for Studying Protein Structures and Protein–Protein Interactions: Where Are We Now and Where Should We Go from Here?* 2018. **57**(22): p. 6390-6396.

335. Hackbart, M., X. Deng, and S.C. Baker, *Coronavirus endoribonuclease targets viral polyuridine sequences to evade activating host sensors.* Proc Natl Acad Sci U S A, 2020. **117**(14): p. 8094-8103.

336. Guo, Y.R., et al., *The origin, transmission and clinical therapies on coronavirus disease 2019 (COVID-19) outbreak - an update on the status.* Mil Med Res, 2020. **7**(1): p. 11.

337. Hussain, S., et al., *Identification of novel subgenomic RNAs and noncanonical transcription initiation signals of severe acute respiratory syndrome coronavirus.* J Virol, 2005. **79**(9): p. 5288-95.

338. Sawicki, S.G., D.L. Sawicki, and S.G. Siddell, *A contemporary view of coronavirus transcription.* J Virol, 2007. **81**(1): p. 20-9.

339. Mahajan, M., et al., *NMR structure and localization of a large fragment of the SARS-CoV fusion protein: Implications in viral cell fusion.* Biochimica et Biophysica Acta (BBA) - Biomembranes, 2018. **1860**(2): p. 407-415.

340. Fehr, A.R. and S. Perlman, *Coronaviruses: an overview of their replication and pathogenesis.* Methods Mol Biol, 2015. **1282**: p. 1-23.

341. Martin, M.J.E.j., *Cutadapt removes adapter sequences from high-throughput sequencing reads.* 2011. **17**: p. 10-12.

342. Yee, B.A., et al., *RBP-Maps enables robust generation of splicing regulatory maps.* Rna, 2019. **25**(2): p. 193-204.

343. Van Nostrand, E.L., et al., *Principles of RNA processing from analysis of enhanced CLIP maps for 150 RNA binding proteins.* Genome Biol, 2020. **21**(1): p. 90.

344. Van Nostrand, E.L., et al., *A large-scale binding and functional map of human RNA-binding proteins.* Nature, 2020. **583**(7818): p. 711-719.

345. Robinson, J.T., et al., *Integrative genomics viewer.* Nature Biotechnology, 2011. **29**(1): p. 24-26.

346. Quinlan, A.R. and I.M. Hall, *BEDTools: a flexible suite of utilities for comparing genomic features.* Bioinformatics, 2010. **26**(6): p. 841-842.

347. Livak, K.J. and T.D. Schmittgen, *Analysis of Relative Gene Expression Data Using Real-Time Quantitative PCR and the 2−ΔΔCT Method.* Methods, 2001. **25**(4): p. 402-408.

348. Hughes, C.S., et al., *Ultrasensitive proteome analysis using paramagnetic bead technology.* Mol Syst Biol, 2014. **10**(10): p. 757.

349. Hughes, C.S., et al., *Single-pot, solid-phase-enhanced sample preparation for proteomics experiments.* Nature Protocols, 2019. **14**(1): p. 68-85.

350. Werner, T., et al., *Ion Coalescence of Neutron Encoded TMT 10-Plex Reporter Ions.* Analytical Chemistry, 2014. **86**(7): p. 3594-3601.

351. Dayon, L., et al., *Relative quantification of proteins in human cerebrospinal fluids by MS/MS using 6-plex isobaric tags.* Anal Chem, 2008. **80**(8): p. 2921-31.

352. Reichel, M., et al., *In Planta Determination of the mRNA-Binding Proteome of Arabidopsis Etiolated Seedlings.* Plant Cell, 2016. **28**(10): p. 2435-2452.

353. Franken, H., et al., *Thermal proteome profiling for unbiased identification of direct and indirect drug targets using multiplexed quantitative mass spectrometry.* Nat Protoc, 2015. **10**(10): p. 1567-93.

