

BIOTECHNOLOGICAL DRUG PLATFORMS

**A DISSERTATION SUBMITTED TO
THE GRADUATE SCHOOL OF ENGINEERING AND SCIENCE
OF BILKENT UNIVERSITY
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR
THE DEGREE OF
DOCTOR OF PHILOSOPHY
IN
MATERIALS SCIENCE AND NANOTECHNOLOGY**

By

Recep Erdem Ahan

November 2023

BIOTECHNOLOGICAL DRUG PLATFORMS

By Recep Erdem Ahan

November 2023

We certify that we have read this dissertation and that in our opinion it is fully adequate, in scope and in quality, as a thesis for the degree of Doctor of Philosophy.

Urartu Özgür Şafak Şeker (Advisor)

Ayşe Elif Erson Bensan

Bahar Değirmenci Uzun

Ceren Özkul Koçak

Onur Çizmecioğlu

Approved for the Graduate School of Engineering and Science:

Orhan Arıkan
Director of the Graduate School

ABSTRACT

BIOTECHNOLOGICAL DRUG PLATFORMS

Recep Erdem Ahan

PhD in Materials Science and Nanotechnology

Advisor: Urartu Özgür Şafak Şeker

November 2023

Biopharmaceuticals, also known as biotechnological drugs, have revolutionized the treatment of many diseases by providing access to new mechanisms of action that can target the underlying biological processes behind the diseases. Technological advancement in biological sciences opens new paths to uncover new biopharmaceutical modalities in nature as well as to augment the existing modalities new functions. The implementation of engineering principles i.e., synthetic biology approaches have been transforming biopharmaceutical research wherein “smart” therapeutics are developed and deployed for treatment of previously intractable diseases. However, there are still unmet clinical needs that require novel and advanced biopharmaceuticals. In this thesis, I explored different biopharmaceuticals to characterize and/or advance their capabilities for diverse indications. Firstly, we have developed a prophylactic agent from the lectin protein, griffithsin, as for ancestral and the emerged strains of SARS-CoV-2. Secondly, we have advanced genetic technologies to engineer the probiotic *Escherichia coli* (*E. coli*) strain, Nissle 1917 (EcN), for therapeutical purposes. We developed a stable recombinant DNA transfer system based on cryptic plasmids of EcN. Furthermore, a synthetic protein

secretion system was envisioned and functionally validated in EcN to shuttle therapeutical proteins to diseases site. Finally, peptide tags for extracellular protein secretion as well as a cell surface protein display system were developed for Lachnospiraceae species which are parts of the healthy human gut microbiome. The technologies and methodologies described herein will pay the way for inventing and/or discovering novel biopharmaceuticals to treat current and future diseases.

Keywords: biotechnological drugs, cell engineering, synthetic biology, living therapeutics, protein therapeutics, biopharmaceuticals



ÖZET

BİYOTEKNOLOJİK İLAÇ PLATFORMLARI

Recep Erdem Ahan

Malzeme Bilimi ve Nanoteknoloji, Doktora

Tez Danışmanı: Urartu Özgür Şafak Şeker

Kasım 2023

Biyofarmasötikler bir başka adıyla biyoteknolojik ilaçlar, hastalıklara neden olan biyolojik mekanizmaları hedef alabilecek yeni moleküler etken yollarına erişim sağlayarak birçok hastalığın tedavisinde devrim yaratmıştır. Biyoloji bilimlerindeki teknolojik ilerleme, doğadaki keşfedilmemiş yeni biyofarmasötik ilaçların bulunması ve mevcut biyofarmasötiklere yeni işlevler kazandırılmasına için imkân sağlamaktadır. Biyofarmasötik araştırmalar mühendislik ilkelerinin, bir başka deyişle sentetik biyoloji yaklaşımlarının uygulanması ile daha önce tedavisi mümkün olmayan hastalıkların tedavisinde kullanabilecek yeni “akıllı” terapötiklerin geliştirilmesine olanak sağlayan için bir değişim geçirmektedir. Ancak hâlâ yeni ve gelişmiş biyofarmasötikler geliştirilmesi klinikte tedavi edilemeyen hastalıklar için elzemdir. Bu tezde, farklı biyofarmasötik sınıflarının çeşitli endikasyonlara karşı etkinliği karakterize edilmiş ve/veya moleküler düzeyde etkinlikleri artırılması için yeni teknolojik araçlar geliştirilmiştir. İlk olarak, SARS-CoV-2'nin ataları ve sonradan ortaya çıkan türleri için lektin proteini olan Griffithsin profilaktik bir ajan kullanabileceği kanıtlanmıştır. İkinci olarak, probiyotik *Escherichia coli* (*E. coli*) türü Nissle 1917' yi (EcN) tedavi amaçlı olarak tasarlayacak yeni genetik teknolojiler

