COVID-19 Protein Expression Analysis in Cell-Free System

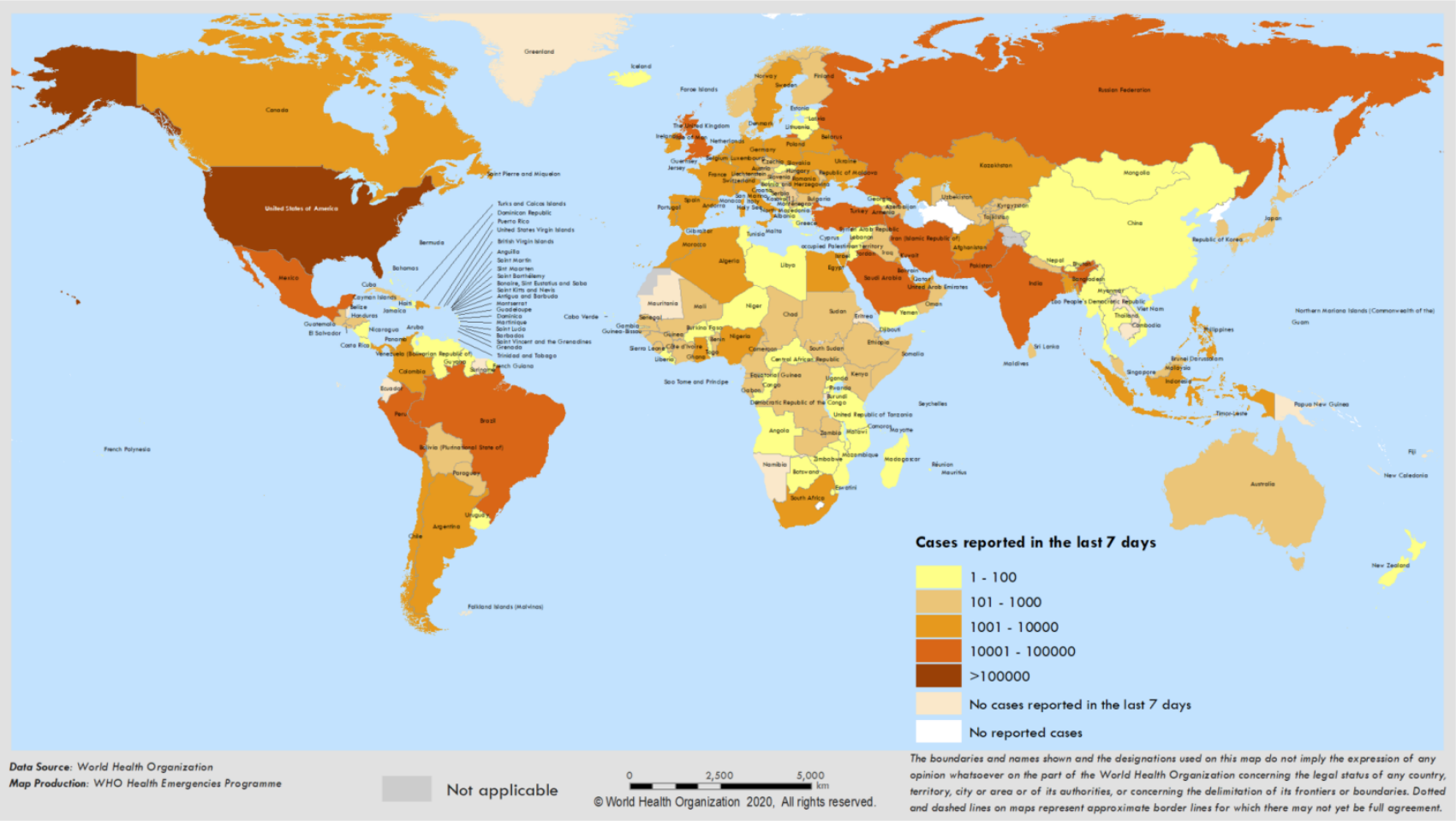
ChemE 7770 Final Project
Tina Ye (ty369) & Xiaojing Ma (xm89)
May 14th 2020

Outline

1. Motivation and Background Information.
2. Rationale of the Simulation Approach.
3. Quantitative and Qualitative Analysis.
4. Key Findings and Conclusion.

COVID-19 is A Current Pandemic

Figure 1. Number of confirmed COVID-19 cases reported in the last seven days by country, territory or area, 6 May to 12 May**



- 4.34M confirmed cases; 297K death by May 13th.
- Transmission: contact with contaminated objects or between persons. (CDC)
- Affects our normal life and the economy.