354. Ritchie, M.E., et al., *limma powers differential expression analyses for RNA-sequencing and microarray studies.* Nucleic Acids Res, 2015. **43**(7): p. e47.

355. Huber, W., et al., *Variance stabilization applied to microarray data calibration and to the quantification of differential expression.* Bioinformatics, 2002. **18 Suppl 1**: p. S96-104.

356. Mi, H., et al., *Large-scale gene function analysis with the PANTHER classification system.* Nat Protoc, 2013. **8**(8): p. 1551-66.

Acknowledgments

First and foremost, I want to thank Jun. Prof. Mathias Munschauer to choose me as one of his first PhD students. It was an insightful and educational experience being in your lab and I value the time and effort you have put to support me in my journey. Thank you for your trust and your guidance.

The person who supported me beyond anything I could imagine is Dr. Nora Schmidt. You are a role model in our lab and the most welcoming and supportive person I have met in my time here. I appreciate all of your effort and time in the lab and the discussions . Thank you for reviewing my thesis and the mental support you brought to me.

Next, I want to thank Yuanjie Wei and Dr. Alexander Gabel for the bioinformatic analysis and scientific discussions. Without you the true potential of my experimental data could not have been revealed as they are now.

Also, I want to thank our whole LRIB team members. Colleagues that have become friends and family. Work would not be as fun without supportive and cheerful people around. Thanks for the laughter and joy you brought to me, especially in the frustrating and though times. I am also deeply grateful to my gossip girls, besties and my sporting ace that made each day end with a smile.

Another thanks goes to my second left hand in the literal sense. Without you, Sebastian Stegmaier my project would have come to a halt and valuable time could be saved thanks to your support.

Special thanks have the be addressed to my first scientific supervisor Dr. Burkhardt Laufer. You have built the fundament of my confidence and endurance in pursuing the next 10 years of my life. I can´t thank you enough for my short but sweet time under your guidance. I really appreciate that you decided to take me as an untrained schoolgirl and gave me the toolkit to become a woman in science.

Zu guter Letzt meine Familie und meinem Liebsten. Die Unterstützung die ich durch euch während meines gesamten Studiums erhalten habe werde ich euch nie vergessen. Danke, dass ich mich an so vielen Orten gleichzeitig zu Hause fühlen konnte. David, die Kraft die du mir gegeben hast ist unendlich. Ich freue mich auf die Zukunft, wissend, dass ihr alle für mich da sein werdet.

CURRICULUM VITAE

Name: Sabina Ganskih; Ms

Adress: Auf der Schanz 58, 97076 Würzburg

Mobile: 0151 - 41652866

E- mail: ganskih@mailbox.org

Date of Birth: 06/12/1995

Education and internships

Oct. 2019 – ongoing Graduate School of Life Science

Dr. rer. nat.: Mathias Munschauer

May 2018 – Aug. 2018 internship at **Abbvie** Deutschland GmbH & Co. KG R&D Dr. Knut Biber

Sep. 2016 – July 2019 **Ruprecht- Karls University Heidelberg**

Master Molecular Biology- Infectious Diseases; M.Sc. Molecular Biosciences: Prof. Dr. Yonne Samstag

Sep. 2013 – Jul. 2016 **Ruprecht- Karls University Heidelberg**

B.Sc Bioscience: Prof. Oliver Fackler

Jul. 2013 – Aug. 2013 internship at **Ascendis pharma**

Employment

Oct. 2019 – ongoing	Helmholtz Institute for RNA- based Infection Research
AG Mathias Munschauer	
April 2018	research assistant for **Zendia** GmbH R&D Guido Boehse and AG Markus Ganter
Oct. 2016 – Nov. 2016	research assistant for the **Department of Medicine**; Tracking of HIV (un-)infected cells for research; AG Oliver Fackler
Aug. 2016 – Dec. 2017	research assistant for the **Department of Medicine**; Genotyping & phenotyping of mouse strains; AG Oliver Fackler
Oct. 2015 **–** Dec. 2015	research assistant for the **Department of Medicine**; Preparation of medical course; AG Oliver Fackler
Oct. 2015	research assistant for the **Department of Bioscience** Tutor for elementary students