geliştirilmiştir. Stabil DNA transfer yapılabilmesi için EcN kriptik plazmitlerinin vektörler olarak kullanılmasına dayanan bir yöntem önerilmiş ve çalıştığı ispatlanmıştır. Buna ek olarak, terapötik proteinlerin EcN bakterisi kullanılarak üretilerek hedeflenen bölgeye salgılanması bir sentetik protein salgılama sistemi tasarlanmıştır. Son olarak, sağlıklı insan bağırsak mikrobiyomunun bir parçası olan Lachnospiraceae türleri için hücre dışı protein salgılanmasına yönelik peptit etiketleri ve hücre yüzeyi protein taşıma sistemi geliştirilmiştir. Burada açıklanan teknolojiler ve metodolojiler, mevcut ve gelecekteki hastalıkları tedavi etmek için yeni biyofarmasötiklerin icat edilmesine ve/veya keşfedilmesine olanak sağlayacağı öngörülmektedir.

Anahtar kelimeler: biyoteknolojik ilaçlar, hücre mühendisliği, sentetik biyoloji, yaşayan ilaçlar, protein tabanlı ilaçlar, biyofarmasötikler

ACKNOWLEDGEMENTS

In my pursuit of becoming a scientist, my graduate life in the Synthetic Biosystems Laboratory (SBL) at Bilkent University taught me invaluable lessons. Endless attempts to optimize new concepts, night-long characterization experiments, and never-ending PCR runs showed me the hardship; meanwhile, I learned to be patient and stubborn after each failed experiment. Although all these experiences helped me to push forward whenever I felt myself in the abyss of desperation, the most precious thing that I am acquainted with is teamwork. I witnessed that the best science can only be done with teams, and science is not some sort of competition but rather collaborative efforts of people. Therefore, I want to thank everybody who made this journey bearable for me.

First and foremost, I want to express my deepest gratitude to my thesis advisor, Urartu Şeker. His contagious curiosity and courage forced me to leave my comfort zone and explore the things that I was completely unfamiliar with. Each new horizon that we set with Urartu drove me to learn many stuffs I wouldn't have learned otherwise. While doing all these, he showed utmost support and always left room for failure and growth. Hence, I am thankful for the trust he put in me and gave me the chance to become a true independent researcher and biotech entrepreneur. I am grateful to have the privilege to be one of the early members of his laboratory.

I am thankful to my thesis committee members, Onur Çizmecioğlu and Ayse Elif Erson-Bensan. They helped me to navigate the next steps of thesis research, and their critical comments on my data in each meeting improved and shaped the research in this thesis.

I want to thank present and past members of SBL for the support they showed me during the time I spent at Bilkent University. Particularly, I am thankful to Gizem Makas for her support in everything related to laboratory coordination and administration. She made our life at SBL easy by taking care of everything except our thesis. I am grateful to meet and get to know Esra Yuca, and Burcu Ergün. I always remember their kindness and our discussions on science during coffee breaks.

