Report purification of methyltransferase

Introduction

mRNA cap structures are currently receiving immense attention. In bacteria, new cap structures have been discovered that were thought non-existent. In eukaryotes, more and more enzymes involved in the synthesis of cap structures are being discovered and their role in translation initiation and innate immunity is being characterized. Technologies for analytics of cap structures, in particular by LC-MS, have made significant progress, and thus, the content of mRNA populations in terms of different cap structures, as well as their dynamics under different stress conditions are now accessible to analysis.

Most recently, synthetic mRNA in its capacity of antiviral vaccine has gained immense public attention. Translation efficiency and immunogenicity are key aspects to be investigated, which hold significant stakes in public health. Capping of these mRNAs is key to their stability *in vivo* and thus key to their efficiency as vaccines. We aim to use enzymatic synthesis of new cap-analogues for initiation of transcription by T7 RNA polymerase to prepare stable mRNAs for further applications.

We chose the NSP1 protein in Venezuelan equine encephalitis virus (VEEV) which plays import roles in capping single stranded RNA as it bears not only guanylyltransferase but more importantly methyltransferase activity (Ortega Granda *et al.,* 2021).

Designing

It was decided to order original plasmid pUC57 with VEEV NSP1 gene from GenScript. After a codon optimization of its sequence for expressing in bacteria cells, confirmation of localizations of NSP1 gene in the original vector was done by Sanger sequencing using the pUC57 specific oligonucleotides. Gibson assembly method was used to subclone NSP1 gene into pET28a vector with double His-tag, lipoyl domain and TEV-cleavage site for subsequent purification.

Oligonucleotides for pUC57 sequencing:

Forward: 5´-GTA AAA CGA CGG CCA GT-3´

Reverse: 5´-GGA AAC AGC TAT GAC CAT G-3´

Forward oligonucleotide for pUC57 recognition outside MCS for NSP1 confirmation:

5´-CGG CAT CAG AGC AGA TTG TAC-3´

> Original plasmid puc57 with NSP1 from GenScript (YP1015)