Project Tasks

1. Investigate different constraints' effect on optimal productivity of RdRp, nsp1, and nsp10 in the cell free system.

2. Propose modeling to reveal the expression of COVID19 nonstructural protein nsp1 varies with time.

3. Review literature's modeling on the effects of nsp1 on the expression of GFP-ssrA protein.

SARS-CoV-2 Proteins for this Study

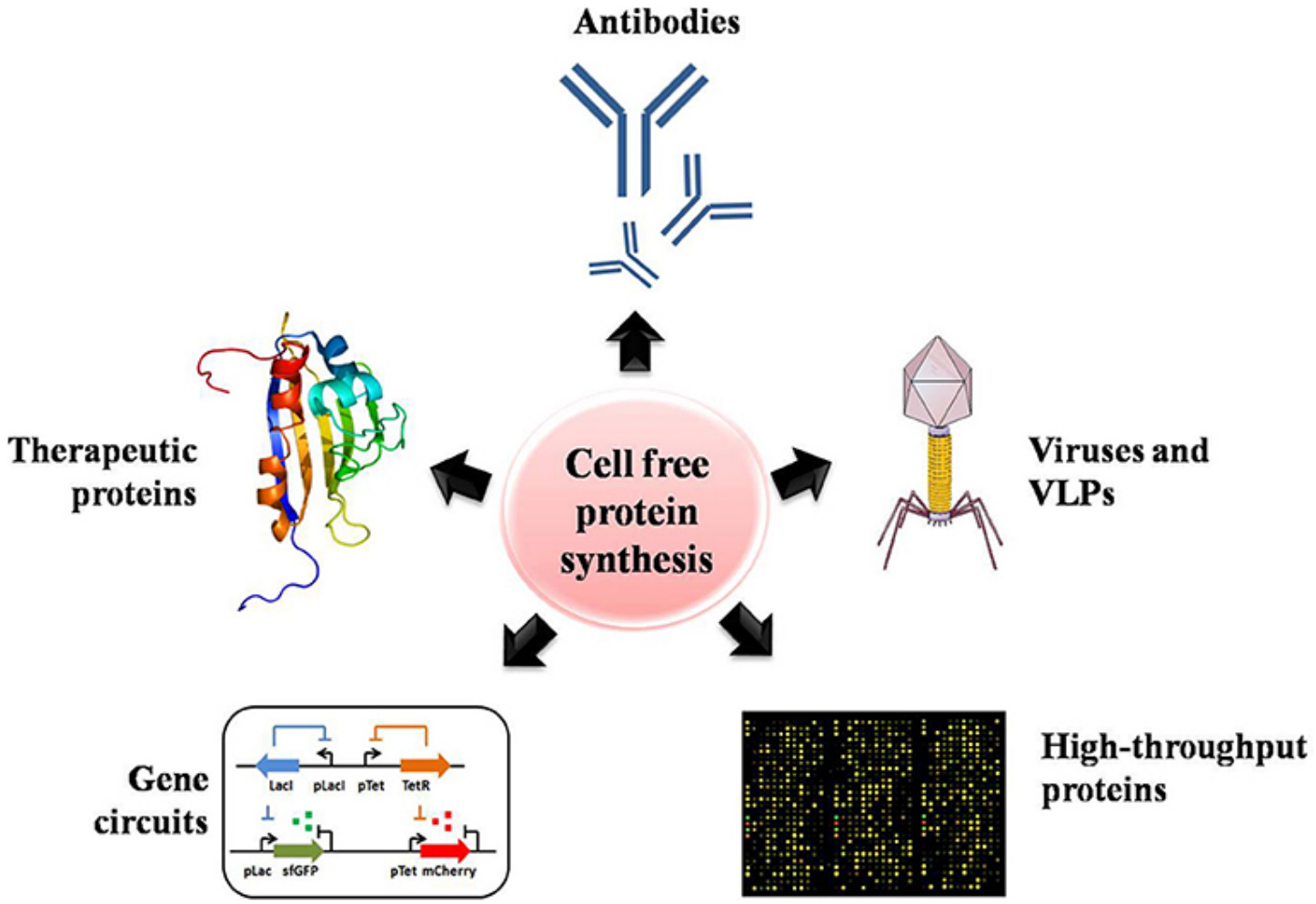
- **RdRp** (the RNA-dependent RNA polymerase, also named nsp12):
  - central component of coronaviral replication/transcription machinery
  - primary target for the antiviral drug, remdesivir.

- **nsp1** (non-structural protein 1): inhibit host gene expression
  - degradation of expressed RNA transcripts and host endogenous mRNAs

- **nsp10** (non-structural protein 10): a critical cofactor for activation of multiple replicative enzymes
  - binds and stimulates both the nsp14 and nsp16 activities.

Yan Gao et al, 2020; Wataru Kamitani et al, 2006; Bouvet M et al, 2014

Cell Free Protein Synthesis as Model Condition



- The commercial *E. coli* TX-TL cell-free protein synthesis (CFPS)

