Utjecaj proteina Nsp1 virusa SARS-CoV-2 na translaciju u ljudskim stanicama

Bujanić, Lucija

University of Zagreb
Faculty of Science
Department of Biology

Lucija Bujanić

The effects of SARS-CoV-2 protein Nsp1 on translation in human cells

Master's thesis

Zagreb, 2021.

Sveučilište u Zagrebu
Prirodoslovno-matematički fakultet
Biološki odsjek

Lucija Bujanić

Utjecaj proteina Nsp1 virusa SARS-CoV-2 na translaciju u ljudskim stanicama

Diplomski rad

Zagreb, 2021.

The research for this thesis was generated in the Laboratory for non-coding RNAs and mechanisms of posttranscriptional gene regulation at the Berlin Institute for Medical Systems Biology (part of Max-Delbrück Center for Molecular Medicine in the Helmholtz Association) under the supervision of dr. sc. Marina Chekulaeva, and co-supervised by dr. sc. Inga Urlić. The thesis was submitted for review to the Department of Biology at the Faculty of Science of University of Zagreb in order to obtain the title of Master in Molecular Biology (mag. biol. mol.).

Acknowledgments

First of all, I would like to thank my supervisor, Marina Chekulaeva, for giving me the opportunity to pursue my master's thesis under her mentorship and for allowing me to participate in this project from its beginning. I would also like to thank my co-supervisor, dr. sc. Inga Urlić, for enthusiastically accepting this role and being supportive throughout this whole process.

I would like to thank Samantha for her endless moral support, advice and all the tips and tricks she taught me. A big thank you to Erik, Sayaka and Kasia for patiently answering every question I had and selflessly sharing your knowledge and protocols. And a big thank you to each of them for the amazing company they offered during my stay in their lab. Thank you to David, Artem, Nicolai, Inga and Seva for all the helpful discussions, advice and their input over Zoom meetings.

Next, I would like to thank my friends - Ema, Josipa and all the Karlas for having supported me without exception in all my endeavors and continuing to do so. Thank you to Iris and Lucija for their compassion and understanding throughout this journey, it was a pleasure to share it with you.

Thank you to Iheb for all of his daily pep talks whenever they were needed. Thank you for your unwavering support, love and comfort regardless of where you are.

Most of all, I would like to thank my family who have supported me throughout my education in more aspects than I can count. Thank you for constantly reminding me of just how much you love and believe in me.

TEMELJNA DOKUMENTACIJSKA KARTICA

Sveučilište u Zagrebu
Prirodoslovno-matematički fakultet
Biološki odsjek Diplomski rad

Utjecaj proteina Nsp1 virusa SARS-CoV-2 na translaciju u ljudskim stanicama

Lucija Bujanić

Rooseveltov trg 6, 10000 Zagreb, Hrvatska

Severe acture respiratory syndrome coronavirus 2 (SARS-CoV-2) je koronavirus koji uzrokuje teške respiratorne probleme kod ljudske populacije i uzrok je trenutne pandemije. Zbog toga postoji potreba za istraživanjem mehanizma infekcije virusa kako bi se mogle utvrditi potencijalne mete terapeutika. U ovom radu istražena je uloga proteina Nsp1 virusa SARS-CoV-2 korištenjem dvostrukog sustava proteina reportera luciferaze. Protein Nsp1 je jedan od proteina koji su eksprimirani u najranijim fazama infekcije virusom te se pretpostavlja kako ima ulogu u inhibiciji translacije. Ovaj rad pokazuje da protein Nsp1 ima sposobnost selektivnog odabira virusnih transkripata od staničnih te samo za njih omogućuje translaciju, dok za stanične transkripte istu inhibira. Također, pokazano je da je za proces prepoznavanja bitna uloga 5' vodeće virusne sekvence i da su aminokiseline na mjestima RK124 i KH164 esencijalne za funkcionalnost proteina. Nadalje, ovaj rad pokazuje da postoji očuvanje funkcije između proteina Nsp1 virusa SARS-CoV-2 i homolognog proteina virusa SARS-CoV-1 (*severe acture respiratory syndrome 1*). Zajedno ovi rezultati ukazuju na važnost proteina Nsp1 u virusnoj infekciji i naglašavaju ulogu u progresiji bolesti.

(32 stranice, 11 slika, 2 tablice, 33 literaturna navoda, jezik izvornika: engleski)
Rad je pohranjen u Središnjoj biološkoj knjižnici
Ključne riječi: koronavirus, 5' vodeća sekvenca, metoda dvostrukog reportera, luciferaza
Voditelj: dr. sc. Marina Chekulaeva
Suvoditelj: izv. prof. dr. sc. Inga Urlić
Ocjenitelji: izv. prof. dr. sc. Inga Urlić
izv. prof. dr. sc. Petra Korać
izv. prof. dr. sc. Petra Peharec Štefanić

Rad prihvaćen: 7. srpnja 2021.

BASIC DOCUMENTATION CARD

University of Zagreb
Faculty of Science
Department of Biology

Master's Thesis

The effects of SARS-CoV-2 protein Nsp1 on translation in human cells

Lucija Bujanić

Rooseveltov trg 6, 10000 Zagreb, Hrvatska

Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) is a coronavirus that causes severe respiratory issues in the human population and is the reason behind the ongoing pandemic. Because of that, there is a worldwide need to decipher its mechanisms of infection in order to be able to determine potential therapeutics targets. In this thesis, the role of the SARS-CoV-2 protein Nsp1 is explored using a dual luciferase reporter assay approach. Nsp1 is one of the earliest viral proteins expressed upon infection and it is presumed to have a role in inhibiting translation. This thesis shows that Nsp1 selects and allows translation of viral transcripts and blocks translation of cellular transcripts. Another finding is that the 5' viral leader sequence is essential for this recognition process and that the amino acid sites RK124 and KH164 are essential for the protein's functionality. In addition, it is shown that there is conservation of function between SARS-CoV-2 Nsp1 and its homolog from SARS-CoV-1 (severe acture respiratory syndrome coronavirus 1). Together, the results indicate the importance of Nsp1 in viral infection and emphasise its role in disease progression.

(32 pages, 11 figures, 2 tables, 33 references, original in: English)
Thesis is deposited in Central Biological Library
Keywords: coronavirus, 5' leader sequence, reporter assay, luciferase
Supervisor: dr. sc. Marina Chekulaeva
Co-supervisor: dr. sc. Inga Urlić, assoc. prof.
Reviewers: dr. sc. Inga Urlić, assoc. prof.
dr. sc. Petra Korać, assoc. prof.
dr. sc. Petra Peharec Štefanić, assoc. prof.
Thesis accepted: 7th July 2021.

Table of contents

Abbreviations

CDS – coding sequence

CoV – coronavirus

Fluc – Firefly luciferase

ORF – open reading frame

RLU – relative light unit

Rluc – *Renilla* luciferase

S1 – SARS-CoV-1

S2 – SARS-CoV-2

SARS – Severe acute respiratory syndrome

SL – stem loop

uORF – upstream open reading frame

1. Introduction

1.1. SARS-CoV-2 virus

A novel coronavirus, severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), was identified in 2019. It was discovered to cause a severe respiratory disease later named coronavirus disease 2019 (COVID-19) (Huang et al., 2020, Zou et al., 2020, Andersen et al., 2020). Due to its high transmission rate in the human population, the spread of the virus has caused an ongoing pandemic with more than 100 million confirmed cases and over 3 million deaths to date according to data issued by the World Health Organisation (covid19.who.int). There has since been a worldwide effort to investigate the mechanisms of infection and disease progression in order to identify targets for potential therapeutic agents and to determine possible ways of slowing down disease transmission.

Coronaviruses are a family of single-stranded RNA viruses that infect and cause diseases in various mammals. They are characterized by distinctly large genomes of around 30 kb, which is the largest known among RNA viruses (Graham and Baric, 2010). SARS-CoV-2 belongs to a group of beta-coronaviruses and is made up of an enveloped single-stranded positive-sense RNA (Figure 1). Upon infection of a host cell, the genomic RNA is translated by cellular machinery and viral proteins are synthesized (Gorbalenya et al., 2020, Lim et al., 2016).

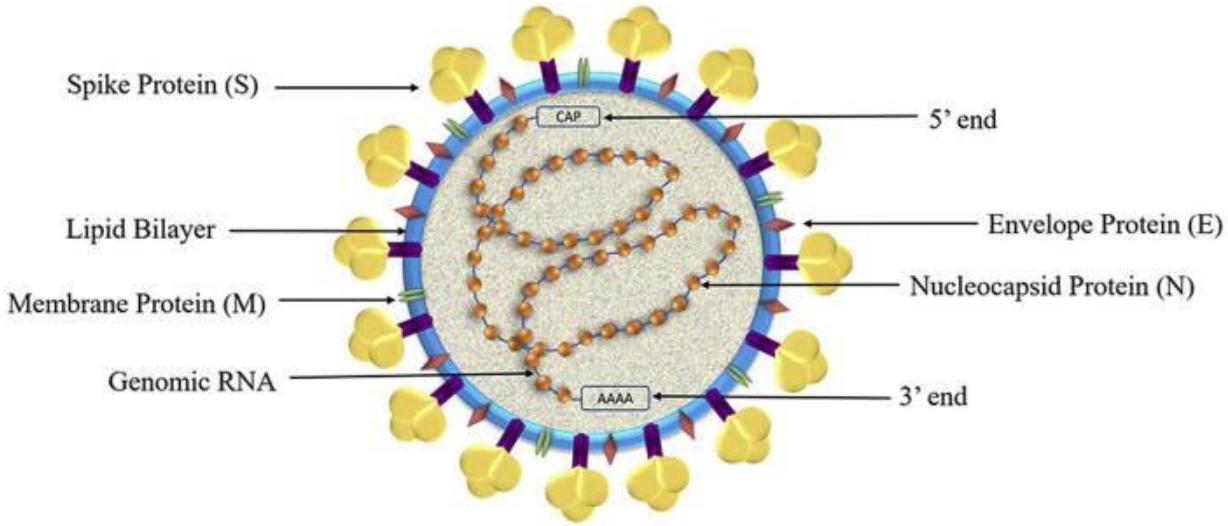

Figure 1. Schematic representation of a SARS-CoV-2 virion. (Satarker and Nampoothiri, 2020)

The genomic RNA encodes a total of 27 proteins which can be divided into three groups: structural proteins, non-structural proteins and accessory proteins (Bar-On et al., 2020). SARS-CoV-2 has 4 structural proteins: nucleocapsid protein (N, which binds viral RNA) and 3 integral membrane proteins – envelope (E), membrane (M) and spike (S) protein. Additionally, there are 16 non-structural proteins (Nsp1-Nsp16) and they comprise of components involved in host immune response control, viral replication and viral RNA synthesis (da Silva et al., 2020, Nakagawa et al., 2016). Lastly, there are 7 accessory proteins (ORF3a-ORF8), however their role is still undetermined (Finkel et al., 2020).