TCGCGCGTTTCGGTGATGACGGTGAAAACCTCTGACACATGCAGCTCCCGGAGACGGTCACAGCTTGTCTGTAA
GCGGATGCCGGGAGCAGACAAGCCCGTCAGGGCGCGTCAGCGGGTGTTGGCGGGTGTCGGGGCTGGCTTAACT
ATGCGGCATCAGAGCAGATTGTACTGAGAGTGCACCATATGGAAAAAGTTCACGTGGATATTGAAGAGGATAGC
CCGTTTCTGCGTGCACTGCAGCGTAGCTTTCCGCAGTTTGAAGTGGAAGCCAAACAGGTTACCGACAACGATCA
CGCCAATGCACGCGCCTTTAGCCACCTGGCCAGCAAACTGATTGAAACCGAAGTTGATCCGAGCGATACCATTC
TGGATATTGGCAGCGCCCCGGCGCGCCGCATGTATAGCAAACATAAATATCACTGCATTTGCCCGATGCGCTGTGC
GGAAGATCCGGACCGTCTGTATAAATATGCGACCAAACTGAAAAAAAACTGTAAAGAAATTACCGATAAAGAAC
TGGATAAAAAAATGAAAGAACTGGCCGCCGTTATGAGCGATCCGGACCTTGAAACCGAAACCATGTGCCTGCAT
GATGATGAAAGCTGCCGCTACGAAGGCCAGGTGGCCGTGTACCAGGATGTGTATGCCGTGGATGGTCCGACCAG
CCTGTATCATCAGGCCAATAAAGGCGTGCGTGTGGCCTACTGGATTGGCTTTGATACCACCCCGTTTATGTTCAAA
AATCTGGCGGGCGCGTATCCGAGCTACAGCACGAACTGGGCCGATGAAACCGTTCTGACCGCGCGCAACATTGG
CCTGTGCAGCAGCGATGTTATGGAGCGCAGTCGTCGCGGCATGAGCATTCTGCGCAAAAAATATCTGAAACCGA
GCAACAACGTTCTGTTTAGCGTGGGCAGCACCATTTACCACGAAAAACGCGATCTGCTGCGCAGCTGGCATCTG
CCGAGCGTCTTTCATCTGCGCGGTAAACAGAACTACACCTGCCGCTGTGAAACCATTGTGAGCTGTGATGGCTAT
GTGGTAAAACGCATCGCGATTTCACCGGGCCTGTACGGCAAACCGAGCGGCTACGCCGCGACCATGCATCGCGA
AGGATTTCTGTGCTGTAAAGTGACCGATACGCTGAACGGCGAACGCGTGAGCTTTCCGGTGTGTACCTACGTGC
CGGCGACCCTGTGCGATCAGATGACTGGCATTCTGGCGACCGATGTGAGCGCGGATGATGCGCAGAAACTGTTG
GTGGGCCTGAACCAGCGCATTGTGGTGAATGGCCGCACCCAGCGTAACACCAATACCATGAAAAACTACCTGCT
GCCGGTGGTTGCGCAGGCCTTTGCGCGCTGGGCCAAAGAATACAAAGAAGATCAGGAGGATGAACGCCCGCTG
GGCCTGCGCGATCGCCAGCTGGTGATGGGATGCTGCTGGGCGTTTCGCCGTCATAAAATTACCAGCATTTACAAA
CGCCCGGATACCCAGACCATTATTAAAGTTAACAGCGATTTCCATAGCTTTGTTCTGCCGCGTATTGGCAGCAACA
CCCTCGAAATTGGTCTGCGCACCCGCATTCGCAAAATGCTGGAAGAACACAAAGAACCGAGCCCGCTGATTACC
GCGGAAGATGTGCAGGAAGCGAAATGTGCCGCCGATGAAGCGAAAGAAGTGCGCGAAGCGGAAGAACTGCGC
GCGGCGCTGCCTCCGCTGGCAGCGGATGTGGAAGAACCGACCCTGGAAGCCGACGTGGACCTGATGCTGCAGG
AAGCGGGCGCGTAAGCTTGCGGCCGCACTCGAGGGCGTAATCATGGTCATAGCTGTTTCCTGTGTGAAATTGTTA
TCCGCTCACAATTCCACACAACATACGAGCCGGAAGCATAAAGTGTAAAGCCTGGGGTGCCTAATGAGTGAGCT
AACTCACATTAATTGCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCGTGCCAGCTGCATTAATGAA
TCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCTCTTCCGCTTCCTCGCTCACTGACTCGCTGCGC
TCGGTCGTTCGGCTGCGGCGAGCGGTATCAGCTCACTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGA
TAACGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGC
GTTTTTCCATAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGA
CAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTA
CCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTC
GGTGTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCG
GTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATT
AGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGAAC
AGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACA
AACCACCGCTGGTAGCGGTGGTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAG
ATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGTCATGAGAT
TATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAAATCAATCTAAAGTATATATGAGTAA
ACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTCGTTCATCCATAG
TTGCCTGACTCCCCGTCGTGTAGATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATAC
CGCGAGACCCACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAGCGCAGAAG
TGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCTAGAGTAAGTAGTTCGCCAGTT
AATAGTTTGCGCAACGTTGTTGCCATTGCTACAGGCATCGTGGTGTCACGCTCGTCGTTTGGTATGGCTTCATTCA
GCTCCGGTTCCCAACGATCAAGGCGAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTC
CTCCGATCGTTGTCAGAAGTAAGTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCATAATTCTCTTAC
TGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTCATTCTGAGAATAGTGTATGCGG
CGACCGAGTTGCTCTTGCCCGGCGTCAATACGGGATAATACCGCGCCACATAGCAGAACTTTAAAAGTGCTCATC
ATTGGAAAACGTTCTTCGGGGCGAAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACT
CGTGCACCCAACTGATCTTCAGCATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAA
TGCCGCAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATACTCTTCCTTTTTCAATATTATTGAAG
CATTTATCAGGGTTATTGTCTCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCG
CGCACATTTCCCCGAAAAGTGCCACCTGACGTCTAAGAAACCATTATTATCATGACATTAACCTATAAAAATAGGC
GTATCACGAGGCCCTTTCGTC