Research techniques

- Experience in flow cytometrical experiments
- Experience in cell sorting
- Experience in immunofluorescence imaging
- Strong background in methods for RNA research
- Strong background in methods for protein research
- Strong background in methods for virus research
- Strong knowledge in Inkscape
- Intermediate knowledge in imaging analysis software Fiji and CellProfiler
- Basic knowledge in RStudio

Languages

- German, native tongue
- English, fluently
- Russion, fluently
- Spanish, pre-intermediate
- Latinum
- Graecum

Interests and additional skills

Jan. 2020 – Dec. 2021	Chair-Person and Main organizer of the GSLS Symposium Eureka!2020 and Eureka!2021
Jun. 2015 – Oct. 2016	Elected member of the student representatives for the **Department of Bioscience**
Sep. 2013 – Jul. 2017	Active member of the student's association
Nov. 2012 – Jul. 2014	Coach at the martial arts school **Black Scorpions Mannheim**
Sept. 2011 – Jul. 2012	Committee of the School Government (SMV) **Karl- Friedrich- Gymnasium**

Sports

During my childhood, I was very interested in sports. I´ve been member of a volleyball club for 8 years and from the age of 14 team taken the tasks of team captain and children trainer. Additionally, I enjoyed standard dancing courses for one and a half years. Later on, I decided to switch to martial arts for self- defense and as efficient workout, until I moved to Heidelberg for my studies. In our martial arts club, I have coached children at the age of 8-16 years for a year and occasionally the adults for about 2 years.

During my master program I went back to dancing where I know regularly attend on international workshops and getting to know many people from all over the world.

List of Publications

Published

Gonzalez-Perez, A.C., et al., *The Zinc Finger Antiviral Protein ZAP Restricts Human Cytomegalovirus and Selectively Binds and Destabilizes Viral UL4/UL5 Transcripts.* mBio, 2021. **12**(3).

Schmidt, N., et al., *The SARS-CoV-2 RNA-protein interactome in infected human cells.* Nat Microbiol, 2021. **6**(3): p. 339-353.

Orlik, C., et al., *Keratinocytes costimulate naive human T cells via CD2: a potential target to prevent the development of proinflammatory Th1 cells in the skin.* Cell Mol Immunol, 2020. **17**(4): p. 380-394.

In preparation

Schmidt, N., Ganskih, S., Wei, Y., Gabel, A., et al. *The host protein SND1 binds SARS-CoV-2 negative-sense RNA and promotes viral RNA synthesis through NSP9*

Trifault et al., *Nucleolar detention of NONO shields DNA double-strand breaks form aberrant transcripts*

Aydin et al., *SHIFTR enables the unbiased identification of proteins bound to specific RNA regions in transcriptomes of virus and host*