1.2. Translation of SARS-CoV-2 proteins

One translation strategy employed by SARS-CoV-2 is translation of the positive-sense genomic RNA, where polypeptides are synthesized from two overlapping open reading frames (ORF1a and ORF1b, Figure 2) on the genomic RNA and afterwards cleaved into non-structural proteins. Translation of ORF1a enables production of early proteins which are involved in suppresion of host innate immune response. ORF1b translation generates proteins involved in RNA synthesis. Translation of ORF1b is possible because of a programmed -1 frameshift which enables translation to continue after the stop codon of ORF1a (Finkel et al., 2020). The mechanism utilizes a cis-acting RNA element that is able to redirect ribosomes to shift back by one base in the 5' direction and has been described in all other known coronaviruses (de Breyne et al., 2020).

Figure 2. Schematic overview of SARS-CoV-2 genome organization. UTR – untranslated region, ORF – open reading frame, S – spike protein, E – envelope protein, M – membrane protein, N – nucleocapsid protein.

Another translation strategy of SARS-CoV-2 requires a negative-sense RNA intermediate which is synthesized from the genomic positive-sense RNA. The intermediate RNA is then used as a template for the generation of positive-sense subgenomic RNAs (sgRNAs) via a discontinuous, cotranscriptional process. sgRNAs are utilized as mRNAs for translation of ORFs

downstream of ORF1 and mostly encode for structural and accessory proteins which are expressed later during the viral replication cycle (Sola et al., 2015).

Translation of viral proteins begins immediately after the genome is released into the cytoplasm upon infection when ORF1 polyproteins are synthesized and cleaved into non-structural proteins. The most N-terminal cleavage product is the protein Nsp1 and it is not present in all coronaviruses, only alpha- and beta-coronaviruses (Kamitani et al., 2009). Its size is highly variable, from around 110 aa in alpha-coronaviruses to up to 245 aa in some beta-coronaviruses. However, Nsp1 protein in SARS-CoV-1 and SARS-CoV-2 viruses is similar in size (180 aa) and although some sequences are different between the two homologs, they share 84.4 % sequence identity suggesting high conservation of function between them (de Breyne et al., 2020, Thoms et al., 2020).

1.3. Role of Nsp1 protein

The role and mechanism of action of Nsp1 in SARS-CoV-1 infection have been explored extensively since the virus was first identified in 2002 and subsequently caused an epidemic in southern China (Drosten et al., 2003, Tanaka et al., 2012). Interest in the protein stems from its early-cycle synthesis and the fact that its expression is detectable at 6 hours post-infection, which is roughly the same time as global translation inhibition begins in infected cells. In fact, it has been shown that transient expression of SARS-CoV-1 Nsp1 is sufficient to promote global inhibition of host protein synthesis (Kamitani et al., 2006, Narayanan et al., 2008). Two amino acids, K164 and H165, have been identified and confirmed to be essential for this inhibitory effect of the protein. It has also been shown that mutating the mentioned amino acids into alanine reverts global translation shutoff in host cells without affecting viral replication, viral RNA transcription or expression of Nsp1 itself (Narayanan et al., 2008). Additionally, SARS-CoV-1 Nsp1 has shown the same inhibitory effect on endogenous cellular mRNAs as well as transfected mRNAs and mRNAs from ectopic plasmid expression (Kamitani et al., 2006, Narayanan et al., 2008).

There are two mechanisms by which SARS-CoV-1 Nsp1 controls host gene expression - translational arrest and RNA degradation, as shown in Figure 3 (Nakagawa et al., 2016). Although the two mechanisms are linked processes, what differentiates between them is the amino acid site mutations required for the abolishment of each of them.

The first mechanism by which Nsp1 contributes to translation repression is through binding of 40S ribosomal subunit which results in reduction of the number of polysomes available for translation in the cell. Western blot evidence shows that this binding ability of Nsp1 is abolished in the case of KH164AA mutation, meaning these two amino acids are essential for ribosome binding, and consequently, translation inhibition and viral infection (de Breyne et al., 2020, Kamitani et al., 2009). A second mechanism by which Nsp1 controls gene expression is through induction of cellular mRNA degradation (Kamitani et al., 2006, Narayanan et al., 2008). Two exposed and charged amino acids, R124 and K125, have been identified as essential for this function. Research shows that mutations in those sites are sufficient to prevent host mRNA degradation, without affecting translation inhibition which indicates clear distinction between the two mechanisms (Lokugamage et al., 2012).

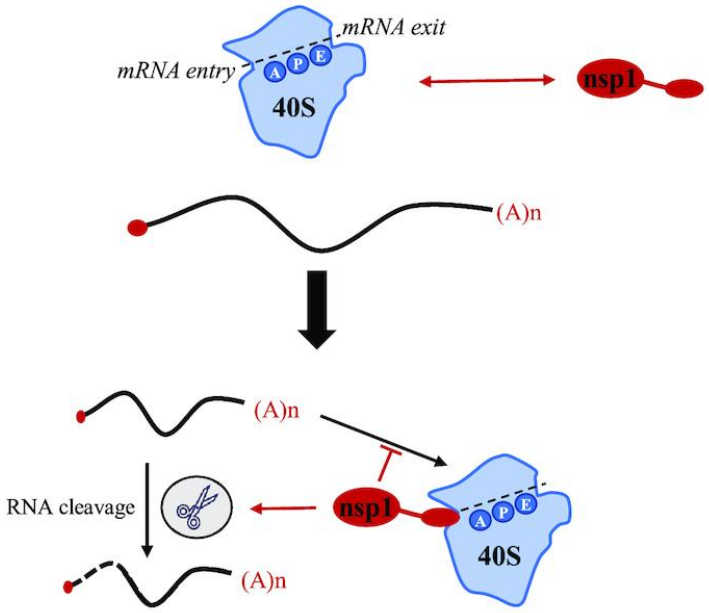

Figure 3. Overview of SARS-CoV-1 Nsp1-mediated host gene expression control. 40S ribosomal subunit is showed with mRNA entry and exit channels and three tRNA sites: acceptor (A), peptidyl (P) and exit (E) site. It is shown how Nsp1 binds to the 40S ribosomal subunit, blocking the mRNA entry channel. In addition, Nsp1 promotes cleavage of RNA and as a result cellular protein expression is reduced. (de Breyne et al., 2020)

1.4. Leader sequences in viral 5' UTR regioin

Despite translation repression and cellular mRNA degradation in SARS-CoV-1-infected cells, viral mRNAs are capable of evading degradation. In fact, viral protein synthesis is maintained during infection, suggesting that a mechanism of protection of viral mRNAs exists. This mechanism is not completely clear yet, however it seems that viral mRNAs contain a *cis*-acting element which enables the escape from host cell translation shutoff. The *cis*-acting element is believed to be within the 5' UTR of the viral mRNA because it was shown that the SARS-CoV-1 5' UTR sequence alone was sufficient to protect transcripts against RNA cleavage (Huang et al., 2011).

The length of the 5' UTR varies between different coronaviruses. Additionally, the length differs between mRNAs of each coronavirus and for SARS-CoV-2 its size ranges between 75 nucleotides and 265 nucleotides (Kim et al., 2020). For SARS-CoV-1 mRNAs it ranges between 72 and 264 nucleotides (Yang et al., 2009). Despite the variety in 5' UTR length, the initial ~72 nucleotides are strictly conserved and refered to as the leader sequence (de Breyne et al., 2020).

The leader sequence folds itself into secondary structures which enable RNA-RNA interactions as well as RNA-protein interactions during translation and replication (Yang and Leibowitz, 2015). The leader sequence in SARS-CoV-1 folds into several stem loop structures and it has been shown that Nsp1 is able to interact with the stem loop 1 (SL1) structure (Figure 4). This specific interaction appears to be sufficient for escaping translational repression. In fact, research shows that the amino acids required for 40S binding (K164, H165) and for RNA cleavage (R124, K125) are necessary for SARS-CoV-1 SL1 binding as well (Tanaka et al., 2012). This process is not completely clear yet and the exact mechanism is still unknown, but what is evident is that the stem loop structures of the SARS-CoV-1 leader sequence are able to protect viral mRNA from Nsp1-mediated translation inhibition.

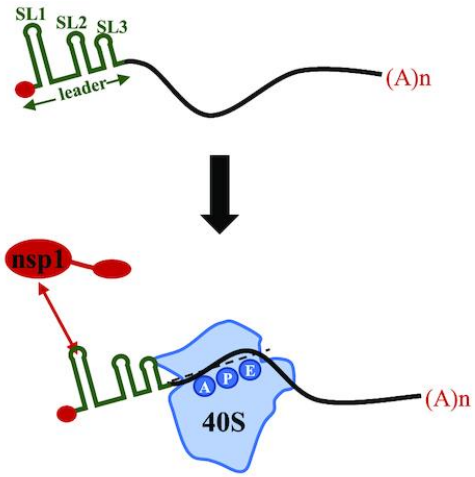

Figure 4. Overview of SARS-CoV-1 mRNA translation. 40S ribosomal subunit is shown with three tRNA sites: acceptor (A), peptidyl (P) and exit (E) site. It is shown how Nsp1 interaction with the leader sequence allows translation of viral mRNA. (de Breyne et al., 2020)

1.5. Reporter assays for gene expression research

Reporter genes are a frequently used method in molecular biology, pharmaceutical and biomedical research. Detection of the reporter genes is usually done through either direct visualization, as is the case of fluorescent proteins, or through detection of the enzymatic activity of the reporter protein, as in the case of galactosidases or luciferases. In general, the reporter protein coding sequences are cloned into an expression vector next to the sequence of interest (protein, UTR, peptide sequence) resulting in generation of a fusion protein. The reporter protein then enables detection of the fused target sequence/protein which in turn makes it possible to study the expression of the target protein (Allard and Kopish, 2008).

Luciferase reporters can be used to study protein expression and regulation of expression by cloning them next to a protein or sequence of interest and then transfecting into a cell culture system. Luciferase activity is measured by adding the proper substrate and detecting a light signal using a luminometer. The detected light is quantified and values obtained are taken as a measure of expression of the target protein (Allard and Kopish, 2008, Carter and Shieh, 2015). There is variability in conditions within each well of a cell culture plate, so in order for the test to be reliable, each condition should be repeated in triplicates. Utilizing luciferase reporters in this way enables research of protein expression in various conditions through a relatively simple and quick approach (Carter and Shieh, 2015).

In this thesis, a dual luciferase reporter approach is used, meaning that two reporter enzymes' - *Photinus pyralis* (firefly) and *Renilla reniformis* luciferase, activities are measured sequentially from the same sample. This is possible due to the structural differences between the two enzymes, in addition to them requiring different substrates for reactions. Thus, by adding appropriate reagents, firefly luciferase (Fluc) activity is measured first and the reaction is then quenched and *Renilla* luciferase (Rluc) activity is measured. This way, it is possible to use one reporter to investigate translation changes in different experimental conditions, while the contransfected reporter acts as an internal control. This approach enables reliable interpretation of obtained experimental results because it minimizes experimental variability caused by discrepancies in cell viability, transfection and lysis efficiency. (Allard, 2008).