NdeI restricton site: CATATG

XhoI restricton site: CTCGAG

>final subcloned plasmid pET28a (YP691) + NSP1 insert

ATTTTTCGTCTAGAAATAATTTTGTTTAACTTTAAGAAGGAGATATACCATGGGCAGCAGCCATCATCATCMTCWT
CACTCTCGTGCTTGGCGTCATCCTCAATTTGGTGGTCATCATCATCATCATCACAGCGGCGCTTTTGAATTTAAGC
TGCCGGACATTGGCGAAGGCATCCACGAAGGTGAAATTGTCAAATGGTTTGTGAAACCGGGCGATGAAGTGAA
CGAAGACGATGTATTGTGCGAAGTGCAAAATGACAAGGCGGTTGTCGAAATTCCCTCCCCGGTCAAAGGGAAA
GTGCTTGAAATCCTCGTCCCGGAGGGAACAGTGGCAACGGTCGGGCAAACGCTCATCACGCTCGATGCGCCGG
GTTATGAAAACATGACGACCGGCAGCGACACCGGCGAAAACCTGTACTTCCAGGGTAGCCATATGGAAAAAGTT
CACGTGGATATTGAAGAGGATAGCCCGTTTCTGCGTGCACTGCAGCGTASCTTTCCGCAGTTTGAAGTGGAAGCC
AAACAGGTTACCGACAACGATCACGCCAATGCACGCGCCTTTAGCCACCTGGCCARCAAACTGATTGAAACCGA
AGTTGATCCGARCGATACCATTCTGGATATTGGCAGCGCCCCGGCGCGCCGCATGTATAGCAAACATAAATATCAC
TGCATTTGCCCGATGCGCTGTGCGGAAGATCCGGACCGTCTGTATAAATATGCGACCAAACTGAAAAAAAACTG
TAAAGAAATTACCGATAAAGAACTGGATAAAAAAATGAAAGAACTGGCCGCCGTTATGAGCGATCCGGACCTTG
AAACCGAAACCATGTGCCTGCATGATGATGAAAGCTGCCGCTACGAAGGCCAGGTGGCCGTGTACCAGGATGTG
TATGCCGTGGATGGTCCGACCAGCCTGTATCATCAGGCCAATAAAGGCGTGCGTGTGGCCTACTGGATTGGCTTT
GATACCACCCCGTTTATGTTCAAAAATCTGGCGGGCGCGTATCCGAGCTACAGCACGAACTGGGCCGATGAAAC
CGTTCTGACCGCGCGCAACATTGGCCTGTGCAGCAGCGATGTTATGGAGCGCAGTCGTCGCGGCATGAGCATTC
TGCGCAAAAAATATCTGAAACCGAGCAACAACGTTCTGTTTAGCGTGGGCAGCACCATTTACCACGAAAAACGC
GATCTGCTGCGCAGCTGGCATCTGCCGAGCGTCTTTCATCTGCGCGGTAAACAGAACTACACCTGCCGCTGTGA
AACCATTGTGAGCTGTGATGGCTATGTGGTAAAACGCATCGCGATTTCACCGGGCCTGTACGGCAAACCGAGCG
GCTACGCCGCGACCATGCATCGCGAAGGATTTCTGTGCTGTAAAGTGACCGATACGCTGAACGGCGAACGCGTG
AGCTTTCCGGTGTGTACCTACGTGCCGGCGACCCTGTGCGATCAGATGACTGGCATTCTGGCGACCGATGTGAGC
GCGGATGATGCGCAGAAACTGTTGGTGGGCCTGAACCAGCGCATTGTGGTGAATGGCCGCACCCAGCGTAACA
CCAATACCATGAAAAACTACCTGCTGCCGGTGGTTGCGCAGGCCTTTGCGCGCTGGGCCAAAGAATACAAAGAA
GATCAGGAGGATGAACGCCCGCTGGGCCTGCGCGATCGCCAGCTGGTGATGGGATGCTGCTGGGCGTTTCGCCG
TCATAAAATTACCAGCATTTACAAACGCCCGGATACCCAGACCATTATTAAAGTTAACAGCGATTTCCATAGCTTT
GTTCTGCCGCGTATTGGCAGCAACACCCTCGAAATTGGTCTGCGCACCCGCATTCGCAAAATGCTGGAAGAACA
CAAAGAACCGAGCCCGCTGATTACCGCGGAAGATGTGCAGGAAGCGAAATGTGCCGCCGATGAAGCGAAAGAA
GTGCGCGAAGCGGAAGAACTGCGCGCGGCGCTGCCTCCGCTGGCAGCGGATGTGGAAGAACCGACCCTGGAA
GCCGACGTGGACCTGATGCTGCAGGAAGCGGGCGCGTAAGCTTGCGGCCGCACTCGAGCACCACCACCACCAC
CACTGAGATCCGGCTGCTAACAAAGCCCGAAAGAGCKKAGKTKSCCC