- Advantage:
  - high efficiency
  - flexibility
  - low cost

1. FBA Sequence-Specific Modeling

Major Reactions Involved for Protein Synthesis

Simplified, fundamental reaction set for protein production

Transcription initiation:	$G + \text{RNAP} \xrightarrow{v_1} G^*$
Transcription:	$G^* + n\text{NTP} \xrightarrow{v_2} \text{mRNA} + G + \text{RNAP} + 2n\text{P}_i$
mRNA decay:	$\text{mRNA} \xrightarrow{v_3} n\text{NMP}$
Translation initiation:	$\text{mRNA} + \text{rib} \xrightarrow{v_4} \text{rib}^*$
Translation:	$\text{rib}^* + a\text{AAtRNA} + 2a\text{GTP} \xrightarrow{v_5} a\text{tRNA} + 2a\text{GDP} + 2a\text{P}_i + \text{rib} + \text{mRNA} + \text{protein}$
tRNA charging:	$\text{AA} + \text{tRNA} + \text{ATP} \xrightarrow{v_6} \text{AMP} + 2\text{P}_i + \text{AAtRNA}$
Exchange fluxes:	$\text{AA}_{ext} \xrightarrow{b_1} \text{AA}$
	$\text{NTP}_{ext} \xrightarrow{b_2} \text{NTP}$
	$\text{protein} \xrightarrow{b_3} \text{protein}_{ext}$
	$\text{NMP} \xrightarrow{b_4} \text{NMP}_{ext}$
	$\text{ATP}_{ext} \xrightarrow{b_5} \text{ATP}$
	$\text{AMP} \xrightarrow{b_6} \text{AMP}_{ext}$
	$\text{GTP}_{ext} \xrightarrow{b_7} \text{GTP}$
	$\text{GDP} \xrightarrow{b_8} \text{GDP}_{ext}$
	$\text{P}_i \xrightarrow{b_9} \text{P}_{i\,ext}$

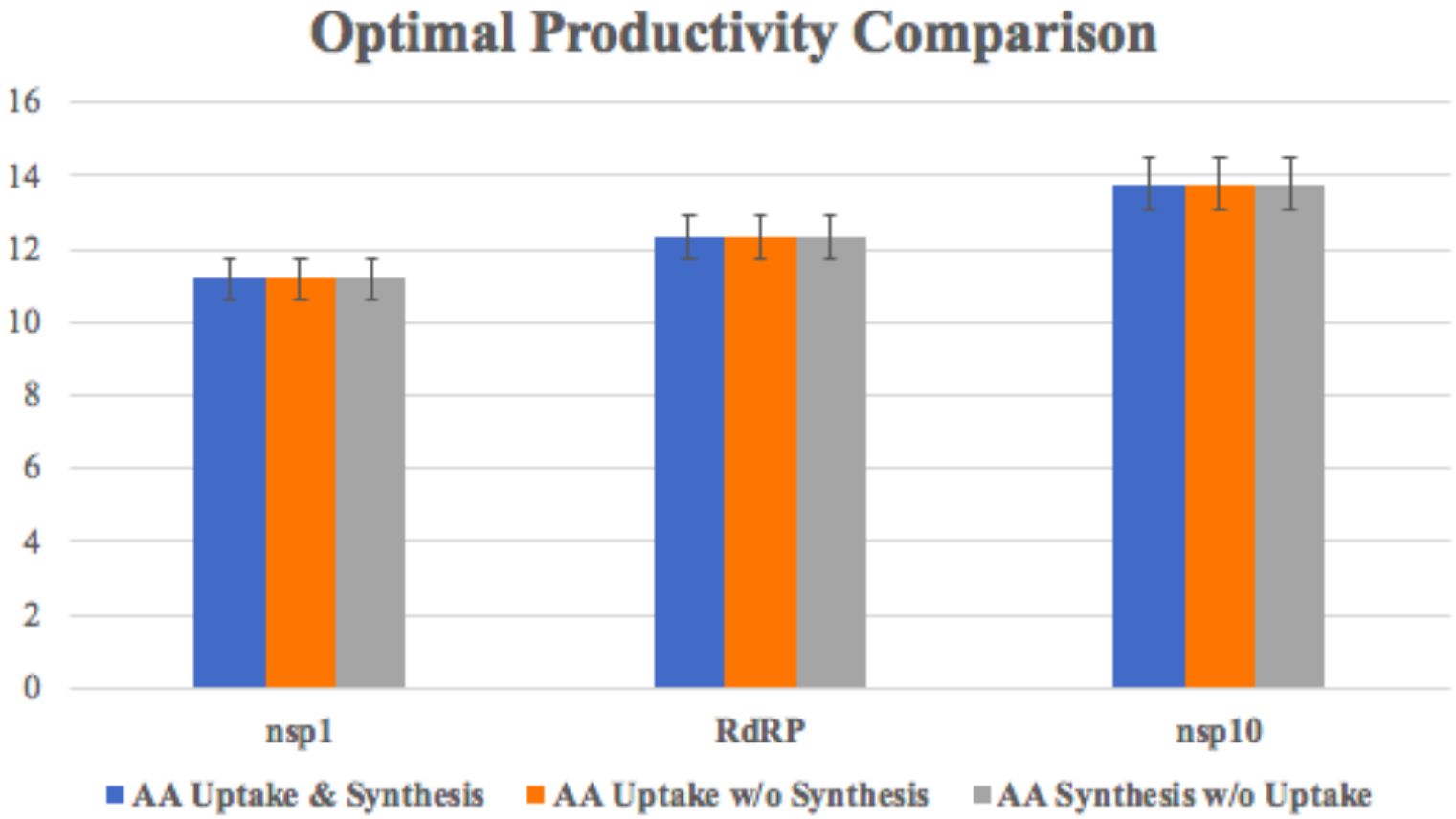
Eqn for internal production of RNAP, tRNA, and rRNA showed in Appendix.