For the purpose of this thesis, *Renilla* luciferase was used to represent viral transcripts by cloning the SARS-CoV-2 leader sequence upstream of the reporter coding sequence (CDS), under a CMV promoter. Its activity was compared to Rluc activity under the same promoter where no upstream ORFs (uORFs) were added. In that case, the reporter represents cellular transcripts. Different experimental conditions were achieved by cotransfecting the reporter plasmids with varying amounts of Nsp1-expressing plasmid from both SARS-CoV-1 and SARS-CoV-2 as well as their mutants (KH164AA, RK124AA, R124A). With both Rluc reporters, firefly luciferase was used as an internal control of the experiment.

Typically, the addition of an upstream ORF decreases the likelihood of translation of the main reading frame because when eukaryotic ribosomes scan mRNA, translation is initiated after the first start codon is read. Since uORF start codons are usually out of frame with the main ORF, this causes the ribosome to initiate translation from the uORF start codon and then scan past the main start codon (Calvo et al., 2009). Consequently, for the experiments in this thesis, it means that lower *Renilla* luciferase activity should be recorded for the reporter containing the uORF when compared to the *Renilla* luciferase reporter without an uORF due to the nature of the constructs.

2. Research aim

The aim of this thesis is to investigate the effect that the SARS-CoV-2 protein Nsp1 has on translation initiation in human cells. By implementing a reporter-based approach, this research explores different expression patterns of viral and non-viral (cellular) transcripts in the presence of Nsp1 and the following mutants: RK124AA, RK124A, KH164AA. Viral transcripts are represented by a *Renilla* luciferase reporter with upstream SARS-CoV-2 leader sequence and non-viral transcripts are represented by a *Renilla* luciferase reporter without the upstream leader sequence. Additionally, experiments are replicated with the homolog protein from the SARS-CoV-1 virus as a comparison of function conservation between the Nsp1 proteins.

3. Materials and methods

3.1. General cloning protocols

3.1.1. Agarose gels

Agarose gels (1 %) were prepared by dissolving laboratory grade agarose (AppliChem, A8963) in TBE buffer (0.089 M Tris base, 0.089 M Boric acid, 0.002 M EDTA, pH = 8.3). DNA samples were loaded on the gel using 6x Gel Loading Dye (NEB, B7024S). Electrophoresis was done with 110 V using PowerPac power supply from Bio-Rad. DNA was extracted from the gel using Jena Bioscience extraction kit (PP-202L) following manufacturer's instructions (gel containing the DNA fragment is excised and dissolved, then loaded on a column and after washing the column, the DNA is eluted).

3.1.2. PCR reactions

PCR reactions were done using Pfu polymerase in Pfu buffer (200 mM Tris-HCl, 100 mM $(NH_4)SO_4$, 100 mM KCl, 1 % (v/v) Triton X-100, 1 mg/ml BSA, 20 mM $MgSO_4$) obtained from the MDC facility. For one 50 µl reaction the following was used: 1 µl of plasmid template DNA (5 ng/µl), 2 µl of forward and reverse primers (10 µM), 5 µl of 10x Pfu buffer, 1 µl of dNTPs mix (10 mM), 0.25 µl of Pfu polymerase and ddH_2O up to 50 µl. Standard PCR conditions were 95 °C for 5 minutes, 95 °C for 30 seconds, 30 seconds at primer-specific annealing temperature, 72 °C for elongation (1 minute for 1 kb) depending on amplicon size, 25 cycles. Final extension was done at 72 °C and after that, samples were held on 12 °C.

PCR reaction for site-directed mutagenesis was done using Q5 polymerase in Q5 buffer. For a 20 µl reaction the following was used: 1 µl of plasmid template DNA (50 ng/µl), 1 µl of forward and reverse primers (10 µM), 0.4 µl of dNTPs mix (10 mM), 4 µl of 5x Q5 Polymerase buffer, 0.2 µl of Q5 polymerase (2 U/µl) and ddH_2O up to 20 µl. Negative control sample was set up without primers to assess the background of undigested template. Conditions used were the following: 30 seconds at 98 °C, 10 seconds at 98 °C. Then 40 seconds per 1 kb at 72 °C repeated for 25 cycles and after that samples were kept at 8 °C. Samples were then digested with DpnI enzyme (FastDigest, FD1704, Thermo Scientific, 0.5 µl of 20 U/µl) for 30 minutes to overnight at 37 °C.

3.1.3. Cloning

Digestions were done using NEB restriction enzymes SbfI (R3642S), NotI (R3189S) in Cutsmart buffer. Between 3 and 5 µg of plasmid DNA was used, digestions were performed at 37 °C for 2 hours or up to overnight. Ligation reactions were done using 50 ng of vector and the amount of insert was adjusted according to a 1:3 molar ratio. For T4 ligation reactions, 1 µl of T4 DNA ligase was used in 10 µl reaction volume for 1 hour at room temperature.

Gibson assembly reactions were done at 50 °C for 1 hour using Gibson assembly master mix in 5 µl reaction volume.

5 µl of both T4 ligation reactions or Gibson assembly reactions was used for transformations of XL-blue *E. coli* cells (competency 10^6). After bacterial cells were thawn on ice, ligation mixes were added and incubated for another 10 minutes on ice. Transformation was done for 1 minute at 42 °C in a water bath, after which the bacteria was put back on ice for another 2 minutes. SOB medium (AE27.1, Carl Roth) was added to each sample (150 µl) and bacterial cells were recovered shaking for 1 hour at 37 °C. After that, bacteria were plated on agar plates with the appropriate antibiotic added. Negative controls were set up for each cloning reaction containing the cut plasmid, but without insert (reaction volume was adjusted using ddH_2O).

The following day, 3 to 5 colonies were picked and grown in LB medium (X968.3, Carl Roth, with appropriate antibiotic added) overnight at 37 °C, with shaking. Plasmid DNA was purified the next day using the Mini-prep kit from Roboklon, according to manufacturer's instructions. Sequences of plasmids were confirmed by Sanger sequencing through LGC Genomics.

3.2. Specific cloning strategies

Nsp1 coding sequences were amplified using oligos listed in Table 1 and SARS-CoV-1 and SARS-CoV-2 cDNA were used as templates (cDNAs were not generated as part of this thesis). PCR products were cut with SbfI (NEB, R3642S) and NotI (R3189S) and cloned into a SbfI/NotI cut pEBG vector via T4 ligation. Nsp1 mutants were generated via site-directed mutagenesis using oligos listed in Table 1.

Table 1. A list of oligo sequences used for PCR reactions. Fw, rev – forward and reverse primer.

Oligo name	Number of oligo in lab collection	Oligo sequence
SbfI-Nsp1-S2- fw	4284	ctggttccgcgtggatctcctgcaggcctcgagatggagagccttgtccctg
NotI-Nsp1-S2-rev	4285	gaccctcactctagagtcgcggccgcttaccctccgttaagctcacgc
SbfI-Nsp1-S1-fw	4314	ctggttccgcgtggatctcctgcaggcctcgagatggagagccttgttcttggtg
NotI-Nsp1-S1-rev	4315	gaccctcactctagagtcgcggccgcttaacctccattgagctcacgag
Nsp1-S2-R124A-fw	4310	cgcaaggttcttcttgctaagaacggtaataaaggagctg
Nsp1-S2-R124A -rev	4311	ttattaccgttcttagcaagaagaaccttgcggtaagcc
Nsp1-S2-KH164AA-fw	4308	aactggaacactgcagctagcagtggtgttacccgtgaac
Nsp1-S2-KH164AA-rev	4309	gtaacaccactgctagctgcagtgttccagttttcttgaa
Nsp1-S1_KH164AA-fw	4350	aactggaacactgcagctggcagtggtgcactccgtgaac
Nsp1-S1_KH164AA-rev	4351	agtgcaccactgccagctgcagtgttccagttttgttcat
Nsp1-S1_R124A-fw	4352	gcaatgttcttcttgctaagaacggtaataagggagccg
Nsp1-S1_R124A-rev	4353	ttattaccgttcttagcaagaagaacattgcggtatgca
Nsp1-S2-RK124AA-fw	4360	cgcaaggttcttcttgctgcgaacggtaataaaggagctg
Nsp1-S2-RK124AA-rev	4361	ttattaccgttcgcagcaagaagaaccttgcggtaagcc
Nsp1-S1-RK124AA-fw	4427	gcaatgttcttcttgctgcgaacggtaataagggagccg
Nsp1-S1-RK124AA-rev	4428	ttattaccgttcgcagcaagaagaacattgcggtatgca

Firefly luciferase reporter plasmid (number 678, as listed in Table 2), *Renilla* luciferase reporter plasmid (number 1028, as listed in Table 2) and flag-tagged sic-peptide expressing

plasmids (number 1238, as listed in Table 2) were not created for the purpose of this thesis, but were available in the lab database. *Renilla* luciferase reporter plasmids with upstream SARS-CoV-2 leader sequence, as well as nucleotide deletions in the CMV promoter, were created by David Koppstein as part of the ongoing project in the lab and the generated plasmids were used in this thesis. The plasmids were generated in the following way: SARS-CoV-2 leader sequence was inserted into a *Renilla* luciferase reporter vector, upstream of the reporter coding sequence and downstream of the CMV promoter. All additional nucleotides between the CMV promoter and the newly inserted leader sequence were removed (Figure 5A). For some experiments, 8 nucleotides from the CMV promoter in *Renilla* luciferase reporter plasmids were removed by PCR with primers lacking the 8 nt sequence (Figure 5B). Nucleotides were removed from the CMV promoter in the control *Renilla* luciferase reporter vector (without upstream viral sequence) as well, and used as negative control. All of the plasmids used are listed in Table 2.

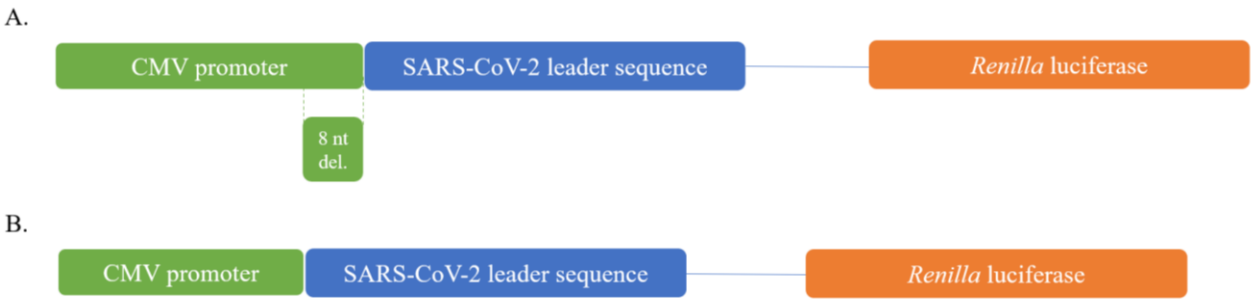

Figure 5. Schematic representation of *Renilla* luciferase reporter constructs with an upstream viral leader sequence. In (A), reporter construct without nucleotide deletions is shown. In (B), additional reporter construct where 8 nucleotides from the CMV promoter have been deleted is shown. The region where 8 nucleotides were deleted is marked in (A).