>protein nsp1 sequence with 2xHis-tag and lipoyl domain and TEV cleavage site

SRRRYTMGSSHHHHHHSRAWRHPQFGGHHHHHHSGAFEFKLPDIGEGIHEGEIVKWFVKPGDEVNEDDVLCEVQN
DKAVVEIPSPVKGKVLEILVPEGTVATVGQTLITLDAPGYENMTTGSDTGENLYFQGSHMEKVHVDIEEDSPFLRALQ
RSFPQFEVEAKQVTDNDHANARAFSHLASKLIETEVDPSDTILDIGSAPARRMYSKHKYHCICPMRCAEDPDRLYKY
ATKLKKNCKEITDKELDKKMKELAAVMSDPDLETETMCLHDDESCRYEGQVAVYQDVYAVDGPTSLYHQANKGVR
VAYWIGFDTTPFMFKNLAGAYPSYSTNWADETVLTARNIGLCSSDVMERSRRGMSILRKKYLKPSNNVLFSVGSTIY
HEKRDLLRSWHLPSVFHLRGKQNYTCRCETIVSCDGYVVKRIAISPGLYGKPSGYAATMHREGFLCCKVTDTLNGE
RVSFPVCTYVPATLCDQMTGILATDVSADDAQKLLVGLNQRIVVNGRTQRNTNTMKNYLLPVVAQAFARWAKEYK
EDQEDERPLGLRDRQLVMGCCWAFRRHKITSIYKRPDTQTIIKVNSDFHSFVLPRIGSNTLEIGLRTRIRKMLEEHKEP
SPLITAEDVQEAKCAADEAKEVREAEELRAALPPLAADVEEPTLEADVDLMLQEAGA*ACGRTR

Cloning

Gibson assembly method was used to subclone NSP1 into pET28a plasmid containg a double His-tag, lipoyl domain and TEV-cleavage site for further purification. The Gibbson assembly is a cloning procedure that allows the cloning of two or more fragments into vector without restriction enzyme digestion. Instead, user-defined overlapping ends of the fragments allow the seamless joining of adjacent fragments. The reaction is carried out under isothermal conditions using three enzymatic activities: a 5’ exonuclease generates long overhangs, a polymerase fills in the gaps of the annealed single-strand regions, and a DNA ligase seals the nicks of the annealed and filled-in gaps (Gibson *et al.*, 2009).