- G is the protein of interest: RdRp, nsp1, or nsp10
- Assume commercial *E. coli* myTXTL extract
- Assume Steady State Condition in Cell-Free System
- Adopted coding template from Varner Lab

T. Allen et al, 2003, M. Vilkhovoy, 2018

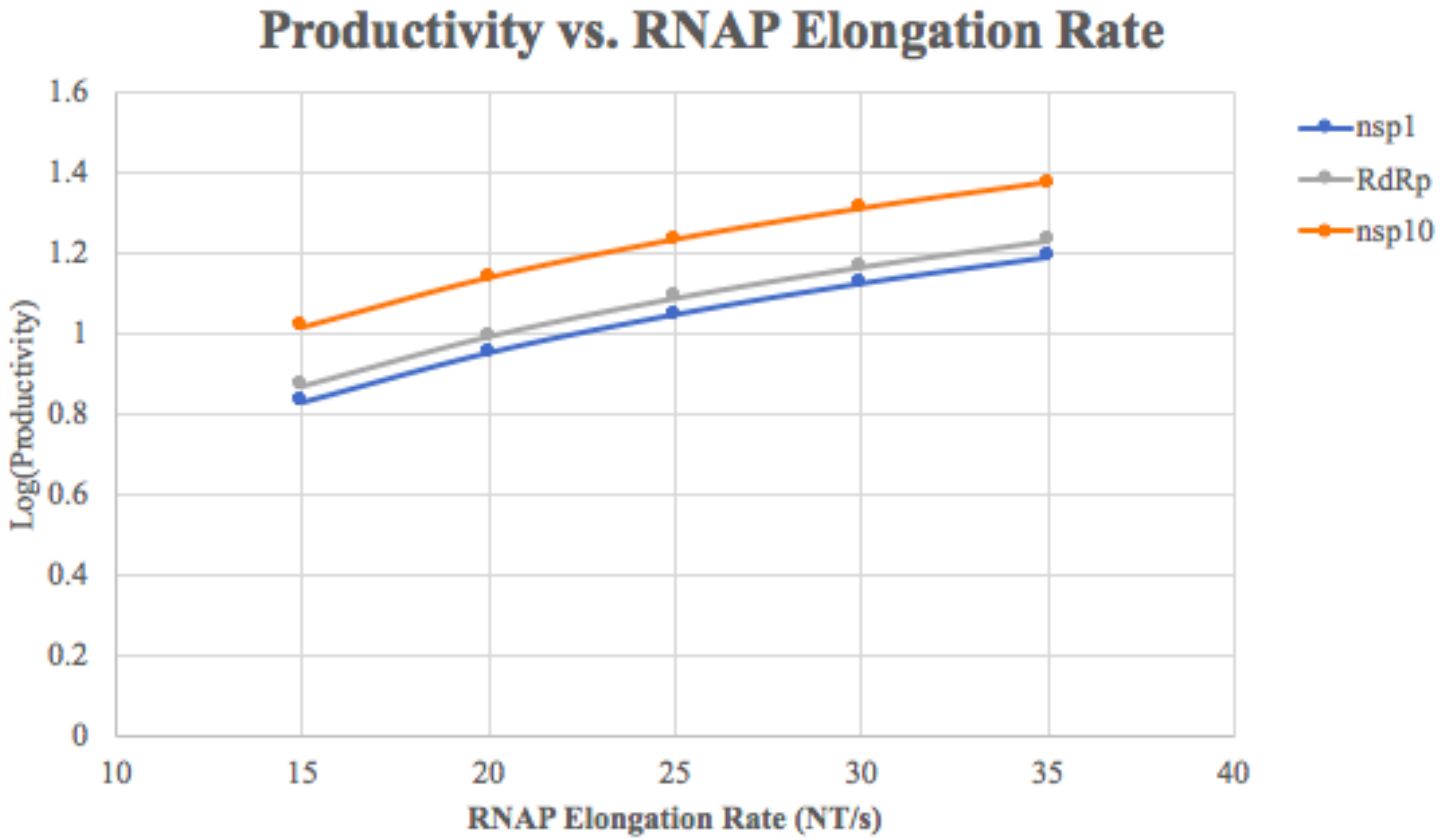
1. FBA Sequence-Specific Modeling

Simulation Result



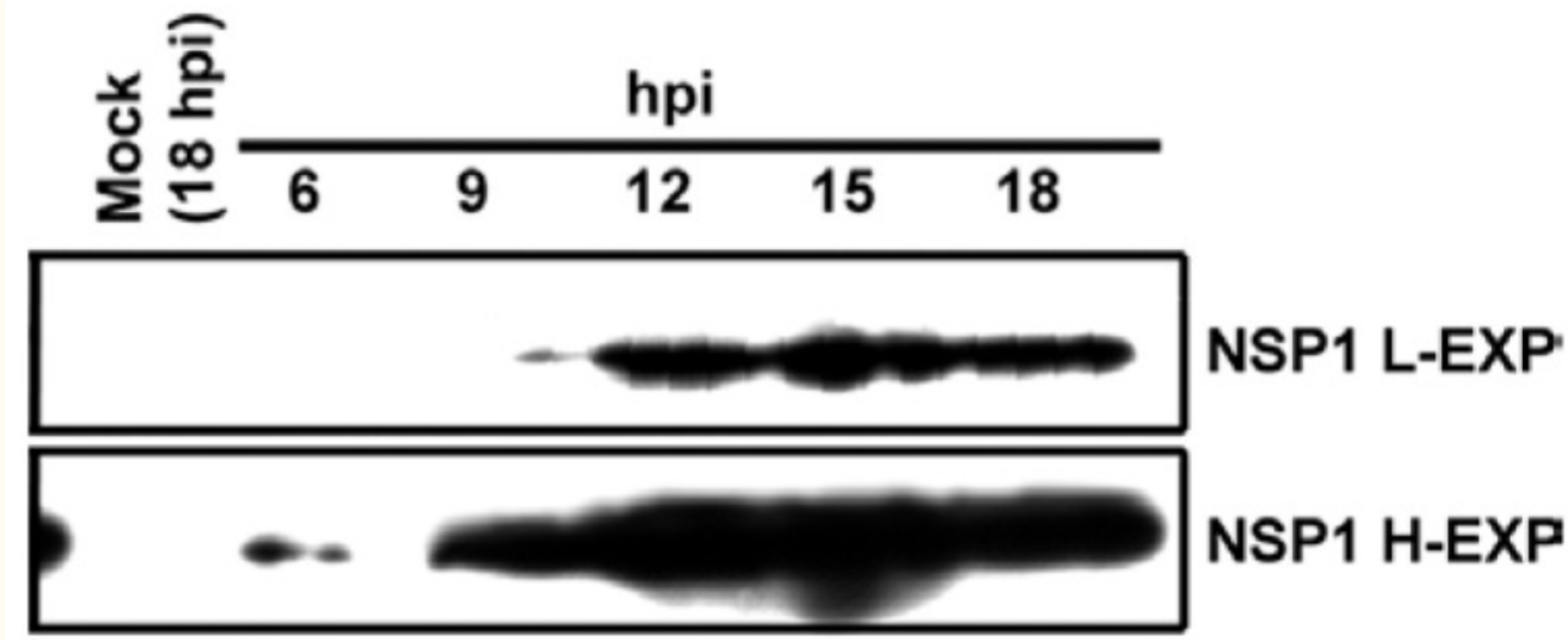
- Optimal productivity inversely proportional to carbon number of the protein.