Table 2. A list of plasmids used in this thesis and short description of each plasmid

Plasmid name	Plasmid number in lab collection	Plasmid description
pEBG-flag-sic	1238	Plasmid expressing flag-tagged sic peptide
pEBG-flag-Nsp1-S2	2280	Plasmid expressing flag-tagged Nsp1 from SARS-CoV-2
pEBG-flag-Nsp1-RK124AA-S2	2334	Plasmid expressing flag-tagged RK124AA mutant of Nsp1 from SARS-CoV-2
pEBG-flag-Nsp1-R124A-S2	2319	Plasmid expressing flag-tagged R124A mutant of Nsp1 from SARS-CoV-2
pEBG-flag-Nsp1-KH164AA-S2	2320	Plasmid expressing flag-tagged KH164AA mutant of Nsp1 from SARS-CoV-2
pEBG-flag-Nsp1-S1	2304	Plasmid expressing flag-tagged Nsp1 from SARS-CoV-1
pEBG-flag-Nsp1-RK124AA-S1	2377	Plasmid expressing flag-tagged RK124AA mutant of Nsp1 from SARS-CoV-1
pEBG-flag-Nsp1-R124A-S1	2332	Plasmid expressing flag-tagged R124A mutant of Nsp1 from SARS-CoV-1
pEBG-flag-Nsp1-KH164AA-S1	2333	Plasmid expressing flag-tagged KH164AA mutant of Nsp1 from SARS-CoV-1
pCiNeo Firefly luciferase	678	Firefly luciferase reporter plasmid
pCiNeo *Renilla* luciferase	1028	*Renilla* luciferase reporter plasmid
pCiNeo-CoV-2-leader-Rluc	2344	*Renilla* luciferase reporter plasmid with upstream SARS-CoV-2 leader sequence
pCiNeo-Rluc-8nt	2417	*Renilla* luciferase reporter plasmid with 8 nucleotides removed from the CMV promoter
pCiNeo-CoV-2-leader-Rluc-8nt	2420	*Renilla* luciferase reporter plasmid with upstream SARS-CoV-2 leader sequence and 8 nucleotides removed from the CMV promoter

3.3. Cell culture and transfection conditions

HEK 293T (human embryonic kidney derivative, highly transfectable) cells were cultured with DMEM Glutamax media (Life technologies, GmbH, 31966047) with 10 % FBS (Life technologies, GmbH, FBS 10270106) and 1x penicilin/streptamycin (Biochrom, PAN P06-07100) at 37°C and humid atmosphere.

For transfection, HEK 293T cells were seeded at 7000 cells / well density in 96-well cell culture plates. The following day, the cells were transfected with plasmid DNA using polyethyleneimine (PEI) as a transfection reagent (181978-100G, Sigma-Aldrich). Between 36 and 50 ng of total plasmid DNA was used per well and between 108 and 150 ng of PEI per well to maintain a molar ratio of 1:3. DNA and PEI were mixed well in DMEM Glutamax (100 µl per well) and incubated for 10 minutes. After that, the transfection mix was added to cells.

For each well, either 0.5, 2, 8 or 32 ng of plasmid DNA expressing Nsp1 CDS (wild-type or mutant) was transfected. Firefly (between 3 and 17 ng) and *Renilla* luciferase (1 ng) reporter plasmids were co-transfected with Nsp1-expressing plasmid. Plasmid expressing flag-tagged sic peptide was used in negative control samples instead of Nsp1-expressing plasmids. To ensure that the same amount of DNA is used for transfection in each well, the negative control vector was additionally used as a „filler“ plasmid in samples where less than 32 ng of Nsp1-expressing plasmid was used for transfection. The full list of plasmids used for transfection is listed in Table 2.

22 hours after transfection, the medium was removed and cells were lysed in 40 µl of 1x Passive lysis buffer (Promega, E1941). Samples were left lightly shaking for 30 minutes at room temperature and 3 µl of each sample was used for luciferase reporter readout in a 96-well plate.

3.4. Luciferase reporter assay readout

For dual luciferase reporter assay, Centro XS^3 LB 960 luminometer was used to measure *Renilla* and Firefly luciferase activity and Microwin 2000 program to analyse the data. For Firefly luciferase activity detection, 5 µl of luciferin were added to 5 ml LARII buffer (75 mM Hepes, 0.1 mM EDTA, 4 mM $MgSO_4$, 530 µM ATP, 270 µM Coenzyme A, 470 µM DTT, pH = 8). For *Renilla* luciferase activity detection, 25 µl of coelenterazine were added to 5 ml STOP&GLO buffer (2.2 mM Na_2EDTA, 220 mM K_XPO_4, 0.44 mg/ml BSA, 1.1 M NaCl, 1.3

mM NaN_3, pH = 5). Settings for automatic reagent uptake and measurement time were selected and detection of luciferase activity was done in the linear range of 10.000-1.000.000 RLU for Firefly luciferase and 25.000-80.000.000 RLU for *Renilla* luciferase.

3.5. Western blot

Technical replicates were pooled and total 15 µl was used for loading on SDS-PAGE gel (equal to 12.5 % of total sample). Samples were mixed with 5x SDS loading dye and loaded onto 12.5 % SDS-polyacrylamide gels. Gels were run first at 80 V for 30 minutes and subsequently at 100 V for 1 hour. Gels were then transferred onto PVDF membranes using semi-dry transfer for 30 minutes (25 V, 1 A). Membranes were blocked for 1 hour with 3 % milk in TBST buffer (25 mM Tris, 150 mM NaCl pH = 7.5, 0.1 % Tween-20) and incubated with primary antibodies at 4 °C overnight in appropriate dilutions. The following day, membranes were washed 3x for 15 minutes using TBST buffer and incubated with secondary antibodies for 1 hour at room temperature. Membranes were washed again 3x for 15 minutes with TBST buffer. Membranes were incubated with ECL reagent (100 mM Tris pH = 8.5, 2.6 µl of 30% H_2O_2, 0.225 mM coumaric acid, 2.5 mM luminol) to detect the signal and imaged using Amersham Imager 600.

The following primary antibodies were used: anti-flag (Sigma, F3165), anti-beta-actin (Sigma, A2228). Anti-mouse-HRP (Abcam, ab99632) was used as a secondary antibody.

4. Results

4.1. Initial luciferase reporter assays

To test the effects of Nsp1 protein on translation in HEK 293T cells, plasmids expressing *Renilla* and Firefly luciferase were cotransfected with Nsp1-expressing plasmids (wild-type or mutant) for reporter assay experiments. Each reporter assay was repeated in at least two independent experiments. In each experiment, samples were prepared in three technical replicates and average values were calculated from those and used for analysis. The values were normalized to the negative control sample for which a plasmid expressing a flag-tagged sic-peptide was used in combination with a *Renilla* luciferase reporter without an upstream leader sequence. For each reporter used, a control sample was prepared where no Nsp1 was added.

Initial reporter assay experiments were done using Rluc and Fluc reporter constructs without any nucleotide deletions in the CMV promoter region (Figure 6). In those experiments, it was shown that *Renilla* luciferase expression significantly decreases in the presence of Nsp1 (60 % drop when 0.5 ng of Nsp1-expressing plasmid DNA is added). The decrease becomes more apparent with higher concentrations of Nsp1-expressing plasmid DNA used, with a 90 % drop observed when 32 ng of plasmid DNA were used, suggesting that Nsp1 has a great impact on protein synthesis. A similar negative trend in Rluc expression was observed with the gradual increase of plasmid DNA expressing Nsp1 with the RK124AA mutation. However, RLU values for Rluc remained several fold higher than they were in the presence of wild-type Nsp1 across all concentrations tested. On the other hand, KH164AA mutant did not inhibit translation of *Renilla* luciferase, instead it was observed that Rluc expression actually increases compared to control values when 32 ng of Nsp1-KH164AA-expressing plasmid DNA were used.

Short upstream ORFs reduce the likelihood of translation of the main ORF as was observed with the addition of the SARS-CoV-2 leader sequence upstream of *Renilla* luciferase sequence. An 80 % decrease in expression of Rluc was observed when compared to the expression level of Rluc without any upstream sequences added. Additionally, this experiment showed that the Rluc reporter with an upstream ORF maintained the same negative trends in expression as was the case with the previous reporter (without an upstream leader sequence). When concentrations of the Nsp1-expressing plasmid were increased, expression of this reporter was downregulated as was the case with *Renilla* luciferase reporter with no upstream leader

sequence. A similar negative trend was observed with the RK124AA Nsp1 mutant and the values were similar to the samples where wild-type Nsp1 was used. With this *Renilla* reporter, KH164AA mutant once again caused the expression of Rluc to increase compared to the control sample, suggesting that the mutations in the protein prevent its inhibitory effect on translation.

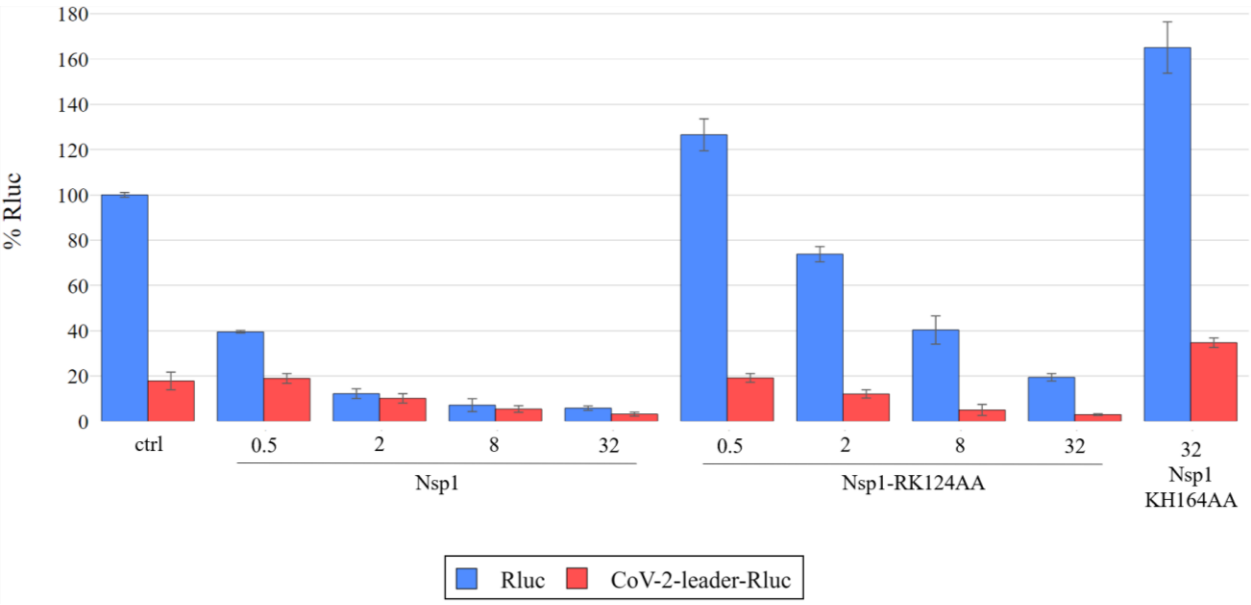

Figure 6. SARS-CoV-2-Nsp1-mediated repression of *Renilla* luciferase in reporter assay. Rluc reporter with no upstream ORFs (in blue) represents cellular transcripts and Rluc reporter with 5' viral UTR upstream of Rluc CDS (in red) represents viral transcripts. Reporter expression was measured for 0.5, 2, 8 and 32 ng of Nsp1-expressing plasmid (and RK124AA mutant). For KH164AA mutant only 32 ng was used. No Nsp1 was added to control (ctrl) samples. Values are normalized to the negative control sample where Rluc reporter with no upstream ORFs was used. Error bars represent standard deviation calculated from three technical replicates of each sample. The data was replicated in two separate experiments.