Oligonucleotides used for Gibson assembly:

Forward: 3'-CCTGTACTTCCAGGGTAGCCATATGGAAAAAGTTCACGTGGATATTGAAG-5

Reverse: 5´--cagtggtggtggtggtggtgctcgagTGCGGCCGCAAGCTTACGCGCCCG -3´

Overlap from vector

Overlap from insert

Restriction site

Oligonucleotides specific to NSP1 used for confirmation of succesful cloning by colony PCR:

Forward: 5´-GCC ATA TGG AAA AAG TTC ACG TGG-3´

Reverse: 5´-GCA AGC TTA CGC GCC CGC-3´

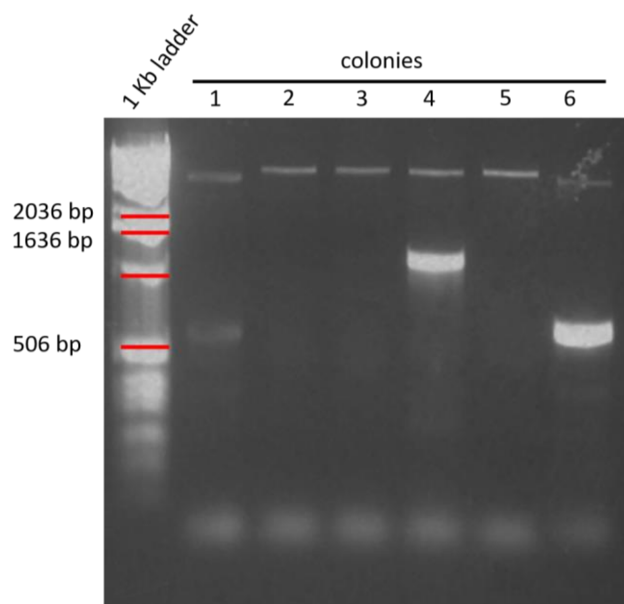

Figure 1: *1% agarose gel analysis of colony PCR after Gibson assembly shows subcloned insert into pet28a plasmid in colony 4 in size aroun 1600 bp.*

Expression

After confirmation of succesfuly subcloned NSP1 by Sanger sequencing, an expression test with T7 expression E.Coli strain RIL was conducted with two post-induction temperatures – 25° and 16° Celsius. Precultures incubated at 16° C exhibited higher expression of the nsp1 protein, prompting further growth of cell cultures and were scaled-up for higher yield.