  - nsp1:180 aa, 872 C
  - RdRp: 141 aa, 704 C
  - nsp10: 139 aa, 636 C
- AA uptake w/ or w/o de novo synthesis seems to not affect productivity.

1. FBA Sequence-Specific Modeling



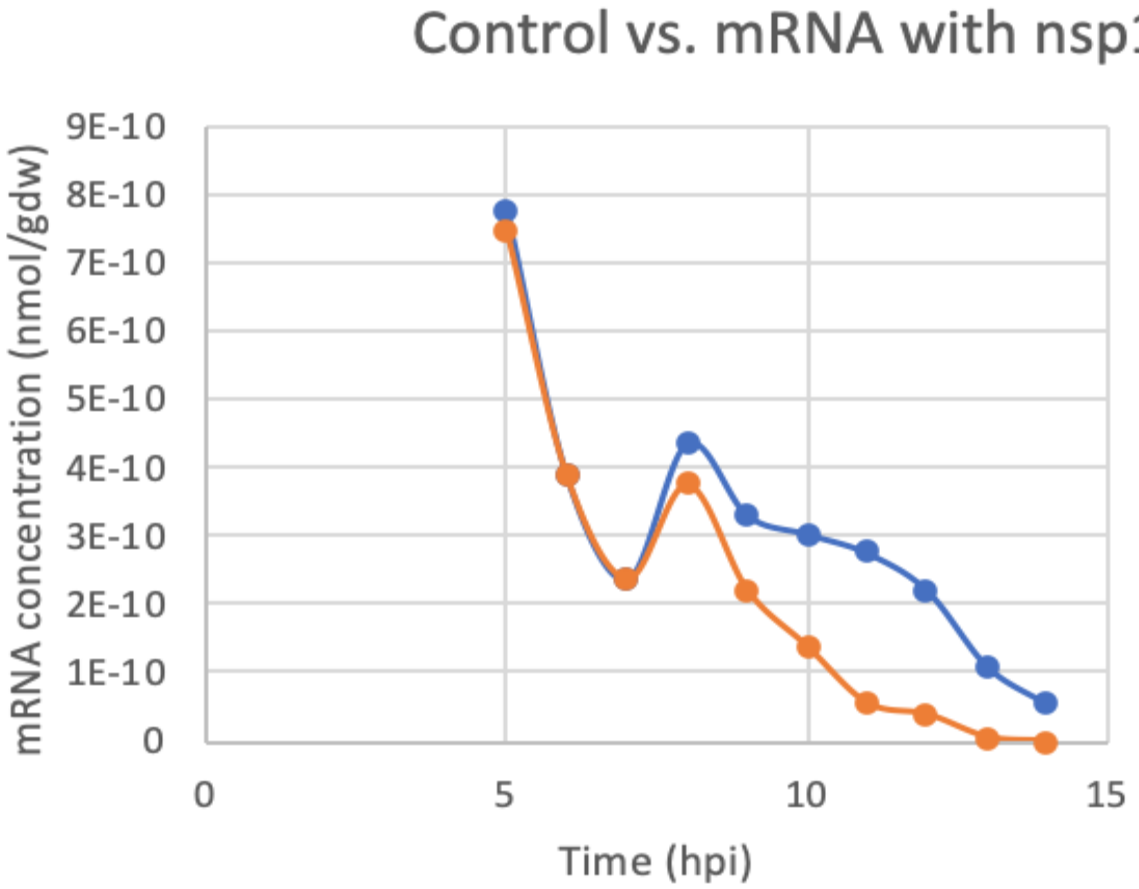
- Maximum optimal productivity increases proportionally to translation elongation rate.
- RNAP elongation rate has greater influence on productivity than for RNAP concentration.