Spending long times at UNAM put a significant toll on me, but luckily, I was among the friends who were kind and joyful. I am grateful to Cemile Elif Özçelik, Gökçe Özkul, and Dogus Akboğa for the “breaks” that became farther each year from the UNAM’ location whenever I needed to refresh my mind after a long and overwhelming experiment. I would like to thank Ebru Şahin Kehribar for her friendship, career advice, and the coffee breaks. Surprise birthday cakes, snack parties at the middle office, coffees brewed with new fancy espresso machine, barbeque party at Urartu’s place, and late lunches were delightful with Ebru Şahin Kehribar, Cemile Elif Özçelik, Gökçe Özkul, Çisil Köksaldı, Mehmet Emin Bakar, Dogus Akboğa, Ahmet Hinçer, Nedim Kurt, Ece Avcı, Merve Erden, Merve Yavuz, Suat Tuçer, Senem Şen, Asli Semerci, Melis Karaca, Cem Dirse Polat, Nazlıcan Tunc, Abdurrahman Atilla, Damla Albayrak and Derin Akman. Everybody that I knew at the SBL were wonderful and brilliant scientists. I hope we can keep in touch in the future.

I would like to thank Cemile Özçelik, Dogus Akboğa, Zeynep Edizcan, Ömer Ergül, and İrem Niran Çağıl for their help in my thesis. Their valuable insights and their help in performing tricky and annoying experiments are truly indispensable to me. I always felt gifted to work alongside them.

I am grateful to Mark Mimee for hosting me as an exchange student in his laboratory at UChicago. It was a pleasure for me to work with Jack Arnold on the Lachno engineering project, and I want to thank Jack for his time to show me how to deal with anaerobic bugs. I am thankful to Ella for accompanying me to go to weekly seminars around the campus. I also want to thank Jaehyun, Joyce, Sadiksha, and Sandra for tips on what to do around Chicago and their hospitality. Visiting Mark's lab was a valuable experience for me.

Some friendships are meant to be life-long, and I am lucky to have met a few of them at my high school. I want to express my thankfulness to Deniz Ağırdan, Efe Bent, and Umut Eskiçırak for their support and friendship that has been over 15 years.

I want to express my deepest gratefulness to my family. Without their support and unconditioned love, I wouldn't have been able to study this long. I want to thank my mother Atikhan, my father Kadir, my big brother and his wife Oguzhan and Arzugül, my "little" sister Zülfiye and my nephew Rifat.

Lastly, I want to thank my biggest but uncredited collaborator, my wife, Merve. Her continuous support and love keep me going even the times I thought I didn't have the material to do PhD. She was the main reason that I was immune to hopelessness, failures, and exhaustion. I am glad that you are with me to listen to my ramblings on my work, and I am happy to continue my science journey with company of you.

I thank Bilkent University UNAM for the Graduate fellowship to support me during my PhD. My stay at UChicago was supported by TUBITAK (The Scientific and Research Council of Turkey) BİDEB 2214-A program and I thank TUBITAK for the fundings. The thesis section related to development of griffithsin protein as anti-SARS-Cov-2 blocking agent was support by TUBITAK COVID Platform

(TÜBİTAK, Grant No: T1004 -18AG020) and the Ministry of Industry and Technology of Turkey.



TABLE OF CONTENTS

ACKNOWLEDGEMENTS.....	v
TABLE OF CONTENTS.....	ix
LIST OF FIGURES	xvi
LIST OF TABLES	xxii
CHAPTER 1	1
Introduction.....	1
1.1. Protein-Based Therapeutics.....	3
1.1.1. Antibodies and Synthetic Antibody-like Scaffolds	4
1.1.2. Cytokines and Protein Hormones	8
1.1.3. Recombinant Human Enzymes	10
1.1.4. Miscellaneous Proteins	10
1.2. Nucleic Acid Based Therapeutics	11
1.2.1. mRNA Therapeutics.....	13
1.3. Virus and Bacteriophage Based Therapeutics.....	15
1.3.1. Human Virus as Viral Vectors for Therapies.....	16
1.3.2. Engineered Bacteriophages	20
1.4. Cell Based Therapeutics.....	21
1.4.1. Engineered Human Cells	22
1.4.2. Engineered Microbial Cells	25
1.5. Thesis Scope, and Structure.....	28
CHAPTER 2	31
Griffithsin as a Broad-Spectrum Antiviral Agent.....	31
2.1. Objective of the Study	32

