Supplementary Information

Leveraging mRNAs sequences to express SARS-CoV-2 antigens in vivo

Chunxi Zeng[1], Xucheng Hou[1], Jingyue Yan[1], Chengxiang Zhang[1], Wenqing Li[1], Weiyu Zhao[1], Shi Du[1], Yizhou Dong[1,2*]

[1] Division of Pharmaceutics & Pharmacology, College of Pharmacy, The Ohio State University, Columbus, OH, USA

[2] Department of Biomedical Engineering, Center for Clinical and Translational Science, Comprehensive Cancer Center, Dorothy M. Davis Heart & Lung Research Institute, Department of Radiation Oncology, The Ohio State University, Columbus, OH, USA

Email: dong.525@osu.edu

Supplementary Table 1. Translation capacity per mRNA molecule of 4248 genes. The genes were sorted based on their translation capacity per mRNA. The 30 genes with the highest translation capacity were list below. The full dataset is available in the Additional Information file. The information and data of Gene Names, protein length [amino acids], average mRNA half-life [h], and average translation rate constant [molecules/(mRNA*h)] are adapted from the literature [1]. Translation capacity per mRNA (total amino acids/mRNA) is calculated using the following equation: protein length (amino acids) × average half-life (h) × translation rate constant (protein copy number/(mRNA×h)).

Gene Names	Protein length [amino acids]	Average mRNA half-life [h]	Average translation rate constant [molecules/(mRNA*h)]	Translation capacity per mRNA (total amino acids/mRNA)
Rps27a	156	10.26	224528.75	359371736
Hist1h1c	212	3.85	147337.15	120256582
H2afv	129	19.97	29040.14	74811176
Rab11b	218	15.3	19290.92	64342935
Dnaja1	397	11.31	12921.59	58018844
Naca	2187	14.15	1869.6	57856735
Hnrnpd	355	22.13	5169.69	40613860
Rpl15	204	9.04	20242.96	37331257
Hmgb2	210	10.11	17120.18	36347854
Rpl31	125	9.86	25219.48	31083009
Ahcy	432	15.84	4414.95	30210973
Tep1	2629	14.78	764.43	29703166
Dync1h1	4644	23.58	270.23	29591677
Serbp1	407	31.92	2155.29	28000321
Ncl;Nuc	707	27.85	1394.12	27450153
Plec1	4691	10.2	548.39	26239474
Psma3	255	12.33	8076.35	25393256
Rpl18a	176	21.5	6465.83	24466701
Tln1	2541	24.99	383.78	24369873
Cse1l	971	11.13	2247.8	24292492
Rps15a	130	5.3	27988.42	19284021
Dpysl2	572	6.05	5327.25	18435481
Myo9a	2631	6.7	1029.41	18146131
Elavl1	326	9.44	5683.23	17489799
Ywhaz	245	11.11	6288.32	17116493
Cad	2225	26.89	273.81	16382121
Flnb	2602	20.05	307.7	16052740
Pdia3	505	25.68	1232.02	15977328
Lmna	665	18.79	1277.63	15964434

Supplementary Table 2. Sequences of 5' UTRs used in this study.