4.2. SARS-CoV-2 Nsp1 effect on shorter Rluc reporter plasmids expression

The previous experiment was repeated, this time using *Renilla* luciferase reporter constructs that have 8 nucleotide deletions in the CMV promoter region. Once again, as seen in Figure 7 (in blue), it was shown that in the presence of Nsp1, the expression of *Renilla* luciferase drops proportionally with increasing concentration of Nsp1-expressing plasmid used. In fact, a decrease in Rluc expression of over 90% was measured when 32 ng of Nsp1-expressing plasmid

DNA was used. Interestingly, the RK124AA mutant again showed a similar effect on Rluc expression as wild-type Nsp1, causing a downward trend in Rluc expression when the amount of plasmid expressing the mutant was increased. However, the values were up to 4 times higher when compared to samples with corresponding concentrations of plasmid expressing wild-type Nsp1. Additionally, in this experiment, a set of samples with R124A mutant was included and in that case we see that the values and the trend in expression of Rluc is equal to the one where RK124AA mutant was used, suggesting that a single mutation in the site is sufficient to achieve the same effect. The KH164AA mutant seems to not affect the expression of Rluc as the values are close in value to negative control (the difference between them is not significant).

Again, it was shown that with the addition of an upstream ORF, the expression of *Renilla* luciferase drops 90 % when there is no Nsp1 present (Figure 7, in red). However, in this experiment, presence of Nsp1 promotes translation of *Renilla* luciferase with an upstream leader sequence with a 3 fold increase in expression observed when 0.5 and 2 ng of Nsp1-expressing plasmid DNA was used. This could suggest that SARS-CoV-2 Nsp1 allows translation of viral sequences, although, with increased concentration of Nsp1-expressing plasmid DNA, expression of Rluc decreases. In the case of RK124AA and R124A mutants, Rluc reporter expression decreases with the addition of plasmid DNA expressing the mutant proteins. The visible trend is similar to the trend present when an Rluc reporter without upstream ORFs was used (Figure 7, in blue). There is a downward trend in expression present in both cases, suggesting that either RK124AA or R124A mutations are sufficient to prevent Nsp1-mediated upregulation of viral mRNA translation. Additionally, the results suggest that a single mutantion in the site is sufficient to achieve that effect. Interestingly, the KH164AA mutant samples showed a two fold increase in expression of the reporter protein compared to the control sic-peptide sample.

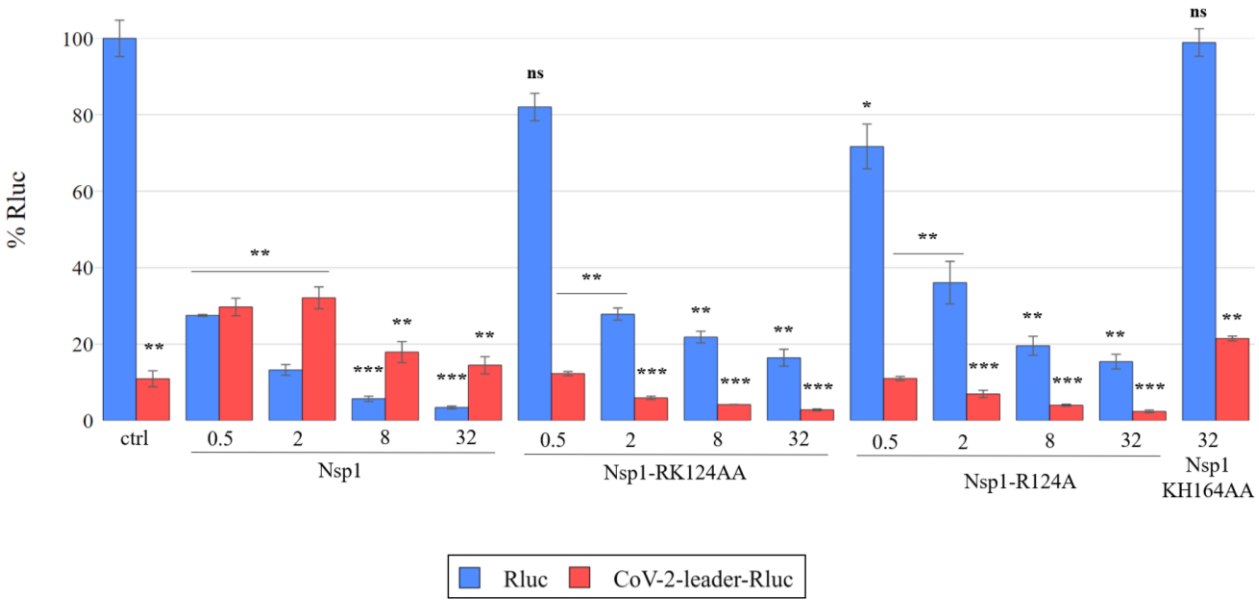

Figure 7. *Renilla* luciferase expression in the presence of SARS-CoV-2 Nsp1 in reporter assay. *Renilla* luciferase reporter with upstream leader sequence escapes SARS-CoV-2-Nsp1-mediated repression. Rluc reporter with no upstream ORFs (in blue) represents cellular transcripts and Rluc reporter with 5' viral UTR upstream of Rluc CDS (in red) represents viral transcripts. Reporter expression was measured for 0.5, 2, 8 and 32 ng of Nsp1-expressing plasmids (and mutants). For KH164AA mutant only 32 ng was used. No Nsp1 was added to control (ctrl) samples. Values are normalized to the negative control sample where Rluc reporter with no upstream ORFs was used. Error bars represent standard deviation of three technical replicates for each sample. The data was replicated in two separate experiments. Significance was calculated with student's t-test (two-tailed, unequal variance). ns stands for nonsignificant, degrees of significance are indicated as following: *0.01 < p < 0.05, **0.001 < p < 0.01, ***p < 0.001.

4.3. SARS-CoV-1 Nsp1 effect on shorter Rluc reporter plasmids expression

The experiment was repeated with SARS-CoV-1-Nsp1-expressing plasmid DNA and the set of reporters with 8 nucleotides removed from the CMV region. In this experiment, similar results were obtained as was the case with SARS-CoV-2 Nsp1 (Figure 8). Wild-type Nsp1 did not inhibit Rluc expression only in the case when there was viral leader sequence upstream of the reporter CDS. Additionally, the trends in expression were equal to those in the previous experiment where SARS-CoV-2 Nsp1 was used and again with higher concentration of Nsp1-expressing plasmid, Rluc expression began to decrease. RK124AA and R124A mutants inhibited

Rluc translation in both reporter constructs as was the case with SARS-CoV-2 Nsp1 mutants. Lastly, the KH164AA mutant samples did not downregulate *Renilla* luciferase expression, in the case of the reporter with the leader sequence it even increased its expression.

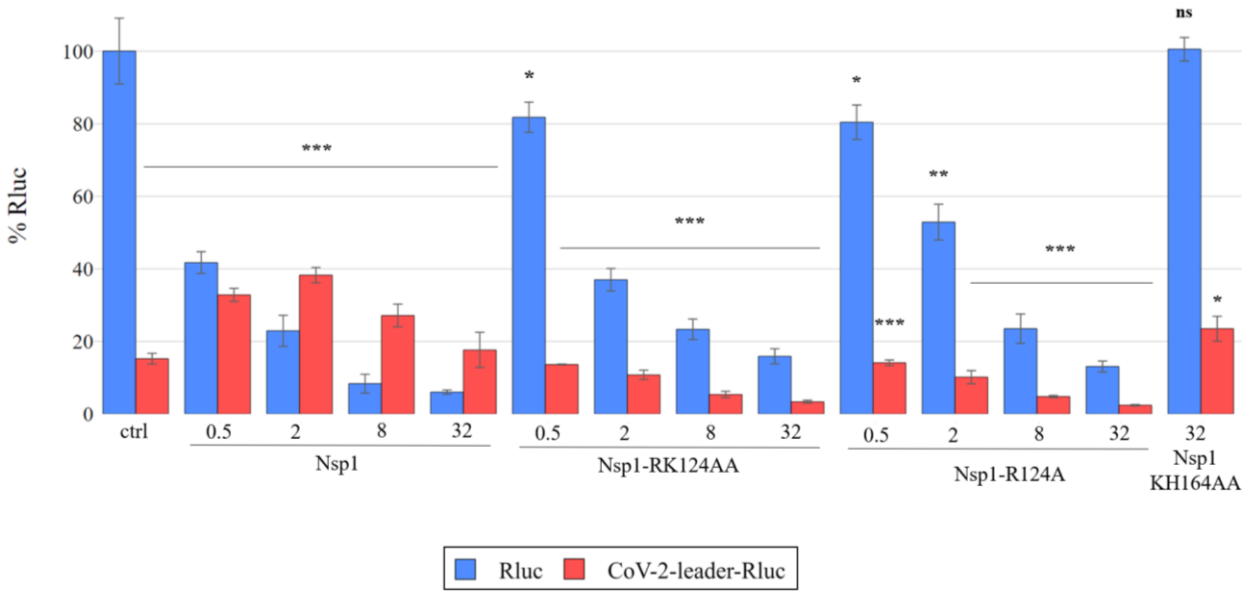

Figure 8. *Renilla* luciferase expression in the presence of SARS-CoV-1 Nsp1 in reporter assay. *Renilla* luciferase reporter with upstream leader sequence escapes SARS-CoV-1-Nsp1-mediated repression. Rluc reporter with no upstream ORFs (in blue) represents cellular transcripts and Rluc reporter with 5' viral UTR upstream of Rluc CDS (in red) represents viral transcripts. Reporter expression was measured for 0.5, 2, 8 and 32 ng of Nsp1-expressing plasmids (and mutants). For KH164AA mutant only 32 ng was used. No Nsp1 was added to control (ctrl) samples. Values are normalized to the negative control sample where Rluc reporter with no upstream ORFs was used. Error bars represent standard deviation calculated from three technical replicates of each sample. The data was replicated in two separate experiments. Significance was calculated with student's t-test (two-tailed, unequal variance). ns stands for nonsignificant, degrees of significance are indicated as following: *0.01 < p < 0.05, **0.001 < p < 0.01, ***p < 0.001.