Appendix

5.1. List of Figures

5.2. List of Tables

5.3. Supplementary Table 1

Table 12 Go Term analysis SND1 protein interactors in naive cells

Analysis Type:	PANTHER Overrepresentation Test (Released 20230705)						
Annotation Version and Release Date:	GO Ontology database DOI: 10.5281/zenodo.7942786 Released 2023-05-10						
Analyzed List:	upload_1 (Homo sapiens)						
Reference List:	Homo sapiens (all genes in database)						
Test Type:	FISHER						
Correction:	FDR						
GO biological process complete	Homo sapiens - REFLIST (20589)	upload_1 (2	upload_1 (expected)	upload_1 (over/und	upload_1 (fold Enrichmen(	upload_1 (raw P-value)	upload_1 (FDR)
negative regulation of mRNA splicing, via spliceosome (GO:0048025)	23	3	0.02	+	> 100	2.70E-06	1.61E-03
negative regulation of mRNA processing (GO:0050686)	25	3	0.03	+	> 100	3.40E-06	1.95E-03
negative regulation of RNA splicing (GO:0033119)	28	3	0.03	+	> 100	4.65E-06	2.49E-03
regulation of alternative mRNA splicing, via spliceosome (GO:0000381)	54	4	0.06	+	69.32	3.98E-07	2.68E-04
positive regulation of RNA splicing (GO:0033120)	45	3	0.05	+	62.39	1.77E-05	8.59E-03
regulation of mRNA splicing, via spliceosome (GO:0048024)	106	7	0.11	+	61.8	1.97E-11	5.09E-08
spliceosomal complex assembly (GO:0000245)	73	4	0.08	+	51.28	1.25E-06	8.10E-04
regulation of mRNA processing (GO:0050684)	134	7	0.14	+	48.89	9.45E-11	1.47E-07
regulation of RNA splicing (GO:0043484)	183	8	0.2	+	40.91	1.34E-11	5.22E-08
negative regulation of mRNA metabolic process (GO:1903312)	96	4	0.1	+	38.99	3.57E-06	1.98E-03
cytoplasmic translation (GO:0002181)	123	4	0.13	+	30.43	9.23E-06	4.62E-03
mRNA splicing, via spliceosome (GO:0000398)	249	7	0.27	+	26.31	6.11E-09	5.58E-06
RNA splicing, via transesterification reactions with bulged adenosine as nucleophile (GO:0000377)	249	7	0.27	+	26.31	6.11E-09	5.27E-06
RNA splicing, via transesterification reactions (GO:0000375)	253	7	0.27	+	25.89	6.80E-09	5.56E-06
RNA splicing (GO:0008380)	381	10	0.41	+	24.56	2.84E-12	1.47E-08
regulation of mRNA metabolic process (GO:1903311)	300	7	0.32	+	21.84	2.14E-08	1.66E-05
mRNA processing (GO:0006397)	459	10	0.49	+	20.39	1.71E-11	5.32E-08
protein-RNA complex assembly (GO:0022618)	199	4	0.21	+	18.81	5.81E-05	2.51E-02
ribonucleoprotein complex biogenesis (GO:0022613)	457	9	0.49	+	18.43	5.48E-10	6.54E-07
protein-RNA complex organization (GO:0071826)	207	4	0.22	+	18.08	6.76E-05	2.84E-02
rRNA processing (GO:0006364)	218	4	0.23	+	17.17	8.23E-05	3.28E-02
mRNA metabolic process (GO:0016071)	622	10	0.66	+	15.05	3.18E-10	4.48E-07
ribosome biogenesis (GO:0042254)	312	5	0.33	+	15	1.77E-05	8.34E-03
RNA processing (GO:0006396)	882	14	0.94	+	14.85	1.80E-14	2.80E-10
gene expression (GO:0010467)	1827	16	1.95	+	8.2	6.96E-13	5.41E-09
RNA metabolic process (GO:0016070)	1662	14	1.78	+	7.88	8.99E-11	1.55E-07
nucleic acid metabolic process (GO:0090304)	2281	16	2.44	+	6.56	2.08E-11	4.62E-08
nucleobase-containing compound metabolic process (GO:0006139)	2796	16	2.99	+	5.36	4.57E-10	5.91E-07
heterocycle metabolic process (GO:0046483)	2972	16	3.18	+	5.04	1.14E-09	1.27E-06
cellular nitrogen compound metabolic process (GO:0034641)	3347	18	3.58	+	5.03	2.45E-11	4.76E-08
cellular aromatic compound metabolic process (GO:0006725)	3016	16	3.22	+	4.96	1.43E-09	1.48E-06
macromolecule biosynthetic process (GO:0009059)	1526	8	1.63	+	4.91	1.15E-04	4.35E-02
organic cyclic compound metabolic process (GO:1901360)	3258	16	3.48	+	4.6	4.53E-09	4.39E-06
cellular component biogenesis (GO:0044085)	2666	11	2.85	+	3.86	3.09E-05	1.37E-02
cellular metabolic process (GO:0044237)	5828	18	6.23	+	2.89	2.91E-07	2.15E-04
macromolecule metabolic process (GO:0043170)	5887	18	6.29	+	2.86	3.44E-07	2.43E-04
nitrogen compound metabolic process (GO:0006807)	6614	18	7.07	+	2.55	2.31E-06	1.44E-03
cellular component organization or biogenesis (GO:0071840)	5774	15	6.17	+	2.43	1.08E-04	4.17E-02
primary metabolic process (GO:0044238)	7165	18	7.66	+	2.35	8.41E-06	4.35E-03
organic substance metabolic process (GO:0071704)	7653	18	8.18	+	2.2	2.40E-05	1.10E-02
metabolic process (GO:0008152)	8086	18	8.64	+	2.08	7.38E-05	3.02E-02