2.2. Introduction.....	33
2.3. Materials and Methods.....	36
2.3.1. Cloning, Expression, and Purification of Griffithsin Protein.....	36
2.3.2. SDS-PAGE and Western Blotting.....	37
2.3.3. Circular Dichorism.....	38
2.3.4. ITC Assay	38
2.3.5. QCM Assay.....	38
2.3.6. rGRFT Binding on SARS-CoV-2 by ELISA.....	40
2.3.7. Virus, Cells, Animals, and Ethical Approval	40
2.3.8. Toxicity and Immunogenicity	41
2.3.9. In Vitro Testing of rGRFT	42
2.3.10. In Vivo Testing of rGRFT.....	43
2.4. Results	44
2.4.1. Cloning of rGRFT into pET22b	44
2.4.2. Expression and Purification of rGRFT	44
2.4.3. Structural Changes of rGRFT upon Binding with Mannose Sugar.....	45
2.4.3. Binding of rGRFT to SARS-CoV-2 Spike Protein	47
2.4.4. Binding of rGRFT to Whole SARS-CoV-2	47
2.4.5. Antiviral Activity of rGRFT against SARS-CoV-2 <i>in vitro</i>	48
2.4.6. Toxicity and Immunogenicity of rGRFT	49
2.4.7. Anti-viral Activity of rGRFT on SARS-CoV-2 <i>in vivo</i> .	54
2.4.8. Anti-viral Activity of rGRFT on SARS-CoV-2 Variants	56
2.5. Discussion	57

CHAPTER 3	62
Genetic Technologies to Modify Probiotic <i>E. coli</i> Nissle Strain	62
3.1. Objective of the Study	62
3.2. Introduction.....	62
3.3. Materials and Methods.....	65
3.3.1. Construction of Plasmids Used in the Study	65
3.3.2. Curing of pMUT1 and pMUT2 from EcN strain	67
3.3.3. Colony PCR for pMUT1 and pMUT2	67
3.3.4. Stability of pMUT1 and pMUT2 in ΔpMUT1ΔpMUT2 EcN Strain	68
3.4. Results	68
3.4.1. Construction of gRNA Expression Plasmid	68
3.4.2. Cloning of gRNAs targeting the pMUT1 and pMUT2 into CRISPR Array Plasmid.....	69
3.4.3. Curing of pMUT1 in WT EcN strain.....	70
3.4.4. Curing of pMUT2 in ΔpMUT1 EcN strain	72
3.4.5. Construction of Recombinant pMUT1 and pMUT2 with Constitutively Active sfGFP Expression Cassette (pMUT1-sfGFP and pMUT2-sfGFP).....	74
3.4.6. Stability of pMUT1-sfGFP and pMUT2-sfGFP in ΔpMUT1ΔpMUT2 EcN Strain	76
3.5. Discussions.....	77
CHAPTER 4	79
Synthetic Secretion Systems for <i>E. coli</i> Nissle.....	79
4.1. Objective of the Study	79
4.2. Introduction.....	80
4.3. Materials and Methods.....	84

4.3.1. Construction of Plasmids Used in the Study	84
4.3.2. Scanning Electron Microscopy	87
4.3.3. Outer Membrane Permeability Assay with FTIC-4kDa	88
4.3.4. Determination of Growth Curve and Cell Motility	88
4.3.5. Western blot analysis for protein secretion with iLOM-SS	89
4.3.6. PnPP Assay for ALP Activity	89
4.3.7. iLOM-SS Characterizations	89
4.3.8. Reversible Activation of iLOM-SS with Toggle Switch	91
4.3.9. Functional Assessments of Compact iLOM-SS in Different Genetic Circuits	92
4.3.10. Determination of Effects of Leaked Periplasmic Content (LPC) on Caco-2 Cell Viability, Monolayer Permeability, and Cell Migration.....	94
4.3.11. Determination of Inflammatory Response of Treated Caco-2 Cells with LPC	96
4.3.12. Caco-2 Invasion Assay	96
4.3.13. ELISA, Western blot analysis and Functional Assay for IL-1Ra	97
4.3.14. Nanobody Secretion with iLOM-SS.....	98
4.3.15. Statistical Analysis	98
4.4. Results	98
4.4.1. Cloning of micL into IPTG Inducible pET22b	98
4.4.2. ESEM Analysis of micL Expressing Cells.....	99
4.4.3. FTIC-4kDa Dextram Diffusion Assay with micL Expressing Cells	100
4.4.4. Secretion of DsbA, bALP, MalS and ChiA with iLOM-SS	102