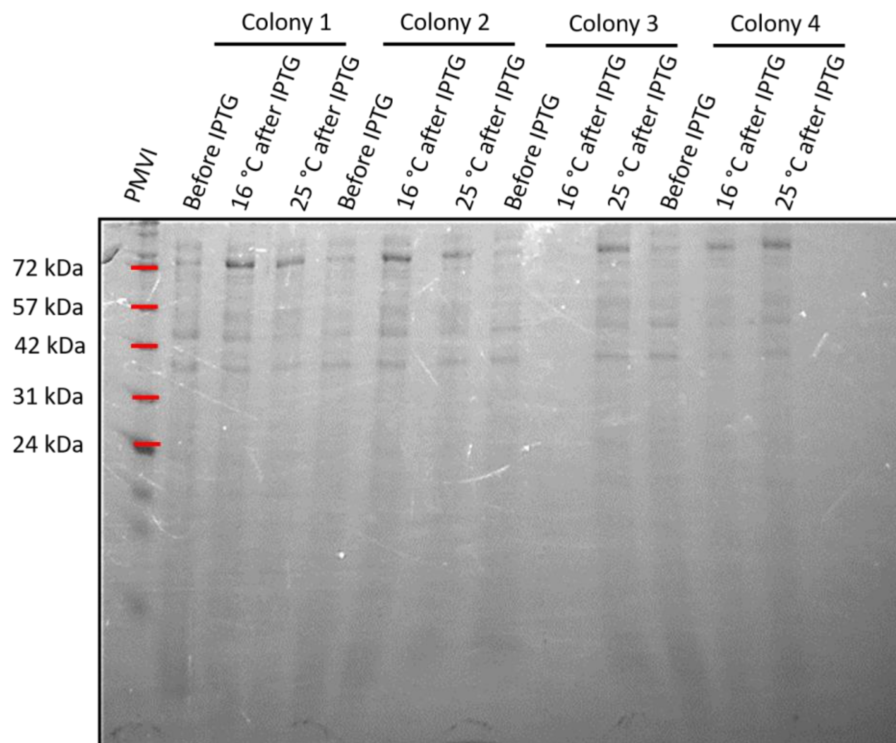

Figure 2: *Expression test of pre-selected colonies showing post-.induction tempertaure 16° C has mostly higher efficiency resulting n higher protein production*

Purification

After bacteria cells lysis with HIS-A buffer (20mM Tris-HCl, pH 7.5, 300mM NaCl, 30 mM imidazole 5% glycerol, and 5mM b-mercaptoethanol) supplemented with protease inhibitors EDTA-free tablet by cell cracking with high pressure, we usid three-step purification procedure to extract the nsp1 protein.

First purification step used Immobilized Metal Ion Affinity Chromatography (IMAC) purification using a 5-ml HisTrap colum to bind the His-tag of the protein with specific buffers (low imidazole buffer: HIS-A; high imidazole buffer: HIS-B with concentration of Imidazole 600mM). Fractions with eluted protein were collected and undergo subsequent cleavage by TEV protease. Second IMAC purification step was performed to ensure additional reduction of all impurities. Cleaved protein was collected in flow-through fractions and proceed further for the final purification step. To further clean purified protein, RNase inhibitor (40U/ul) was added during protein dialysis after second IMAC purification step.

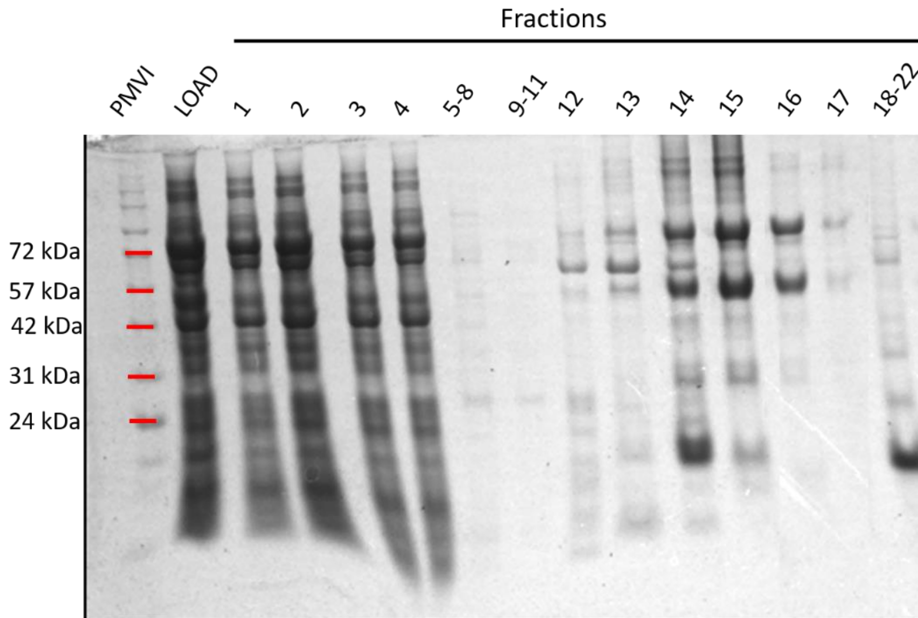

***Figure 3**: Protein gel after first IMAC purification step showing nsp1 protein with TEV cleavage site and lipolyl domain at a size of 76,2 kDa in fractions 12-17.*