5.4. Abbreviations

4SU	4-Thiouridine
5´ TOP	5´ terminal oligopyrimidine
5´ppp	5′ triphosphate
A	Adenine
ABC	Antibody barcode CLIP
ACE2	Angiotensin-converting enzyme 2
AGO	Argonaute
anti 3´ UTR	Antisense sequence of 3´ UTR
anti TRS	Antisense region of transcription regulatory sequence
ASN	Asparagine
Asp	Asparagine
BCoV	Bovine coronavirus
C	Cytosine
Cas9	CRISPR associated protein 9
ChIRP-MS	Comprehensive identification of RNA-binding proteins by mass spectrometry
CLAMP	Coss-link assisted mRNP purification
CLIP	Cross-linking immunoprecipitation
CNBP	CCHC-Type Zinc Finger Nucleic Acid Binding Protein
COVID-19	Coronavirus disease 2019
CRISPR	Clustered regularly interspaced short palindromic repeats
CsA	Cyclosporin A
DMA	Dimethylarginine
DMV	Double membrane vesicle
DNA	Deoxyribonucleic acid
ds	Double stranded
dsRNA	Double stranded RNA

E	Envelope protein
eCLIP	Enhanced cross-linking IP
eIF4B	Translation initiation factor 4B
ER	Endoplasmic reticulum
ERGIC	Endoplasmic reticulum-intermediate compartment
FC	Fold change
fCLIP	Formaldehyde cross-linking IP
FDR	False discovery rate
FP	Fusion peptide
G	Guanine
GO	Gene ontology
GpppA	(5')ppp(5')A RNA Cap
gRNA	Genomic RNA
h	Hour
hCMV	Human cytomegalovirus
hCoV	Human Coronavirus
hCoV-229E	Human coronavirus 229E
HITS-CLIP	High-throughput sequencing of RNA isolated by cross-linking IP
HIV	Human immunodeficiency virus-1
hpi	Hours post infection
HR1/ HR2	Heptapeptide repeat sequence 1/ 2
HyPR-MS	Hybridization purification of RNA-protein complexes followed by mass spectrometry
iCLIP	Individual-nucleotide resolution cross-linking and IP
ICTV	International Committee on Taxonomy of Viruses
IDR	intrinsically disordered region
IF	Immunofluorescence
IFIT5	Interferon-induced protein with tetratricopeptide repeats 5

IgG	Immunoglobulin G
IP	Immunoprecipitation
IRES	internal ribosome entry site
ISG	Interferon-stimulated gene
LaM	La-Motif
LARP4	La-related protein 4
LC-MS/MS	Liquid chromatography mass spectrometry
M	Membrane protein
µs	Microsecond
m^6A	N^6-methyladenosine
m6A	N6-methyladenosine
$^{me7}G_0pppA_{1m}$	N7- and 2′-O-methylated cap; Cap-1 structure
MEME	Multiple Expectation maximizations for Motif Elicitation
MERS-CoV	Middle East respiratory syndrome-related coronavirus
MHV	Murine hepatitis virus
miRNA	Micro RNA
MOI	Multiplicity of infection
mRNA	Messenger RNA
MTDH	Metadherin
N	Nucleocapsid
NDUFA12	NADH dehydrogenase 1 alpha subcomplex subunit 12
neg	Negative
NF-κB	Nuclear factor kappa-light-chain-enhancer of activated B cells
NiRAN	Nidovirus RNA-dependent RNA polymerase-associated nucleotidyltransferase
nm	Nanometer
NMP	Nucleoside monophosphate
nsp	Non-structural protein