4.4.5. Characterization of iLOM-SS Using PhoA as Reporter	104
4.4.6. Conditional and Reversible Secretion by iLOM-SS....	108
4.4.7. Construction of Compact-micL.....	114
4.4.8. Functionality Assessment of Compact-micL in Thiosulfate Sensing Circuit.....	116
4.4.9. Functionality Assessment of Compact-micL in Recombinase Based Memory Circuits.....	118
4.4.10. Functionality Assessment of Compact-micL in Quorum Sensing Circuit	120
4.4.11. Effects of micL Expression on EcN Physiology	123
4.4.12. Safety Assays on Caco-2 cells for iLOM-Induced Leaked Periplasmic Content of EcN cells.....	125
4.4.13. IL-1Ra Secretion with iLOM-SS Equipped EcN.....	128
4.4.14. Nanobody Secretion with iLOM-SS Equipped EcN .	131
4.5. Discussions	133
CHAPTER 5	136
Engineering Human Gut Commensal Lachnospiraceae Species for Protein Delivery	136
5.1. Objective of the Study	136
5.2. Introduction.....	136
5.3. Materials and Methods.....	138
5.3.1. Construction of Plasmids Used in the Study	138
5.3.2. Cultivation of Lachnospiraceae species.....	139
5.3.3. Sample Preparation for Proteomic Analysis for Secreted and Intracellular Proteomes of <i>R. gnavus</i>	142
5.3.4. Identification of Signal Peptides in <i>R. gnavus</i> Genome using SignalP6	142

5.3.5. De novo Design of the Signal Peptide Based on Conserved Regions in Signal Peptides of <i>R. gnavus</i>	142
5.3.6. Manual Literature Search to Find Experimentally Validated Secreted Proteins in Lachnospiraceae Species.....	143
5.3.7. Conjugation of Recombinant Plasmids into Lachnospiraceae Species	143
5.3.8. Nanoluciferase Assay	144
5.3.9. Trypsin Surface Trimming of Lachnospiraceae Species	144
5.4. Results	144
5.4.1. Signal Peptide Selection Approaches for Protein Secretion in Lachnospiraceae	144
5.4.2. Shotgun Proteomic Analysis of <i>R. gnavus</i> Intracellular Soluble Fraction and Secretome.....	145
5.4.3. Bioinformatic Analysis of <i>R. gnavus</i> Genome.....	147
5.4.4. Characterization of Secretion Efficiency of Signal Peptides in <i>R. gnavus</i> and <i>C. comes</i>	148
5.4.5. Protein Display Shuttle for <i>R. gnavus</i> and <i>C. comes</i> ...	153
5.5. Discussions	157
CHAPTER 6	159
Conclusion and Future Perspectives	159
6.1. Overview	159
6.2. Major Contributions to the Present Literature	159
6.3. Future Perspectives.....	162
6.3.1. Broad Antivirals for Future Pandemics	162
6.3.2. Engineering Probiotic and Commensal Bacteria Species for Protein Delivery	163
6.3.3. Engineering Native Human Commensal as Theranostic Cells	164

BIBLIOGRAPHY	165
APPENDIX A	185
Primers Used in the Thesis.....	185
APPENDIX B	189
Genes Used in the Thesis	189
APPENDIX C	195
Plasmid Maps and Sequencing Results.....	195
APPENDIX D	210
General Recipes and Protocols	210
APPENDIX F.....	212
Additional Data	212
Time Course Microscopy Images of Caco-2 Cells During Wound Healing	212
Intracellular Proteome and Secretome of <i>R. gnavus</i>	212