Name	Sequence	Annotation	Source
S27a-45	GGGAGGAAAGCCUCUCUUAAUCGC AUCGGCUGUAUAAGAAAGCCUUUU GAGGCAUUUUUUUUAGUUGAGCAC AUCAUUUCGAGGCCAUUCUGAGGU AAACCGAGAAAAGAGCGUAAAGAA ACCGAGCGAACGAGCAAAUCUGGC ACUGCGUUAGACAGCCGCGAUUCCG CUGCAGCGCGCAGGCACGUGUGUG GCCGCCUAAGGGGCGGGUCCUUCGG CCAGGAGACCCCGUCGGCCACGCUC GGAUCUUCCUUUCCGAUCCGCCAUC GUGGGUGGAGCCGCCGCCACG	underline: putative TOP motif	Ensembl transcript ID ENSMUST00000102845.10
S27a-45'	GGGAGGAAAGAAUCGCAUCGGCUG UAUAAGAAAGCCUUUUGAGGCAUU UUUUUUAGUUGAGCACAUCAUUUC GAGGCCAUUCUGAGGUAAACCGAG AAAAGAGCGUAAAGAAACCGAGCG AACGAGCAAAUCUGGCACUGCGUU AGACAGCCGCGAUUCCGCUGCAGCG CGCAGGCACGUGUGUGGCCGCCUAA GGGGCGGGUCCUUCGGCCAGGAGA CCCCGUCGGCCACGCUCGGAUCUUC CUUUCCGAUCCGCCAUCGUGGGUGG AGCCGCCGCCACG		modifed from S27a-45
S27a-44	GGGUUUCCGAUCCGCCAUCGUGGG UGAGUGUAUGCUCUGUGGCCGCGC UCUGGCUAGUGGCGCUACGCGUCGC UCUCACGGGUGUCGUCGGAUCUAA UCCGUCUCUUUUCGAAUGCAGGUG GAGCCGCCGCCACG	underline: putative TOP motif	Ensembl transcript ID ENSMUST00000102844.3
S27a-44'	GGGGAUCCGCCAUCGUGGGUGAGU GUAUGCUCUGUGGCCGCGCUCUGGC UAGUGGCGCUACGCGUCGCUCUCAC GGGUGUCGUCGGAUCUAAUCCGUC UCUUUUCGAAUGCAGGUGGAGCCG CCGCCACG		modifed from S27a-44
10nt	GGGAGCCACC		De novo design
30nt	GGGAAAGAAACAGGACAGAAAACA GCCACC		De novo design
50nt	GGGAACGACAAGAAACACAUACAA AAGAAACAGGACAGAAAACAGCCA CC		De novo design

70nt	GGGAAGAGAUAAACAUAAACAUAA ACGACAAGAAACACAUACAAAAGA AACAGGACAGAAAACAGCCACC		De novo design
90nt	GGGGAGAAGAGGGAACAGGACACA AGAGAUAAACAUAAACAUAAACGA CAAGAAACACAUACAAAAGAAACA GGACAGAAAACAGCCACC		De novo design
NCA-1	GGACACAAACAAACAAACAAACAC ACAAACACACAAACAAACACACACA CACACAAACAAACAAGCCACC		De novo design
NCA-2	GGACACACACACACACACACACA CACACACACAAACACACACACACAC ACACAAACACACAAGCCACC		De novo design
NCA-3	GGAAAGAGAGAAAGAGAGAAAGAA AGAAAGAGAGAGAGAAAGAGAGAA AGAAAGAAAGAGAGAAGCCACC		De novo design
NCA-4	GGAGAGAGAGAGAGAGAGAAAGAG AGAGAGAGAGAAAGAGAGAGAGAG AGAGAGAGAGAGAGAAGCCACC		De novo design
NCA-5	GGAAACACAUAAACAUAAACAUAA CACACAACAAACACAUACAACACAA ACACAACACAAAACAGCCACC		De novo design
NCA-6	GGAAACACAAUAACAUAAUCAUAC UACACAACUAACACAUACAUCACAU ACACAUCACAUAACAGCCACC		De novo design
NCA-7	GGAAACACAAUAACAUAAUCAUAC UACACAACUAACACAUACAUCACAU ACACAUCACAUAACAGCCACC	Underline: microRNA target sites	De novo design
NCA-8	GGCUACACACUCUCACUCUCAUCACC UCACUACUCACUCUCUCAUCACUCU CACAUCACAUCACUGCCACC	Underline: microRNA target sites	De novo design
NCA-7d	GGCAAAAAUCAAAAUCAAUCAUCA UCACAACAUCAACAAUCAAUCAUCA ACACAUCAUCAAGACACCACC		Derived from NCA-7
NCA-8d	GGCAUCACACUCUCACUCUCAUCUC AACACUCCUCCUCAUUCCAAUCUCU CACACAUCCCAUUAGCCACC		Derived from NCA-8