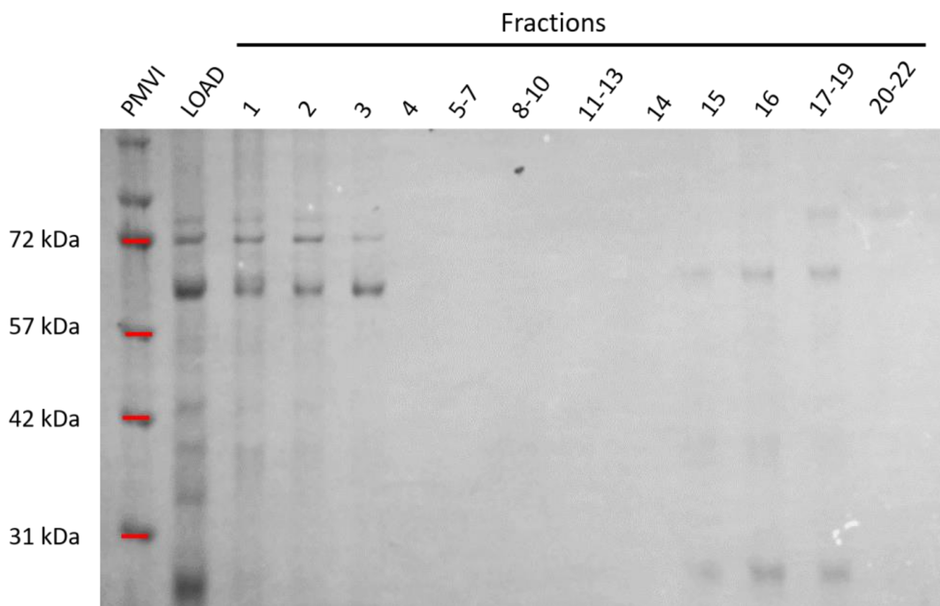

***Figure 4:** Protein gel after TEV cleavage and second IMAC purification step showing nsp1 protein of a size 61,5 kDa in fractions 1-3.*

After dialysis in Q-A buffer (buffer composition), ion exchange chromatography was performed using a 5-ml HiTrap Q-sepharose column (with low-salt buffer Q-A and high-salt buffer Q-B (20mM Tris-HCl, pH 7.5, 150mM/300mM NaCl, 5mM BME) which was chosen based on nsp1 isoelectronic point pI = 6.56.