LIST OF FIGURES

Figure 1-1. Different types of biopharmaceutical substances currently being developed or approved by regulatory agencies.	2
Figure 1-2. Different types of antibodies naturally produced by B cells and their activities.....	5
Figure 1-3. Nucleic acid-based therapeutics and their molecular mode of actions. ...	12
Figure 1-4. Structure of processed eukaryotic mRNA	14
Figure 1-5. Human infecting viruses are utilized as vectors and therapeutical modalities for gene therapy, reprogramming the immune cells and virotherapy.	17
Figure 1-6. Engineered human cells to treat different diseases.	23
Figure 1-7. Topology of an engineered microbes for therapeutical purposes.....	26
Figure 2-1. Expression plasmid for rGRFT expression in E. coli.	44
Figure 2-2. Purification of rGRFT.	45
Figure 2-3. CD analysis of apo-rGRFT and holo-rGRFT with mannose.	46
Figure 2-4. Validation of rGRFT binding to SARS-CoV-2 S protein.	47
Figure 2-5. Validation of rGRFT binding to whole SARS-CoV-2 virus (Ank2 strain).	48
Figure 2-6. <i>In vitro</i> assessment of antiviral activity of rGRFT against SARS-CoV-2 infection.	49
Figure 2-7. Cytotoxic and carcinogenic effects of rGRFT on Vero E6 cells.....	50
Figure 2-8. Assessment of rGRFT safety in vivo with C57BL/6 mice model.....	51
Figure 2-9. Analysis of serum cytokine levels of mice following rGRFT treatment.	52
Figure 2-10. <i>In vivo</i> antiviral activity of rGRFT against SARS-CoV-2 infection shown for direct or contact-based infection models.....	55
Figure 2-11. <i>In situ</i> hybridization of viral RNA in unprotected (b) and protected (a) animals following SARS-CoV-2 infection.	56
Figure 2-12. rGRFT inhibitory activity against SARS-CoV-2 variants <i>in vitro</i> with the Vero E6 infection model.....	57
Figure 2-13. Three dimensional structures of SARS-CoV-2 spike protein and griffithsin protein.....	58
Figure 3-1. Construction of pLacO mock CRISPR Array plasmid for expression of gRNA in the E. coli cells.....	68
Figure 3-2. Insertion of gRNAs into CRISPR array.....	70
Figure 3-3. Illustration of plasmid curing with pCas.	71
Figure 3-4. Curing of pMUT1 from the WT EcN strain.....	72

Figure 3-5. Curing of pMUT2 from the Δ pMUT1 EcN strain.....	73
Figure 3-6. Recombinant cryptic EcN plasmids with the constitutively active sfGFP expression construct.....	75
Figure 3-7. Stability of recombinant pMUT1 and pMUT2 in vitro in the Δ pMUT1 Δ pMUT2 EcN Strain.	76
Figure 4-1. Depiction of iLOM-SS constructed in E. coli.....	83
Figure 4-2. Construction of IPTG inducible micL expression plasmid.	99
Figure 4-3. ESEM images of WT and micL expressing E. coli cells.....	100
Figure 4-4. The iLOM system makes the outer membrane of bacteria leaky, allowing macromolecules to diffuse from the extracellular medium into the periplasm.....	101
Figure 4-5. Construction of pLacO micL/ pNTetO bALP plasmid.....	102
Figure 4-6. Cloning of his-tagged DsbA, MalS, and ChiA proteins at the downstream of pNtetO of pLacO micL/ pNTetO bALP plasmid.	103
Figure 4-7. Western blot analysis of cell-free supernatant obtained from culture expressing the micL sRNA and the proteins.	104
Figure 4-8. Cloning of control plasmids, pET22b pLacO mock RNA / pNTetO bALP and pET22b pLacO mock RNA / pNTetO yebF-bALP.....	105
Figure 4-9. Characterization of iLOM-based secretion system in EcN.....	107
Figure 4-10. Construction of pECJ-micL plasmid that encodes TetR/LacI based toggle switch.	109
Figure 4-11. Knocking out of LacI gene from E. coli genome.	111
Figure 4-12. Leaky membrane can be restored in the absence of micL expression to halt the secretion of periplasmically translocated proteins with the iLOM-SS.	113
Figure 4-13. Construction of pET22b c-micL bALP plasmid.....	115
Figure 4-14. Development of single transcription unit for iLOM-based secretion. .	116
Figure 4-15. Construction of pKD327 bALP and bALP-cmicL plasmids.	117
Figure 4-16. Functional validation of c-micL in thiosulfate sensing TCS based biosensors.	118
Figure 4-17. Construction of registry plasmids with bALP and bALP-micL.	119
Figure 4-18. Functional validation of c-micL in recombinase-based memory recoding circuit.	120
Figure 4-19. Construction of quorum sensing circuit to test functionality of c-micL.	121
Figure 4-20. Functional validation of c-micL activity in quorum sensing circuit....	122
Figure 4-21. Effects of micL expression on EcN physiology.	124
Figure 4-22. In vitro safety assessments of iLOM-based secretion from EcN.	126