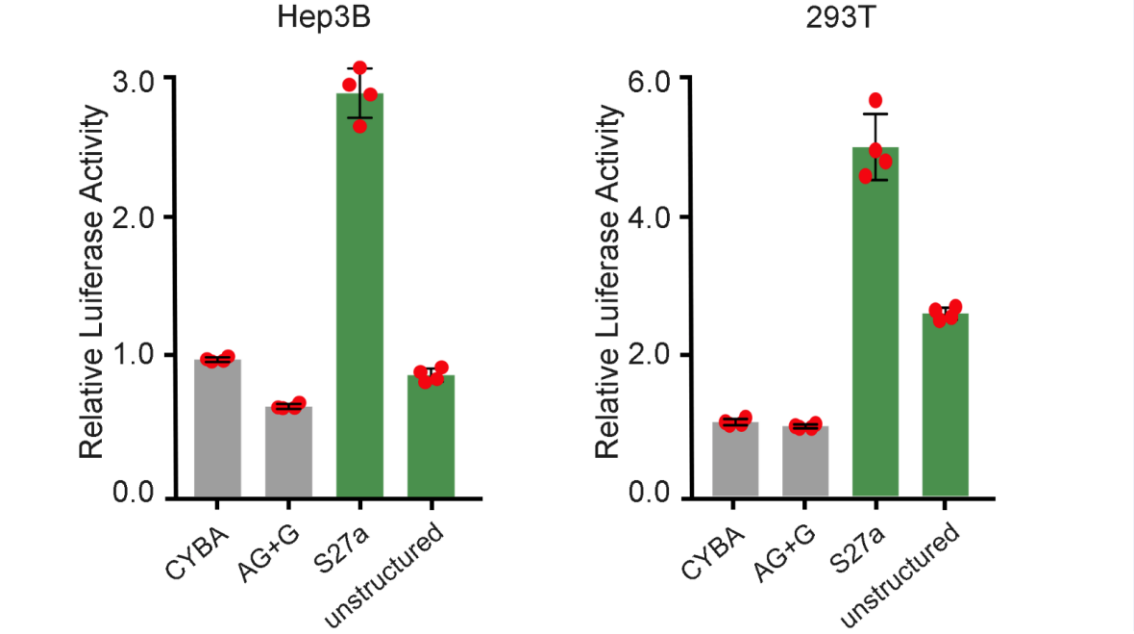
Supplementary Table 3. Sequences of 3' UTRs used in this study. The miRNA target sites were predicted by miRDB[2], a well-established online tool[3].

Name	Sequence	Annotation	Source
S27a	UUGUGUAUGCGUUAAUAAAAAG AAGGAACUCGUA		Ensembl transcript IDs: ENSMUST00000102845.10 and ENSMUST00000102844.3
AAT	CUGCCUCUCGCUCCUCAACCCCU CCCCUCCAUCCCUGGCCCCCUCCC UGGAUGACAUUAAAGAAGGGUU GAGCUGGUCCCUGCCUGCAUGUG ACUGUAAA		Ensembl transcript ID: ENST00000393087.8 Genbank ID NM_000295
TF	AAUCUCAGAGGUAGGGCUGCCAC CAAGGUGAAGAUGGGAACGCAG AUGAUCCAUGAGUUUGCCCUGGU UUCACUGGCCCAAGUGGUUUGUG CUAACCACGUCUGUCUUCACAGC UCUGUGUUGCCAUGUGUGCUGAA CAAAAAAUAAAAAUUAUUAUUG AUUUUAUAUUUCAAAAACUCCAU UCUUUCCUAAAUAUUUUCAACAA AGGAUUUCUUUAUGCAUUCUGCC UAAAUACCUAUGCAACUGAGCCC UUCCUUCUCAGCUCAAGAUUCGU CUGGUCUUUCCCUACAGCUUUGU GUGUGCCAUGGCCACAUCUCCUG GGUACAGUUCAAGGAGACAUCUU UUCUAAAAGGGUCUGCGUGAUCA UUAAAAUAUAAUCAAAUGUA		ENST00000402696.7
HCV	AGGUUGGGGUAAACACUCCGGCC UCUUAAGCCAUUUCCUGUUUUUU UUCUUUCCUUUCCUUCUUUUUUU CCUUUCUUUUUCCCUUCUUUAAU GGUGGCUCCAUCUUAGCCCUAGU CACGGCUAGCUGUGAAAGGUCCG UGAGCCGCAUGACUGCAGAGAGU GCUGAUACUGGCCUCUCUGCAGA UCAUGU		[4]
unstructured	CAAUUUUCCGCUCGAGCUCUAUA UGAUACAACAU		De novo design
S27a + QRE1	UUGUGUAUGCGUUAAUAAAAAG AAGGAACUCGUA	underline: QRE1	[5]