Firefly luciferase was co-transfected with *Renilla* luciferase and Nsp1-expressing plasmid in all experiments and used as an internal control since no deletions or insertions were done on the reporter plasmid. In each experiment, the values obtained for the Firefly luciferase reporter showed the relative trend of expression equal to the one of *Renilla* luciferase plasmid without

any upstream ORFs added. For that reason, the Firefly luciferase values were not used as a baseline for normalization as the obtained data did not show high variability and normalization did not affect statistical significance.

4.4. Nsp1 expression in reporter assay samples

Reporter assay sample triplicates were pooled together and analysed with Western blot to validate the expression of Nsp1 with different amounts of plasmid DNA used. Samples that were transfected with Rluc reporter with no upstream ORFs were used for western blot (data shown in Figures 7 and 8, in blue).

Nsp1 protein has an estimated size of 21 kDa and the flag peptide is 1 kDa in size so a band at approximately 22 kDa was expected. This band was detected for all of the plasmid DNA amounts transfected except for the lowest one (0.5 ng), as visible in Figure 9. Band intensity was stronger with increased amount of Nsp1-expressing plasmid DNA transfected. A similar trend was observed in the case of RK124AA and R124A mutants. KH164AA mutant showed a high intensity band, suggesting high expression of the protein.

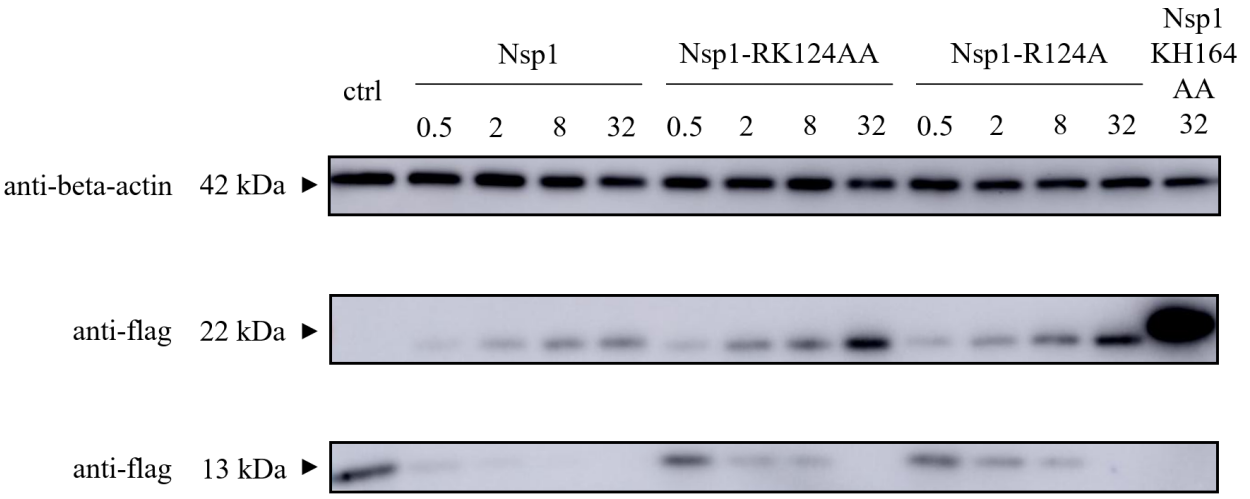

Figure 9. Western blot membrane for pooled reporter assay samples for SARS-CoV-2 Nsp1. Anti-flag antibody is used to visualise flag-tagged Nsp1 and its mutants (22kDa). Expression is not detected for samples with 0.5 ng plasmid DNA used. Band intensity increases with more DNA used for transfection. KH164AA mutant shows a high intensity band indicating high level of expression. Band at 13 kDa corresponds to flag-tagged sic-peptide size which was used as a "filler" plasmid and negative control.

SARS-CoV-1 Nsp1 expression was detected only with higher concentrations of plasmid used (8 and 32 ng, Figure 10). For RK124AA and R124A mutants, a band is visible for the samples in which 2 ng was used, as well. SARS-CoV-1 KH164AA mutant shows a much more intense band than other SARS-CoV-1 samples, as was the case for the mutant in SARS-CoV-2 again suggesting high expression levels.

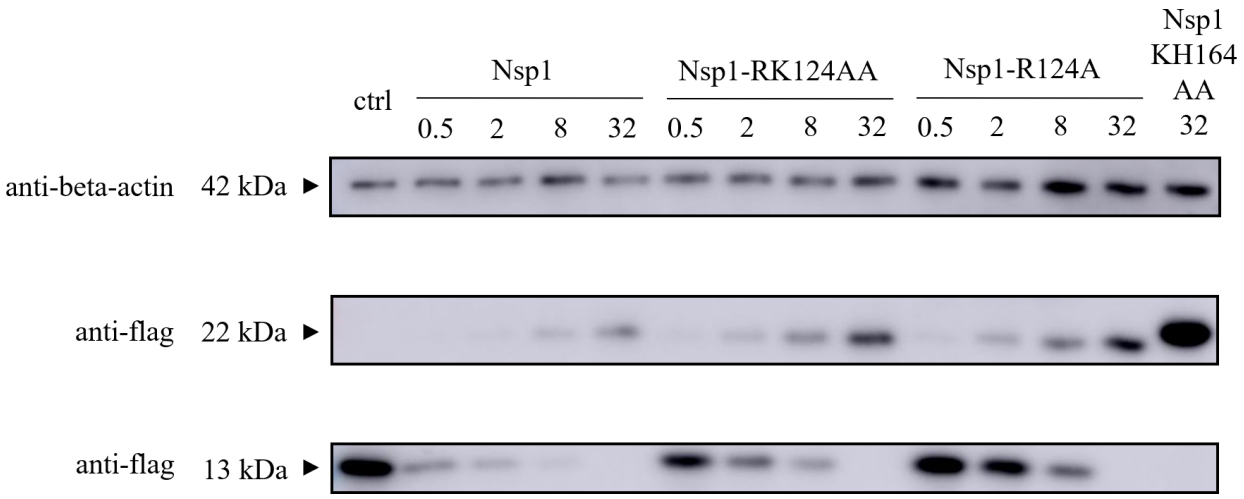

Figure 10. Western blot membrane for pooled reporter assay samples for SARS-CoV-1 Nsp1. Anti-flag antibody is used to visualise flag-tagged Nsp1 and its mutants (22kDa). Expression is not detected for samples with 0.5 and 2 ng Nsp1-expressing plasmid DNA used. For RK124AA and R124A mutants, a band is visible for 2 ng of plasmid DNA used, as well. Band intensity increases with more DNA used for transfection. KH164AA mutant shows a high intensity band indicating high level of expression. Band at 13 kDa corresponds to flag-tagged sic-peptide size which was used as a "filler" plasmid and negative control.

An additional band was visible at 12 kDa for both SARS-CoV-1 and SARS-CoV-2 samples, which corresponds to the size of flag-tagged sic-peptide which was used as a „filler" plasmid to ensure each sample is transfected with an equal amount of total plasmid. Intensity of the band corresponds to the amount of plasmid used, with highest intensity in the negative control samples and gradually decreasing as more Nsp1-expressing plasmid was used.

Beta-actin was used as loading control for all samples and showed bands of the same intensity in each of them.

5. Discussion

The purpose of this thesis was to investigate the effect of SARS-CoV-2 and SARS-CoV-1 Nsp1 protein on cellular and viral mRNA translation in human cells. A reporter-based approach was used to show changes in transcript translation in the presence of Nsp1.

5.1. Addition of upstream ORFs to reporter plasmids

It has previously been shown that the Nsp1 protein in coronaviruses has an important role in inhibiting translation in infected cells (Kamitani et al., 2006, Zust et al., 2007, Narayanan et al., 2008). Data that was generated in this thesis essentially supports these findings suggesting that Nsp1 has a strong inhibitory effect on global protein synthesis. As it is one of the first proteins synthesized upon coronavirus infection, translation inhibition seems to be one of the earliest steps in viral takeover of host cell. Both SARS-CoV-1 and SARS-CoV-2 Nsp1 proteins displayed highly similar effects on reporter expression suggesting conservation of function between the two homologs.

In the initial experiment presented in this thesis, the SARS-CoV-2 leader sequence was cloned upstream of Rluc reporter sequence to represent viral transcripts in infected cells. Rluc reporter sequence without any upstream ORFs represented cellular transcripts in the reporter assay experiment. The results showed that viral and cellular transcripts both exhibited translation inhibition in the presence of Nsp1. That would suggest is that Nsp1 affects both viral and cellular transcripts in the same way, meaning that it causes repression of translation at a global level.

However, these results were not in accordance with available data on Nsp1, particularly extensive research that was done on SARS-CoV-1 Nsp1 which shows that the protein selectively allows translation of viral transcripts (Narayanan et al., 2008, Lokugamage et al., 2015). Since there is high sequence conservation for Nsp1 between SARS-CoV-1 and SARS-CoV-2, it was expected that the protein of the latter virus would have a similar effect.

Due to this issue, a closer look at the leader sequence was necessary. Available research pointed towards secondary structures of the leader sequence as the key reason for translation selectivity (Tanaka et al., 2012). Taking that into account, a potential explanation for the experiment results was that due to the nature of the plasmid construct, the nucleotides of the CMV promoter could potentially interfere with the leader sequence obtaining the proper

secondary structure, even though the initial constructs were generated with that in mind and all nucleotides between the CMV promoter and the leader sequence were removed. However, based on newly published research and the conflicting results that were obtained in the experiment, it became clear that nucleotides within the CMV promoter can still affect the stem loop of the mRNA. The published paper showed that the addition of nucleotides upstream of the leader sequence resulted in ablation of protection from Nsp1-mediated translation inhibition (Banerjee et al., 2020). For this reason, additional 8 nucleotides were deleted from the CMV promoter in order for the 5' UTR to obtain its physiological secondary structures.

When the experiment was repeated using Rluc reporters with 8 nucleotides deleted from the CMV promoter, the results showed that Nsp1 in fact increases translation of viral transcripts while translation of cellular transcripts is repressed. The selectivity in translation initiation explains the importance of Nsp1 for viral infection and shows that the protein is essential in suppressing host defenses by not allowing infected cells to initiate synthesis of cellular proteins, which includes proteins involved in antiviral signaling pathways.

Additionally, the results suggest that Nsp1 loses selectivity for viral transcripts at higher concentrations and begins blocking translation of them as well, which indicates that Nsp1 works best at an optimal concentration in infected cells. It is also worth nothing that the samples in which Nsp1 displayed the strongest effect on viral mRNA translation selectivity, it was not detected on protein level in Western blot, both for SARS-CoV-1 and SARS-CoV-2, further proving what a great impact the protein has on translation even in particularly small amounts. Additionally, it suggests that during viral infection, the protein does not have to be expressed at a high level to suppress host defenses.