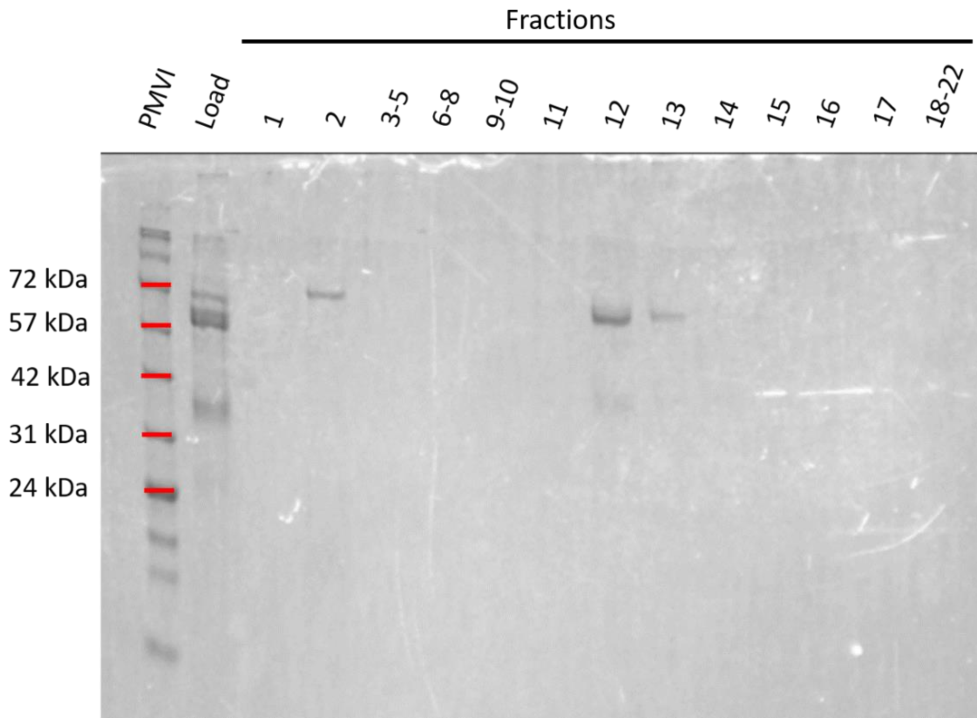

Figure 5*: Protein gel after Q-sepharose purification step showing nsp1 protein in fractions 12 and 13.*

Fractions with eluted NSP1 protein were collected, concentrated and dialysed against storage buffer (20mM Tris-HCl, pH 7.5, 150mM NaCl, and 50 % glycerol). After dialysis, a RNA degradation test performed had negative results. Thus stored protein now is ready for further assays.

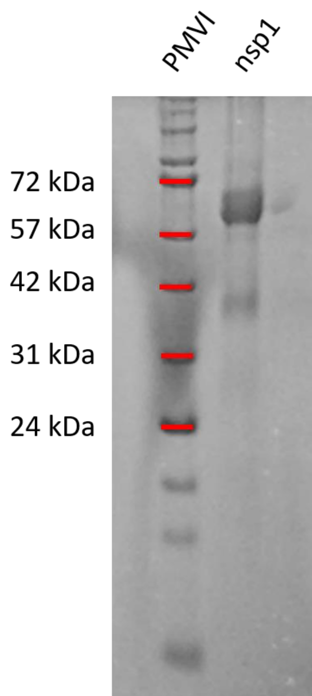

Figure 6: *Protein gel od nsp1 after dialysis in storage buffer.*

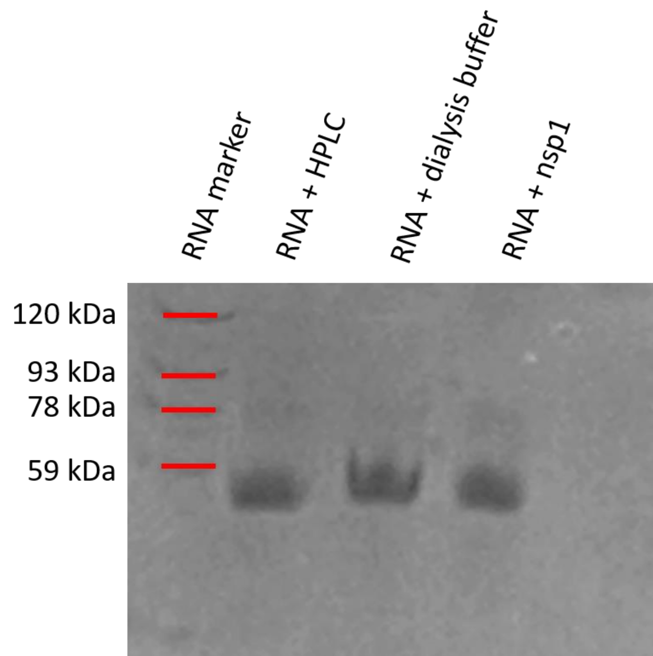

Figure 7: *RNA degradation test indicating no RNase contamination of the sample.*

References:

Ortega Granda O, Valle C, Shannon A, Decroly E, Canard B, Coutard B, Rabah N. *Structure and Sequence Requirements for RNA Capping at the Venezuelan Equine Encephalitis Virus RNA 5' End.* J Virol. 2021 Jul 12;95(15):e0077721. doi: 10.1128/JVI.00777-21. Epub 2021 Jul 12. PMID: 34011549; PMCID: PMC8274618.

Gibson, D., Young, L., Chuang, RY. et al. *Enzymatic assembly of DNA molecules up to several hundred kilobases.* Nat Methods 6, 343–345 (2009). https://doi.org/10.1038/nmeth.1318