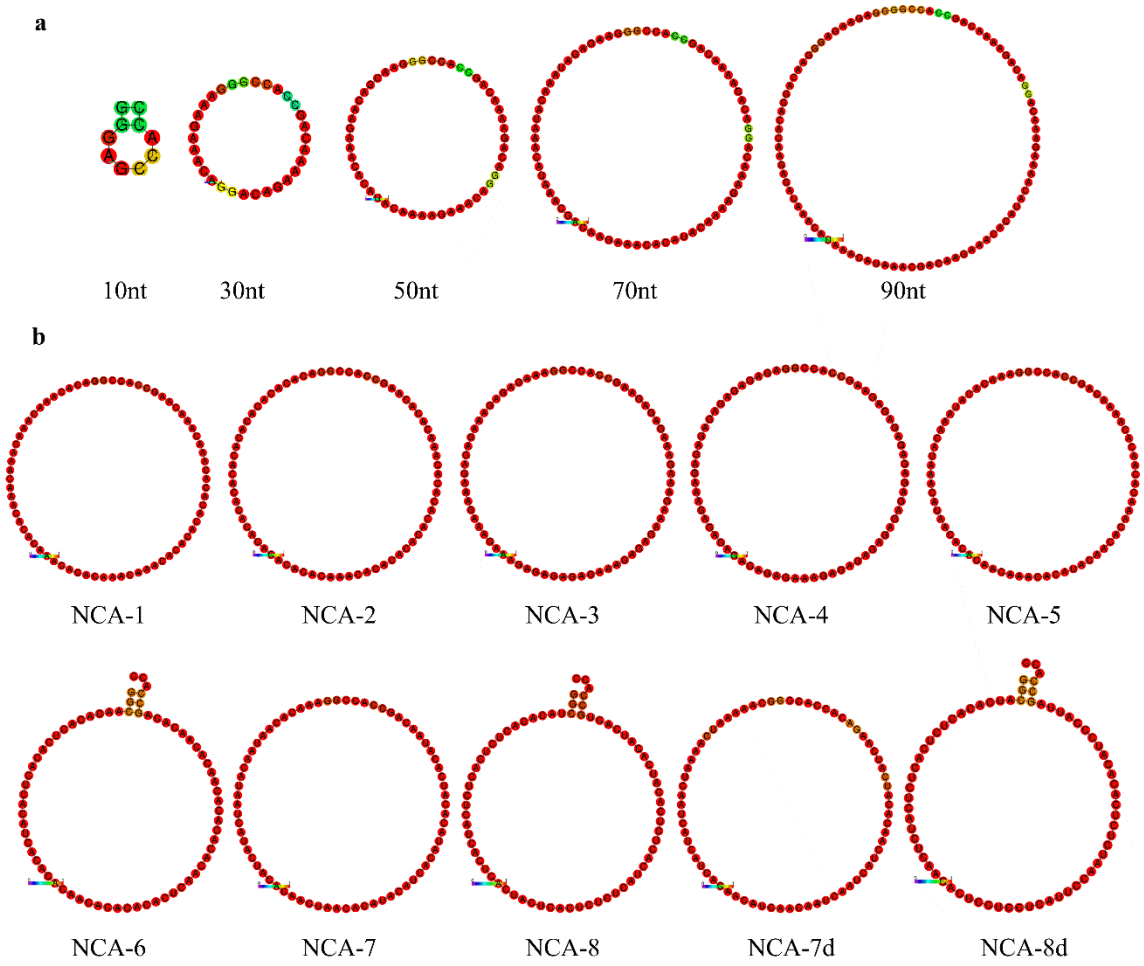
	GCCGUAACCACGUCUACUAACGCCG		
S27a + QRE2	UUGUGUAUGCGUUAAUAAAAAGAAGGAACUCGUAAACUCCAGGACUGUAUUUGUGACUAAUUGUAUAACAGGUU	underline: QRE2	6
S27a + R3U	UUGUGUAUGCGUUAAUAAAAAGAAGGAACUCGUAAAAACUCAAUGUAUUUCUGAGGAAGCGUGGUGCAUAAUGCCACGCAGCGUCUGCAUAACUUUUAUUAUUUCUUUUAUUAAUCAACAAA	underline: R3U	7
S27a + ARE	UUGUGUAUGCGUUAAUAAAAAGAAGGAACUCGUAUAUGUCUGUUUUUGUAUCUUUAUGCUGUAUUUUAACACUUUGUAUUACUUAGGUUAUU	underline: ARE	8
S27a + EMCV	UUGUGUAUGCGUUAAUAAAAAGAAGGAACUCGUAAUAAGAUACACCUGCAAAGGCGGCACAACCCCAGUGCCACGUUGUGAGUUGGAUAGUUGUGGAAAGAGUCAAAUGGCUCUCCUCAAGCGUAUUCAACAAGGGGCUGAAGGAUGCCCAGAAGGUACCCCAUUGUAUGGGAUCUGAUCUGGGGCCUCGGUGCACAUGCUUUACAUGUGUUUAGUCGAGGUUAAAAAACGUCUAGGCCCCCCGAACCACGGGGACGUGGUUUUCCUUUGAAAAACACGAUGAUAAU	underline: EMCV IRES	9
S27a + FMDV	UUGUGUAUGCGUUAAUAAAAAGAAGGAACUCGUACAGCAGGUUUCCCCAACUGACACAAAACGUGCAACUUGAAACUCCGCCUGGUCUUUCCAGGUCUAGAGGGGUAACACUUUGUACUUGGUGACAGGCUAAGGAUGCCCUUCAGGUACCCCGAGGUAACACGCGACACUCGGGAUCUGAGAAGGGGACUGGGGCUUCUAUAAAAGCGCUCGGUUUAAAAAGCUUCUAUGCCUGAAUAGGUGACCGGAGGUCGGCACCUUUCCUUUACAAUUAAUGACCCU	underline: FMDV IRES	10