2. The expression of nsp1 varies with time



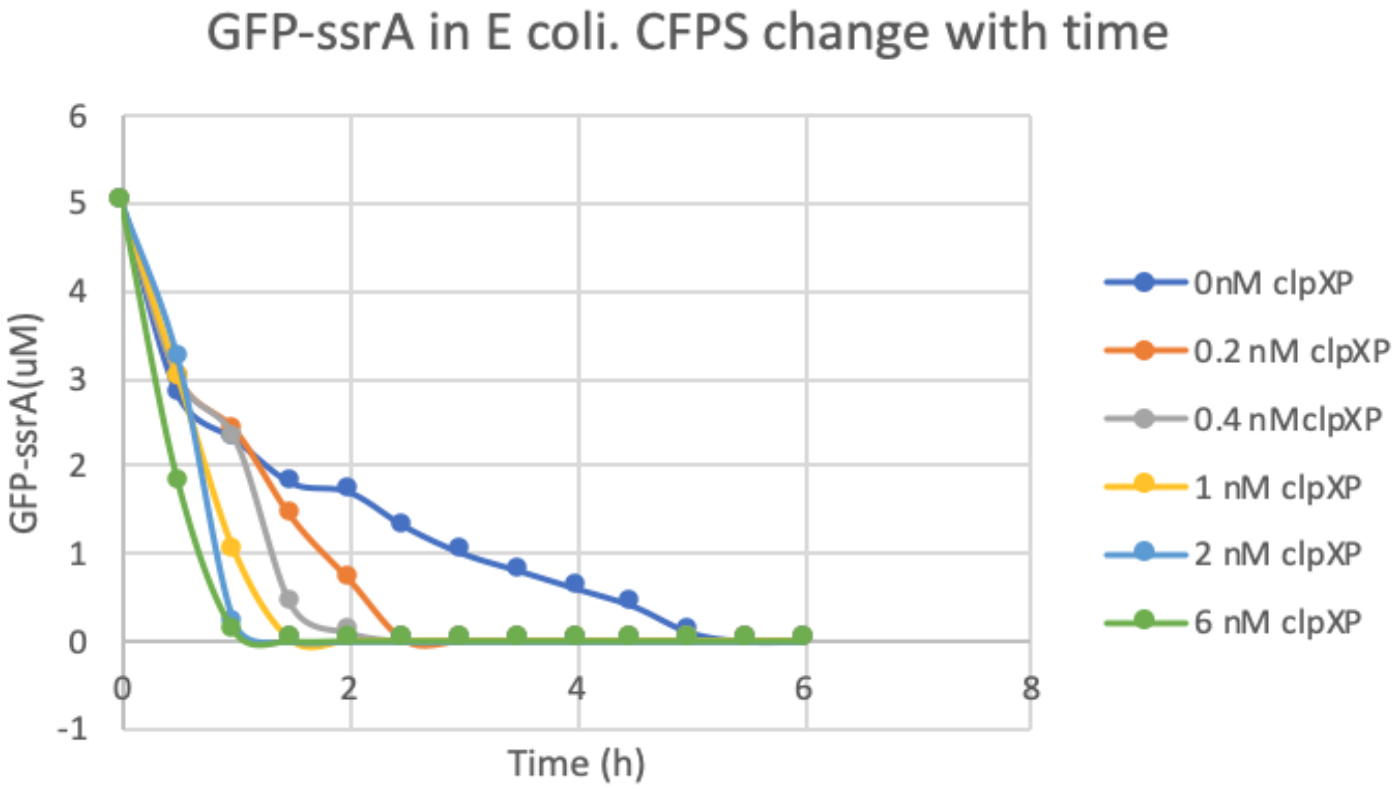
Laura Martínez-Álvarez et al, 2013

2. The expression of mRNA with nsp1



nsp1-CD and nsp1 employed the same mechanism to inhibit translation

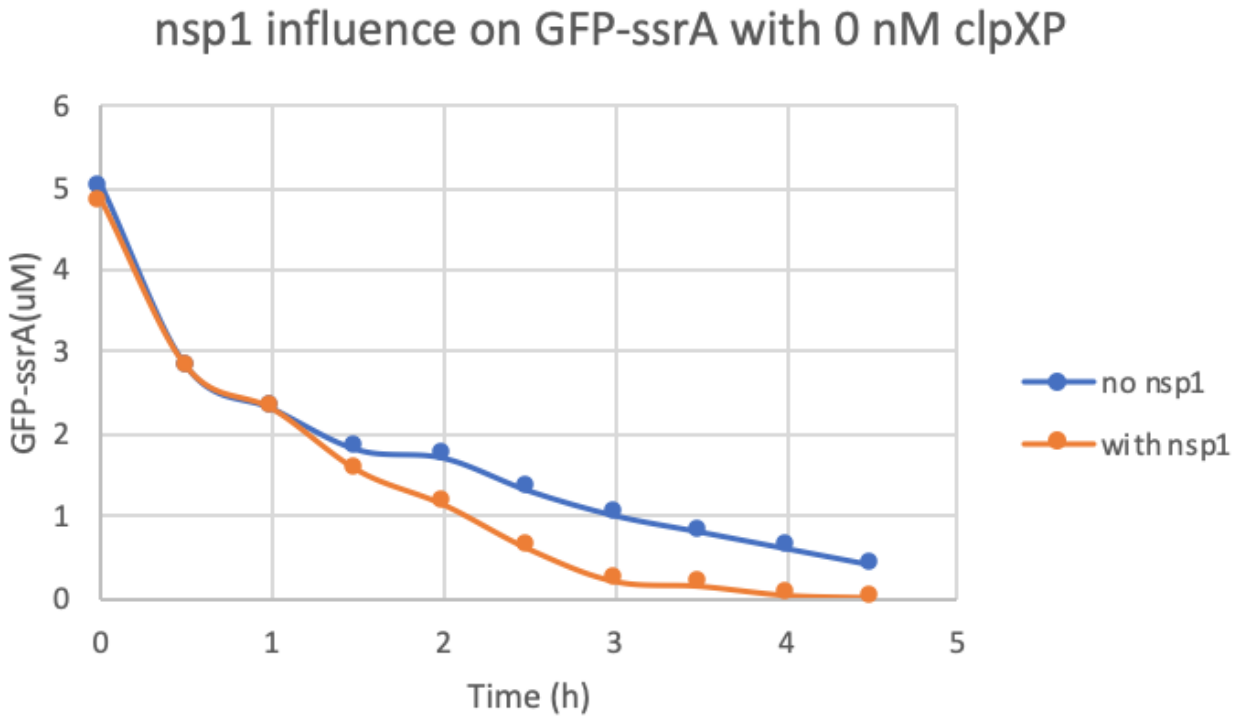
3. The expression of GFP-ssrA protein in CFPS



In E. coli, the ATP-dependent ClpXP protease contributes to degradation of ssrA-tagged proteins.

3. nsp1 Inhibits Production of GFP-ssrA

Time(hpi)	% change	GFP-ssrA(um) with 0nm clpXP	with nsp1
0	0.03571	5	4.82142857
0.5	0	2.8	2.8
1	0	2.3	2.3
1.5	0.13924	1.8	1.54936709
2	0.33333	1.7	1.13333333
2.5	0.54545	1.3	0.59090909
3	0.8	1	0.2
3.5	0.825	0.8	0.14
4	0.95	0.6	0.03
4.5	1	0.4	0



Conclude with Key Findings

- Optimal productivity inversely proportional to carbon number of the protein.
- Translation elongation rate greatly affects productivity compares to RNAP and ribosome abundance.
- NSP1 is a late viral protein.
- The amount of clpXP influence the expression of GFP-ssrA in *E coli*. CFPS.
- The nsp1 promotes degradation of GFP-ssrA in cell free reaction.

Thank you for your attention!

Reference

Bouvet, Mickaël, et al. “Coronavirus Nsp10, a Critical Co-Factor for Activation of Multiple Replicative Enzymes.” *Journal of Biological Chemistry*, 12 Sept. 2014, www.jbc.org/content/289/37/25783.full.