NTP	Nucleoside triphosphate
ORF	Open reading frame
OSBPL3	Oxysterol-binding protein-related protein 3
P body	Processing body
PAMP	Pathogen-associated microbial pattern
PAR-CLIP	Photoactivatable ribonucleoside-enhanced cross-linking and IP
PDS	Pyridostatin
PFU	Plaque forming units
PLA	Proximity ligation assay
PL^{pro}	Papain-like proteases
PMT	Posttranscriptional modification
poly(A)	Polyadenylated
pos	Positive
pp1a/pp1ab	Polyprotein 1a / Polyprotein 1ab
PPIA	Peptidylprolyl isomerase A
proximity-CLIP	Proximity cross-linking IP
PRR	Pattern recognition receptor
RAP-MS	RNA antisense purification coupled with mass spectrometry
RBD	Recognition binding domain
RBDmap	RNA binding domain mapping
RBP	RNA-binding protein
RdRp	RNA-dependent RNA polymerase
RGG	Arginine-Glycine-Glycine
RISC	RNA-induced silencing complex
RISC	RNA-induced silencing complex
RNA	Ribonucleic acid
RNP	Ribonucleoprotein
RPC	RNA-protein complex

RPL7	Ribosomal protein L7
RPS2	40S ribosomal protein S2
RPS8	40S ribosomal protein S8
RRM	RNA recognition motif
RT qPCR	Quantitative reverse transcription polymerase chain reaction
RTC	Replication transcription complex
RTCB	RNA 2',3'-cyclic phosphate and 5'-OH ligase
S	Spike
SARS	Severe acute respiratory syndrome
SARS-CoV-2	Severe acute respiratory syndrome coronavirus 2
SARSr-CoV	Severe acute respiratory syndrome related coronavirus
SARSr-CoV	Severe acute respiratory syndrome-related coronavirus
SDS-PAGE	Sodium dodecyl sulfate polyacrylamide gel electrophoresis
sgRNA	Sub-genomic RNA
SHAPE	2′-hydroxyl acylation analyzed by primer extension
SMI	Size-matched input
SND1	Staphylococcal nuclease domain-containing protein 1
SND1	Staphylococcal Nuclease and Tudor Domain Containing 1
ss	Single stranded
ssRNA	Single stranded RNA
T	Thymine
THRAP3	Thyroid hormone receptor-associated protein 3
TMPRSS2	Transmembrane protease serine 2
TMT	Tandem mass tag
TOP3B	DNA topoisomerase 3-beta-1
TPR	tetratricopeptide repeat
TRA2B	Transformer-2 protein homolog beta
TRS	Transcription regulatory sequence

TRS-B	Transcription regulatory sequence-Body
TRS-L	Transcription regulatory sequence-Leader
TUX-MS	Thiouracil cross(X)-linking mass spectrometry
U	Uracil
UTR	Untranslated region
UV	Ultraviolet
VIR-CLASP	Viral cross-linking and solid-phase purification
VLP	Virus-like particle
VPg	Viral protein genome-linked
vRIC	Viral RNA interactome capture
WHO	World Health Organization
YTHDF1-3	YTH domain-containing family proteins
ZAP	Zinc finger CCCH-type antiviral protein 1
ZC3HAV1	Zinc finger CCCH-type antiviral protein 1
ZnF	Zinc finger