Supplementary Table 4. Sequences of the control UTRs used in this study.

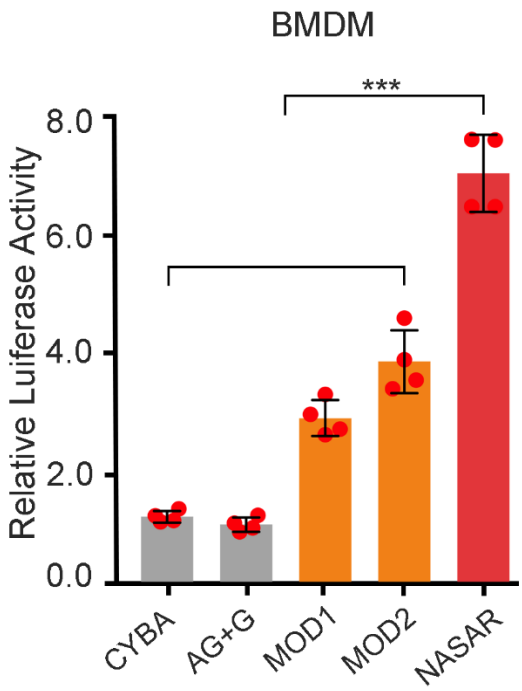
Name	5' UTR	3' UTR	Ref.
AG	GGGACUCUUCUGG UCCCCACAGACUCA GAGAGAACCCACC	GCUGGAGCCUCGGUGGCCAUGC UUCUUGCCCCUUGGGCCUCCCC CCAGCCCCUCCUCCCCUUCCUGC ACCCGUACCCCCGUGGUCUUUG AAUAAAGUCUGAGUGGGCGGCA	11
5AG+G	GGGACUCUUCUGG UCCCCACAGACUCA GAGAGAACGCCACC		12
AG+G	GGGACUCUUCUGG UCCCCACAGACUCA GAGAGAACGCCACC	GCUGGAGCCUCGGUGGCCAUGC UUCUUGCCCCUUGGGCCUCCCC CCAGCCCCUCCUCCCCUUCCUGC ACCCGUACCCCCGUGGUCUUUG AAUAAAGUCUGAGUGGGCGGCA	Modified from 12
CYBA	GGGCGCGCCUAGCA GUGUCCCAGCCGGG UUCGUGUCGCC	CCUCGCCCCGGACCUGCCCUCCC GCCAGGUGCACCCACCUGCAAU AAAUGCAGCGAAGCCGGGA	13
MOD1	GGGAAAUAAGAGA GAAAAGAAGAGUA AGAAGAAAUAUAA GAGCCACC	GCUGCCUUCUGCGGGGCUUGCC UUCUGGCCAUGCCCUUCUUCUC UCCCUUGCACCUGUACCUCUUG GUCUUUGAAUAAAGCCUGAGUA GGAAG	14
MOD2	GGGAAAUAAGAGA GAAAAGAAGAGUA AGAAGAAAUAUAA GAGCCACC	UGAUAAUAGGCUGGAGCCUCGG UGGCCAUGCUUCUUGCCCCUUG GGCCUCCCCCCAGCCCCUCCUCC CCUUCCUGCACCCGUACCCCCGU GGUCUUUGAAUAAAGUCUGA	15