CohenApr, Jon, et al. “COVID-19 Vaccine Protects Monkeys from New Coronavirus, Chinese Biotech Reports.” *Science*, 6 May 2020, www.sciencemag.org/news/2020/04/covid-19-vaccine-protects-monkeys-new-coronavirus-chinese-biotech-reports.

“Disease Background of COVID-19.” *European Centre for Disease Prevention and Control*, 8 May 2020, www.ecdc.europa.eu/en/2019-ncov-background-disease.

Farrell, Christopher M., et al. “Cytoplasmic Degradation of SsrA-Tagged Proteins.” *Wiley Online Library*, John Wiley & Sons, Ltd, 17 Aug. 2005, onlinelibrary.wiley.com/doi/full/10.1111/j.1365-2958.2005.04798.x.

Gao, Yan, et al. “Structure of the RNA-Dependent RNA Polymerase from COVID-19 Virus.” *Science*, American Association for the Advancement of Science, 10 Apr. 2020, science.sciencemag.org/content/early/2020/04/09/science.abb7498.

Ghose, Tia. “How Are People Being Infected with COVID-19?” *LiveScience*, Purch, 7 Apr. 2020, www.livescience.com/how-covid-19-spreads-transmission-routes.html.

Martínez-Álvarez, Laura, et al. “The Shift from Low to High Non-Structural Protein 1 Expression in Rotavirus-Infected MA-104 Cells.” *Memorias Do Instituto Oswaldo Cruz*, Instituto Oswaldo Cruz, Ministério Da Saúde, June 2013, www.ncbi.nlm.nih.gov/pmc/articles/PMC3970611/.

PMC, Europe. *Europe PMC*, europepmc.org/article/MED/23035226.

“UNITED STATES DEPARTMENT OF LABOR.” *Safety and Health Topics | COVID-19 - Background | Occupational Safety and Health Administration*, www.osha.gov/SLTC/covid-19/background.html.