5.2. Effects of Nsp1 mutants

In the experiments of this thesis, Nsp1 protein with the RK124AA mutation seems to lose its ability to promote translation of viral transcripts exclusively. In fact, the mutant protein appears to affect viral and non-viral transcripts the same way, causing a decrease in reporter protein expression. Interestingly, the R124A mutant protein shows the same effect, which indicates that even a single nucleic acid mutation in the site is sufficient for Nsp1 to lose its selectivity in allowing translation of viral transcripts. Since increasing concentrations of plasmids expressing either of the mutant causes a decrease in reporter expression, it is evident that the

mutation does not cause a complete loss-of-function of the protein. It still maintains its inhibitory effect on translation (since the KH site is still active), only it now inhibits translation of viral mRNAs, as well.

Considering that the RK-site is proven to be important for endonucleolytic cleavage of host protein mRNA, it could mean that when it is mutated, cellular mRNAs become more abundant and it reduces the likelihood of Nsp1-mediated recognition of viral mRNAs meaning that the translation initiation rate for them decreases. Some SARS-CoV-1 research points to the RK-site as an important recognition site for viral mRNAs, more specifically that it is included in binding to the stem loop 1 structure in the viral leader sequence (Tanaka, 2012). Since all viral mRNAs have the leader sequence as an integral part of their structure, it appears that a possible explanation for the reduction in translation efficiency could be that the mRNAs are not bound to the RK124AA-Nsp1-40S subunit complex as efficiently. Due to the fact that SARS-CoV-1 and SARS-CoV-2 Nsp1 proteins share 84.4% sequence identity and have both shown the same effect in the experiment, this could also indicate that SARS-CoV-2 Nsp1 has an ability to bind the stem loop 1 structure through its RK-site.

The Nsp1-KH164AA mutant has not caused translation inhibition in any of the experiments in this thesis. In fact, regardless of transcript type, the protein appears to lose its ability to inhibit translation as the RLU values have returned to control levels (or higher) in all experiments. What is interesting is that the mutant has shown higher protein expression level in Western blot compared to wild-type Nsp1 as well as RK124AA and R124A mutants. This increase in expression could explain the higher RLU values in the KH164AA mutant samples. Altogether, this confirms that the KH164AA mutation causes loss of function for Nsp1 and provides insight into SARS-CoV-2 Nsp1 mechanism of host defense suppression.

The KH site is shown to be essential for the ribosome binding ability of both SARS-CoV-1 and SARS-CoV-2 Nsp1 proteins. However, according to available data, SARS-CoV-1 Nsp1 does not seem to associate with the 80S ribosome, suggesting that it exclusively binds to the 40S ribosomal subunit in the pre-initiation stage and releases after 40S and 60S subunits join together (Kamitani, 2009). On the other hand, SARS-CoV-2 Nsp1 has been shown to interact with 80S ribosome in addition to the 40S subunit. Furthermore, an *in vitro* assay has shown that SARS-CoV-2 Nsp1 does not associate with the 60S subunit (Thoms et al., 2020, Schubert et al., 2020).

This would indicate that SARS-CoV-2 Nsp1 binds to the 40S subunit, but does not release after subunit joininig, setting it appart from its SARS-CoV-1 homolog.

5.3. Recognition of viral mRNAs and future perspectives

It is evident that the SARS-CoV-2 leader sequence is able to protect viral mRNAs from host translation shutoff and this effect has previously been determined for SARS-CoV-1 leader sequence (Tanaka et al., 2012). However, in the experiments of this thesis SARS-CoV-1 Nsp1 was used in combination with SARS-CoV-2 leader sequence and it was shown that both proteins inhibited translation of cellular transcripts, while allowing translation of viral mRNAs. This indicates that not only is there a great deal of conservation of function between the two proteins, but that the leader sequence is conserved, as well.

In fact, a recent paper suggests that Nsp1 suppresses translation of endogenous mRNA by blocking the mRNA entry channel with its C-terminal domain, while viral mRNA is protected from this Nsp1-mediated translation inhibition. Additionally, they dissected the SARS-CoV-2 leader sequence and shown that the first stem loop of the leader sequence is sufficient to prevent translation suppresion by Nsp1 (Banerjee et al., 2020). This is in accordance with the data obtained in this thesis, where additional nucleotide sequences from the CMV promoter were evidently sufficient to disturb the protection that viral mRNA has against Nsp1 and it resulted in translation suppression of the reporter protein. An explanation of that could be that Nsp1 did not recognize the reporter mRNA as a viral mRNA in that case since it lacked the proper stem loop structure.

Another recently published paper has shown that Nsp1 does not have an affinity for the SL1 structure unless the protein is bound to the 40S subunit. In addition, they have found that the SARS-CoV-2 5' UTR does not bind to the ribosomal subunit on its own, only in the presence of Nsp1. In the paper, they propose a model for viral evasion of Nsp1-mediated translation repression in which Nsp1 acts as a gatekeeper in the sense that it blocks the mRNA entry channel until SL1 and 40S-Nsp1 interaction is established (Figure 11). They also suggest that Nsp1 remains bound to the ribosome during translation and when the entire mRNA is scanned, the C-terminal domain folds back into the mRNA entry channel to prevent translation initiation of non-viral mRNAs (Tidu et al., 2021). This is contrary to the proposed model from the previously

mentioned publication, where it was suggested that Nsp1 is removed from the ribosome upon translation initiation (Banerjee et al., 2020).

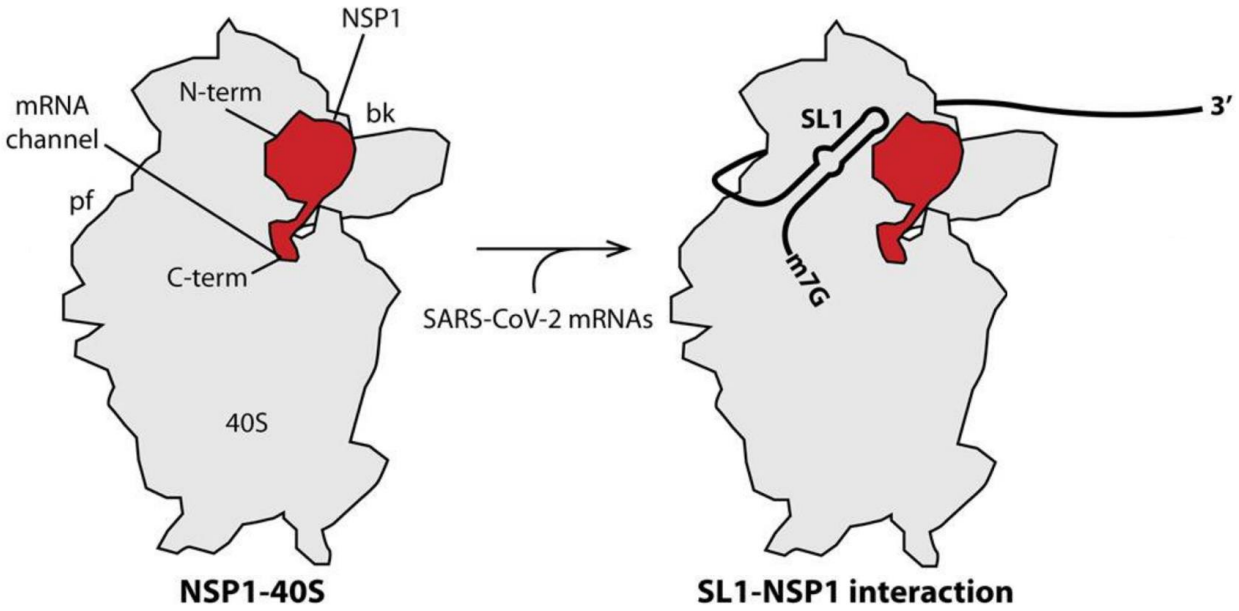

Figure 11. A schematic overview of a proposed model for Nsp1-40S recognition of a SL1 structure in a viral mRNA. Nsp1 binds to the 40S ribosomal subunit blocking the mRNA channel with its C-terminus and causing translation repression. Viral mRNAs contain a stem loop (SL1) structure at their 5' end which is able to interact with Nsp1 enabling entry into the mRNA channel and initiating translation of viral mRNAs exclusively (Tidu et al., 2021).

Despite all the recent breakthroughs, many gaps still remain in the complete mechanism of Nsp1-mediated shutdown of host protein translation. What is evident is that an mRNA requires precise positioning of the 5' leader for translation initiation to be enabled by the Nsp1-40S ribosome complex. Together, these results display the importance of viral mRNA secondary structure formation for successful SARS-CoV-2 infection and a need to better understand the highly coordinated process of Nsp1 translation regulation.

6. Conclusion

The initial goal of this thesis was to determine how SARS-CoV-2 protein Nsp1 affects translation upon infection of human cells. What this reporter-based approach showed was that Nsp1 strongly inhibits translation of cellular transcripts, while simultaneously allowing translation initiation of viral transcripts.

Nsp1 functionality is dependent on its KH-site, which is known for enabling Nsp1 to bind to the 40S ribosomal subunit. When the site is mutated, Nsp1 does not have an inhibitory effect on translation of either viral or cellular transcripts. On the other hand, the RK-site in Nsp1 appears to be important for the selection process, in which Nsp1 differentiates between viral and cellular transcripts and allows viral transcripts to be translated. Additionally, it was shown that it is enough for the arginine in the RK-site to be mutated for the site to lose functionality and the protein to treat viral transcripts the same as cellular.

When the effects of SARS-CoV-2 Nsp1 are compared to the effects of its SARS-CoV-1 homolog, it is clear that there is a great degree of conservation of function between them, although there are mechanistic differences in their activity. This conservation confirms the importance of the protein for viral infection and puts emphasis on determining its mechanism of action as a way of potential development of therapeutics.

An additional finding of this thesis is that secondary structures in the 5' UTR of viral mRNA protect the mRNA from Nsp1-mediated translation inhibition and degradation. The exact mechanism of this process is not completely known yet and it remains unclear at what point Nsp1 dissociates from the 80S ribosome after translation initiation, or in what order the Nsp1-80S ribosome-viral mRNA complex is formed. Many questions about translation disruption during SARS-CoV-2 infection remain unanswered and should continue to be an object of interest in further research.

7. References

Allard, S. T. M. 2008. Bioluminescent reporter genes. worldwide.promega.com/resources/pubhub/enotes/bioluminescent-reporter-genes (Accessed May 29th, 2021)

Allard, S. T. M., Kopish, K. 2008. Luciferase reporter assays: Powerful, adaptable tools for cell biology research. Cell Notes 21, 23-26.

Andersen, K. G., Rambaut, A., Lipkin, W. I., Holmes, E. C., Garry, R. F. 2020. The proximal origin of SARS-CoV-2. Nature Medicine 26, 450-452.