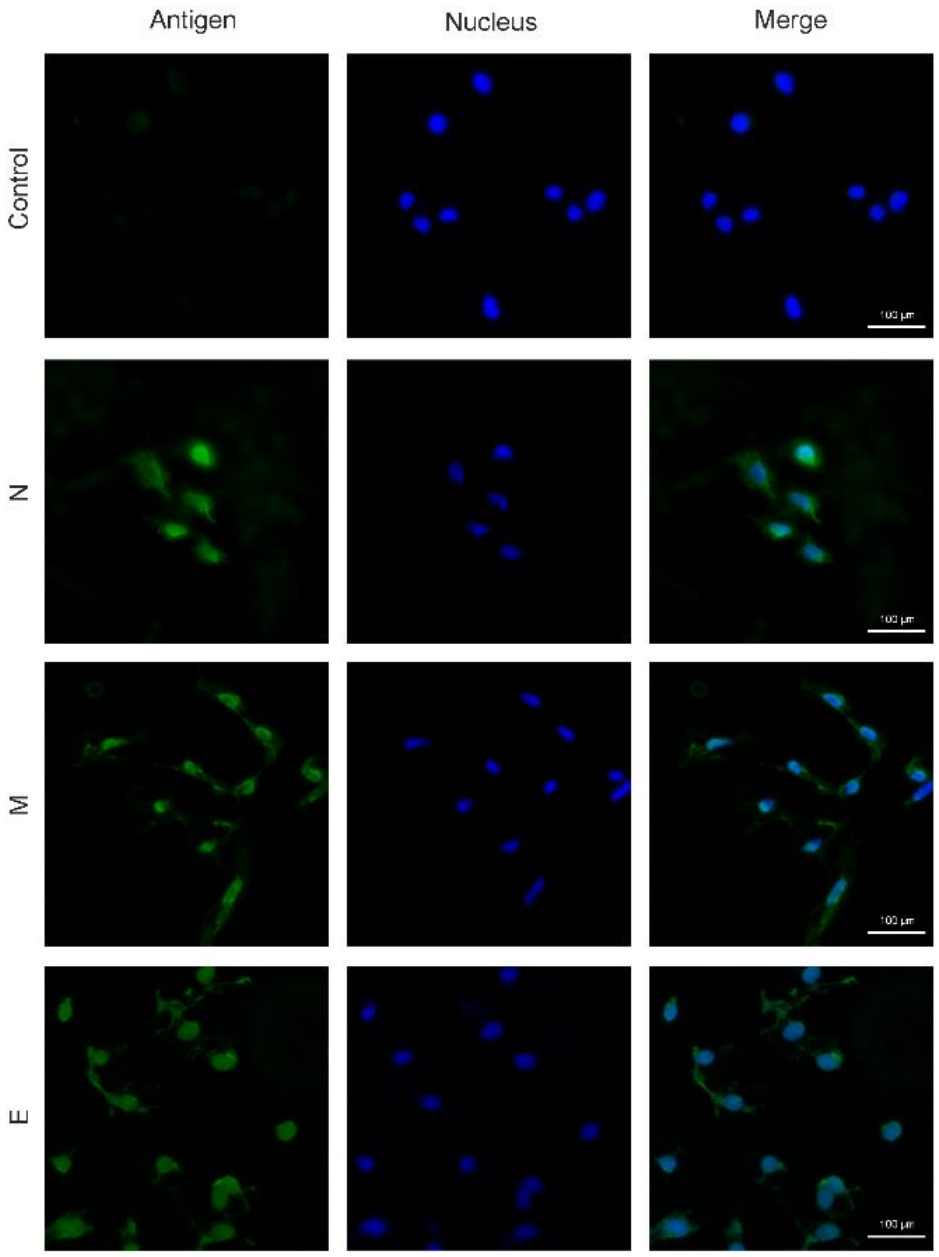
Supplementary Figure 1. Modification of 3' UTR. Relative luciferase activity in Hep 3B and 293T cells. Two different 3' UTR in green are assessed with the same 5' UTR. Relative luciferase activity was normalized to that of CYBA. All data are presented as the mean ± s.d. (n=4).



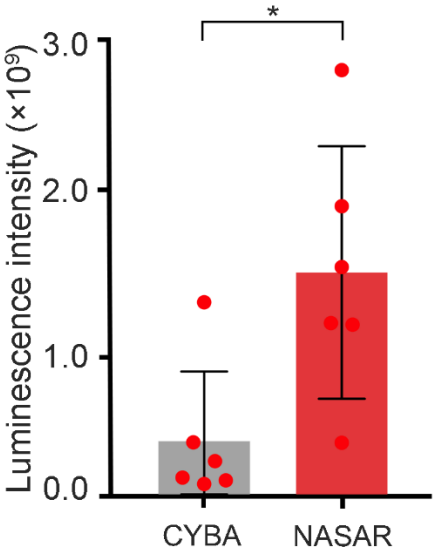
Supplementary Figure 2. Secondary structures prediction of de novo designed 5' UTRs. **a**, 5'UTRs of various lengths. **b**, 5' UTRs designed from nucleotide composition analysis. Minimum free energy (MFE) secondary structures were predicted by RNAfold, a well-established online tool[16]. Colors indicate base-pair probabilities with red and green showing high and low probability, respectively.



Supplementary Figure 3. Evaluation of NASAR mRNA in primary cells. Relative luciferase activity in bone marrow derived macrophages (BMDM). Relative luciferase activity was normalized to that of CYBA. All data are presented as the mean ± s.d. (n=4). Statistical significance in a and c was analyzed by the two-tailed Student's t-test. ****P* < 0.001.