Appendix 1. Equation for Internal Production of RNAP, tRNA, and rRNA

Reactions added to the fundamental system when including the internal production of RNAP, tRNA, and rRNA

RNAP:	$G_P + \text{RNAP} \xrightarrow{v_{1P}} G_P^*$
	$G_P^* + n_P\text{NTP} \xrightarrow{v_{2P}} \text{mRNA}_P + G_P + \text{RNAP} + 2n_P\text{P}_i$
	$\text{mRNA}_P \xrightarrow{v_{3P}} n_P\text{NMP}$
	$\text{mRNA}_P + \text{rib} \xrightarrow{v_{4P}} \text{rib}_P^*$
	$\text{rib}_P^* + a_P\text{AAtRNA} + 2a_P\text{GTP} \xrightarrow{v_{5P}} a_P\text{tRNA} + 2a_P\text{GDP} + 2a_P\text{P}_i + \text{rib} + \text{mRNA}_P + \text{RNAP}$
tRNA:	$G_t + \text{RNAP} \xrightarrow{v_{1t}} G_t^*$
	$G_t^* + n_t\text{NTP} \xrightarrow{v_{2t}} \text{tRNA} + G_t + \text{RNAP} + 2n_t\text{P}_i$
rRNA:	$G_r + \text{RNAP} \xrightarrow{v_{1r}} G_r^*$
	$G_r^* + n_r\text{NTP} \xrightarrow{v_{2r}} \text{rib} + G_r + \text{RNAP} + 2n_r\text{P}_i$

Appendix 2. calculations for task 2

Assumption: 100 mRNA/cell

Table 1				
Time(hpi)	mRNA (% of total)	mRNA (mRNA/cell)	control (% of total)	control (mRNA/cell)
5	13.5	13.5	14	14
6	7	7	7	7
7	4.3	4.3	4.3	4.3
8	6.8	6.8	7.9	7.9
9	4	4	6	6
10	2.5	2.5	5.5	5.5
11	1	1	5	5
12	0.7	0.7	4	4
13	0.1	0.1	2	2
14	0	0	1	1

Appendix 3. calculations for task 2

Calculations					
mRNA with nsp1(mol)	mRNA with nsp1 (nmol)	mRNA with nsp1 (nmol/gdw)	control (mol)	control (nmol)	control (nmol/gdw)
2.24178E-23	2.24178E-14	7.4726E-10	2.3248E-23	2.32481E-14	7.74936E-10
1.1624E-23	1.1624E-14	3.87468E-10	1.1624E-23	1.1624E-14	3.87468E-10
7.14048E-24	7.14048E-15	2.38016E-10	7.1405E-24	7.14048E-15	2.38016E-10
1.12919E-23	1.12919E-14	3.76398E-10	1.3119E-23	1.31186E-14	4.37286E-10
6.64231E-24	6.64231E-15	2.2141E-10	9.9635E-24	9.96347E-15	3.32116E-10
4.15144E-24	4.15144E-15	1.38381E-10	9.1332E-24	9.13318E-15	3.04439E-10
1.66058E-24	1.66058E-15	5.53526E-11	8.3029E-24	8.30289E-15	2.76763E-10
1.1624E-24	1.1624E-15	3.87468E-11	6.6423E-24	6.64231E-15	2.2141E-10
1.66058E-25	1.66058E-16	5.53526E-12	3.3212E-24	3.32116E-15	1.10705E-10
0	0	0	1.6606E-24	1.66058E-15	5.53526E-11

Appendix 4 calculations for task 3

Time(h)	GFP-ssrA(um) with 0nm clpXP	0.2nm clpXP	0.4nm clpXP	1nm clpXP	2nm clpXP	6nm clpXP
0	5	5	5	5	5	5
0.5	2.8	3	3	3	3.2	1.8
1	2.3	2.4	2.3	1	0.2	0.1
1.5	1.8	1.4	0.4	0	0	0
2	1.7	0.7	0.1	0	0	0
2.5	1.3	0	0	0	0	0
3	1	0	0	0	0	0
3.5	0.8	0	0	0	0	0
4	0.6	0	0	0	0	0
4.5	0.4	0	0	0	0	0
5	0.1	0	0	0	0	0
5.5	0	0	0	0	0	0
6	0	0	0	0	0	0

Appendix 5. Gene Sequence for FBA

nsp1 QHD43415_1(L=180) (https://zhanglab.ccmb.med.umich.edu/COVID-19/)

MESLVPGFNEKTHVQLSLPVLQVRDVLVRGFGDSVEEVLSEARQHLKDGTCGLVEVEKGVLPQLEQPYVFIKRSDARTAPHGHV
MVELVAELEGIQYGRSGETLGVLVPHVGEIPVAYRKVLLRKNGNKGAGGHSYGADLKSFDLGDELGTDPYEDFQENWNTKHSSG
VTRELMRELNGG

Nsp10 QHD43415_10 (L=139) (https://zhanglab.ccmb.med.umich.edu/COVID-19/)

AGNATEVPANSTVLSFCAFAVDAAKAYKDYLASGGQPITNCVKMLCTHTGTGQAITVTPEANMDQESFGGASCCLYCRCHIDHP
NPKGFCDLKGKYVQIPTTCANDPVGFTLKNTVCTVCGMWKGYGCSCDQLREPMLQ

RdRp (L=141) (https://www.uniprot.org/uniprot/V5YMF8)

WDYPKCDRAMPNMLRIMASLILARKHSTCCNLSHRFYRLANECAQVLSEMVMCGGSLYVKPGGTSSGDATTAYANSVFNICQAV
TANVNALLSTDGNKIADKYVRNLQHK LYQNLYRNRD VDHEFVSEFY AYLRKHFSMM I