Banerjee, A.K., Blanco, M.R., Bruce, E.A., Honson, D.D., Chen, L.M., Chow, A., Bhat, P., Ollikainen, N., Quinodoz, S.A., Loney, C., Thai, J., Miller, Z.D., Lin, A.E., Schmidt, M.M., Stewart, D.G., Goldfarb, D., De Lorenzo, G., Rihn, S.J., Voorhees, R.M., Botten, J.W., Majumdar, D., Guttman, M. 2020. SARS-CoV-2 disrupts splicing, translation, and protein trafficking to suppress host defenses. Cell 183, 1325-1339.

Bar-On, Y.M., Flamholz, A., Phillips, R., Milo, R. 2020. SARS-CoV-2 (COVID-19) by the numbers. Elife 9, e57309.

Calvo, S. E., Pagliarini, D. J., Mootha, V. K. 2009. Upstream open reading frames cause widespread reduction of protein expression and are polymorphic among humans. Proceedings of the National Academy of Sciences 106, 7507-7512.

Carter, M., Shieh, J. 2015. Chapter 15 - Biochemical assays and intracellular signaling. In Carter, M., Shieh, J. (editors) Guide to research techniques in neuroscience (Second edition). Cambridge, Academic Press, pages 311-343.

Coronaviridae Study Group of the International Committee on Taxonomy of Viruses: Gorbalenya, A.E., Baker, S.C., Baric, R.S., de Groot, R.J., Drosten, C., Gulyaeva, A.A., Haagmans, B.L., Lauber, C., Leontovich, A.M., Neuman, B.W., Penzar, D., Perlman, S., Poon, L.L.M., Samborskiy, D.V., Sidorov, I.A., Sola, I., Ziebuhr, J. 2020. The species Severe acute respiratory syndrome-related coronavirus: classifying 2019-nCoV and naming it SARS-CoV-2. Nature Microbiology 5, 536–544.

covid19.who.int (accessed May 25th, 2021)

da Silva, S.J.R., da Silva, C.T.A., Mendes, R.P.G., Pena, L. 2020. Role of nonstructural proteins in the pathogenesis of SARS-CoV-2. Journal of Medical Virology 92, 1427-1429.

de Breyne, S., Vindry, C., Guillin, O., Condé, L., Mure, F., Gruffat, H., Chavatte, L., Ohlmann, T. 2020. Translational control of coronaviruses. Nucleic Acids Research 48, 12502–12522.

Drosten, C., Günther, S., Preiser, W., van der Werf, S., Brodt, H.R., Becker, S., Rabenau, H., Panning, M., Kolesnikova, L., Fouchier, R.A., Berger, A., Burguière, A.M., Cinatl, J., Eickmann, M., Escriou, N., Grywna, K., Kramme, S., Manuguerra, J.C., Müller, S., Rickerts, V., Stürmer, M., Vieth, S., Klenk, H.D., Osterhaus, A.D., Schmitz, H., Doerr, H.W. 2003. Identification of a novel coronavirus in patients with severe acute respiratory syndrome. New England Journal of Medicine 348, 1967-1976.

Finkel, Y., Mizrahi, O., Nachshon, A., Weingarten-Gabbay, S., Morgenstern, D., Yahalom-Ronen, Y., Tamir, H., Achdout, H., Stein, D., Israeli, O., Beth-Din, A., Melamed, S., Weiss, S., Israely, T., Paran, N., Schwartz, M., Stern-Ginossar, N. 2021. The coding capacity of SARS-CoV-2. Nature 589, 125-130.

Graham, R.L., Baric, R.S. 2009. Recombination, reservoirs, and the modular spike: mechanisms of coronavirus cross-species transmission. Journal of Virology 84, 3134-3146.

Huang, C., Lokugamage, K.G., Rozovics, J.M., Narayanan, K., Semler, B.L., Makino, S. 2021. SARS coronavirus Nsp1 protein induces template-dependent endonucleolytic cleavage of mRNAs: viral mRNAs are resistant to Nsp1-induced RNA cleavage. PLoS Pathogens 7, e1002433.

Huang, C., Wang, Y., Li, X., Ren, L., Zhao, J., Hu, Y., Zhang, L., Fan, G., Xu, J., Gu, X., Cheng, Z., Yu, T., Xia, J., Wei, Y., Wu, W., Xie, X., Yin, W., Li, H., Liu, M., Xiao, Y., Gao, H., Guo, L., Xie, J., Wang, G., Jiang, R., Gao, Z., Jin, Q., Wang, J., Cao, B. 2020. Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. Lancet 395, 497-506.

Kamitani, W., Narayanan, K., Huang, C., Lokugamage, K., Ikegami, T., Ito, N., Kubo, H., Makino, S. 2006. Severe acute respiratory syndrome coronavirus Nsp1 protein suppresses host gene expression by promoting host mRNA degradation. Proceedings of the National Academy of Sciences of the USA 103. 12885-12890.

Kamitani, W., Huang, C., Narayanan, K., Lokugamage, K. G., Makino, S. 2009. A two-pronged strategy to suppress host protein synthesis by SARS coronavirus Nsp1 protein. Nature Structural & Molecular Biology 16, 1134-1140.

Kim, D., Lee, J.Y., Yang, J.S., Kim, J.W., Kim, V.N., Chang, H. 2020. The architecture of SARS-CoV-2 transcriptome. Cell 181, 914-921.

Lim, Y.X., Ng, Y.L., Tam, J.P., Liu, D.X. 2016. Human Coronaviruses: A review of virus–host interactions. Diseases 4, 26.

Lokugamage, K.G., Narayanan, K., Huang, C., Makino, S. 2012. Severe acute respiratory syndrome coronavirus protein Nsp1 is a novel eukaryotic translation inhibitor that represses multiple steps of translation initiation. Journal of Virology 86, 13598-13608.

Nakagawa, K., Lokugamage, K. G., Makino, S., Denison, M. R. 2016. Viral and cellular mRNA translation in coronavirus-infected cells. Advanced Virus Research 96, 2305-2315.

Narayanan, K., Huang, C., Lokugamage, K., Kamitani, W., Ikegami, T., Tseng, C.T., Makino, S. 2008. Severe acute respiratory syndrome coronavirus Nsp1 suppresses host gene expression, including that of type I interferon, in infected cells. Journal of Virology 82, 4471-4479.

Satarker, S., Nampoothiri, M. 2020. Structural proteins in Severe acute respiratory syndrome coronavirus-2. Archives of Medical Research 51, 482-491.

Schubert, K., Karousis, E.D., Jomaa, A., Scaiola, A., Echeverria, B., Gurzeler, L.A., Leibundgut, M., Thiel, V., Mühlemann, O., Ban, N. 2020. SARS-CoV-2 Nsp1 binds the ribosomal mRNA channel to inhibit translation. Nature Structural & Molecular Biology 27, 959-966.

Sola, I., Almazan, F., Zuniga, S., Enjuanes, L. 2015. Continuous and discontinuous RNA synthesis in coronaviruses. Annual Review of Virology 2, 265-288.

Tanaka, T., Kamitani, W., DeDiego, M. L., Enjuanes, L., Matsuura, Y. (2012). Severe acute respiratory syndrome coronavirus Nsp1 facilitates efficient propagation in cells through a specific translational shutoff of host mRNA. Journal of Virology, 86, 11128-11137.

Thoms, M., Buschauer, R., Ameismeier, M., Koepke, L., Denk, T., Hirschenberger, M., Kratzat, H., Hayn, M., Mackens-Kiani, T., Cheng, J., Straub, J.H., Stürzel, C.M., Fröhlich, T.,

Berninghausen, O., Becker, T., Kirchhoff, F., Sparrer, K.M.J., Beckmann, R. 2020. Structural basis for translational shutdown and immune evasion by the Nsp1 protein of SARS-CoV-2. Science 369, 1249-1255.

Tidu, A., Janvier, A., Schaeffer, L., Sosnowski, P., Kuhn, L., Hammann, P., Westhof, E., Eriani, G., & Martin, F. 2020. The viral protein Nsp1 acts as a ribosome gatekeeper for shutting down host translation and fostering SARS-CoV-2 translation. RNA 27, 253–264.

Yang, D., Leibowitz, J.L. 2015. The structure and functions of coronavirus genomic 3' and 5' ends. Virus Research 206, 120-133.

Yang, Y., Hussain, S., Wang, H., Ke, M., Guo, D. 2009. Translational control of the subgenomic RNAs of severe acute respiratory syndrome coronavirus. Virus Genes 39, 10-18.

Zou, L., Ruan, F., Huang, M., Liang, L., Huang, H., Hong, Z., Yu, J., Kang, M., Song, Y., Xia, J., Guo, Q., Song, T., He, J., Yen, H.L., Peiris, M., Wu, J. 2020. SARS-CoV-2 viral load in upper respiratory specimens of infected patients. New England Journal of Medicine 382, 1177-1179.

Zust, R., Cervantes-Barragan, L., Kuri, T., Blakqori, G., Weber, F., Ludewig, B. and Thiel, V., 2007. Coronavirus non-structural protein 1 is a major pathogenicity factor: implications for the rational design of coronavirus vaccines. PLoS Pathogens 3, 1062-1073.

Curriculum Vitae – Lucija Bujanić

From 2011 to 2015 I attended the high school "Gymnasium Josip Slavenski" in Čakovec. From 2015 to 2018 I attended the undergraduate programme in Molecular Biology at the same university and obtained my bachelor's degree. From 2018 to 2021 I attended the graduate programme in Molecular Biology at the Faculty of Science at University of Zagreb.

I have achieved the following work experience. From October 2016 to February 2017 I was a laboratory demonstrator on the course "Zoology", under supervision from dr. sc. Mladen Kučinić at the Faculty of Science at University of Zagreb. From July 2018 to September 2019 I was a summer research intern working on the topic "Transcriptional and translational analysis of the *fat2* gene in *Drosophila* imaginal discs" in the lab of dr. sc. Christian Dahmann at the Technical University Dresden. From July 2019 to September 2019 I was a summer research intern in the lab of dr. sc. Marin Barišić, working on the topic "Generation of fluorescent knock-in HeLa cell lines using CRISPR-Cas9" at the Danish Cancer Society Research Center in Copenhagen. From October 2018 to Septemeber 2019 I was a research intern in the lab of dr. sc. Iva Tolić, working on the topic: "Mitotic spindle helicity reaches its peak at the start of anaphase" at the Ruđer Bošković Institute in Zagreb. From March 2020 to January 2021 I was a student assistant (HiWi) and research intern in the lab of dr. sc. Marina Chekulaeva at the Berlin Institute for Medical Systems Biology (part of Max-Delbrück Center for Molecular Medicine in Berlin) working on the topic "The effects of SARS-CoV-2 protein Nsp1 on translation in human cells".

In 2016 I organized a workshop (swapshop) for high school students at the Summer School of Science in Požega, Croatia. In 2016 and 2017 I organized workshops for children at university open days at the Faculty of Science at the University of Zagreb. In 2019 I had a poster presentation titled "Mitotic spindle helicity reaches its peak tat the start of anaphase" at a conference called: "Mitotic spindle: From living and synthetic systems to theory" in Split, Croatia.

I was awarded the national STEM scholarship in 2017 and 2018 and the Erasmus+ traineeship scholarship in 2018, 2019 and 2020.