Supplementary Figure 4. NASAR mRNAs expressing SARS-CoV-2 antigens. Fluorescent microscopy imaging of Nucleocapsid (N), Membrane (M), and Envelope (E) proteins in 293T cells. Scale bar = 100µm.



Supplementary Figure 5. Quantification of luciferase expression in vivo after i.m. injection to three mice per group (n=6 legs). Both mRNAs were formulated with TT3 nanoparticles. All data are presented as the mean ± s.d. Statistical significance in a and c was analyzed by the two-tailed Student's t-test. *$P < 0.05$.

References

1. Schwanhäusser, B. et al. Global quantification of mammalian gene expression control. *Nature* **473**, 337-342 (2011).
2. Chen, Y. & Wang, X. miRDB: an online database for prediction of functional microRNA targets. *Nucleic Acids Res* **48**, D127-D131 (2019).
3. Shankaraiah, R.C. et al. Metformin prevents liver tumourigenesis by attenuating fibrosis in a transgenic mouse model of hepatocellular carcinoma. *Oncogene* **38**, 7035-7045 (2019).
4. Bai, Y., Zhou, K. & Doudna, J.A. Hepatitis C virus 3'UTR regulates viral translation through direct interactions with the host translation machinery. *Nucleic Acids Res* **41**, 7861-7874 (2013).
5. Galarneau, A. & Richard, S. Target RNA motif and target mRNAs of the Quaking STAR protein. *Nature Structural & Molecular Biology* **12**, 691-698 (2005).
6. Zearfoss, N.R., Clingman, C.C., Farley, B.M., McCoig, L.M. & Ryder, S.P. Quaking Regulates Hnrnpa1 Expression through Its 3′ UTR in Oligodendrocyte Precursor Cells. *PLOS Genetics* **7**, e1001269 (2011).
7. Sokoloski, K.J. et al. Sindbis virus usurps the cellular HuR protein to stabilize its transcripts and promote productive infections in mammalian and mosquito cells. *Cell Host Microbe* **8**, 196-207 (2010).
8. Topisirovic, I. et al. Stability of eukaryotic translation initiation factor 4E mRNA is regulated by HuR, and this activity is dysregulated in cancer. *Molecular and cellular biology* **29**, 1152-1162 (2009).

9. Terenin, I.M., Andreev, D.E., Dmitriev, S.E. & Shatsky, I.N. A novel mechanism of eukaryotic translation initiation that is neither m7G-cap-, nor IRES-dependent. *Nucleic Acids Res* **41**, 1807-1816 (2013).
10. Fernandez, N. et al. Structural basis for the biological relevance of the invariant apical stem in IRES-mediated translation. *Nucleic Acids Res* **39**, 8572-8585 (2011).
11. Babendure, J.R., Babendure, J.L., Ding, J.-H. & Tsien, R.Y. Control of mammalian translation by mRNA structure near caps. *RNA* **12**, 851-861 (2006).
12. Schrom, E. et al. Translation of Angiotensin-Converting Enzyme 2 upon Liver- and Lung-Targeted Delivery of Optimized Chemically Modified mRNA. *Molecular therapy. Nucleic acids* **7**, 350-365 (2017).
13. Ferizi, M. et al. Human cellular CYBA UTR sequences increase mRNA translation without affecting the half-life of recombinant RNA transcripts. *Scientific reports* **6**, 39149 (2016).
14. Lee, J. et al. mRNA-mediated glycoengineering ameliorates deficient homing of human stem cell-derived hematopoietic progenitors. *The Journal of clinical investigation* **127**, 2433-2437 (2017).
15. Richner, J.M. et al. Modified mRNA Vaccines Protect against Zika Virus Infection. *Cell* **168**, 1114-1125.e1110 (2017).
16. Wesselhoeft, R.A., Kowalski, P.S. & Anderson, D.G. Engineering circular RNA for potent and stable translation in eukaryotic cells. *Nature Communications* **9**, 2629 (2018).