Figure 4-23. Construction of pET22b pLacO micL / pNTetO IL-1Ra plasmid.	128
Figure 4-24. Active IL-1Ra secretion using iLOM-SS by EcN.....	130
Figure 4-25. Construction of pET22b pLacO micL / pNTetO nanobody plasmid...	131
Figure 4-26. Nanobody secretion using iLOM-SS by EcN.....	132
Figure 5-1. Intracellular proteome and secretome of <i>R. gnavus</i>	146
Figure 5-2. Comparison of the abundance of the detected proteins in cell-free supernatant and soluble fraction of cell lysate of <i>R. gnavus</i>	147
Figure 5-3. The depiction of most conserved amino acids among sec signal peptides of <i>R. gnavus</i> with logo plot.....	148
Figure 5-4. Cloning of signal peptides at the upstream of nanoluciferase in the conjugative plasmid.	149
Figure 5-5. Genetic engineering of Lachnospiraceae species.	151
Figure 5-6. Colony PCR to verify conjugation into <i>R. gnavus</i> (a) and <i>C. comes</i> (b).	152
Figure 5-7. Intracellular and extracellular luciferase activity of <i>R. gnavus</i> (a) or <i>C. comes</i> (b) cells.	153
Figure 5-8. The protein display shuttle for Lachnospiraceae cells.....	154
Figure 5-9. Construction of the display shuttle plasmids.	155
Figure 5-10. Validation of nanoluciferase surface display with trypsin surface trimming assay.....	156
Table A1. List of primers used in the study.....	185
Figure C-1. Plasmid map of IPTG inducible micL sRNA in pET22b (+) vector and sequence verification of the vector with sanger sequencing.	195
Figure C-2. Plasmid map of IPTG inducible micL sRNA and aTc inducible ChiA gene in pET22b (+) vector and sequence verification of the vector with sanger sequencing.....	195
Figure C-3. Plasmid map of IPTG inducible micL sRNA and aTc inducible DsbA gene in pET22b (+) vector and sequence verification of the vector with sanger sequencing.....	196
Figure C-4. Plasmid map of IPTG inducible micL sRNA and aTc inducible MalS gene in pET22b (+) vector and sequence verification of the vector with sanger sequencing.....	196
Figure C-5. Plasmid map of toggle switch circuit with micL in pECJ3 vector and sequence verification of the vector with NGS.....	196
Figure C-6. Plasmid map of aTc inducible ALP (PhoA) gene fused with c-micL in pET22b (+) vector and sequence verification of the vector with NGS.....	197

Figure C-7. Plasmid map of ALP (PhoA) gene located at inverted proD flanked by attB and attP sites of BxbI and sequence verification of the vector with NGS.	197
Figure C-8. Plasmid map of ALP (PhoA) gene fused with c-micL located at inverted proD flanked by attB and attP sites of BxbI and sequence verification of the vector with NGS.....	197
Figure C-9. Plasmid map of ALP (PhoA) cloned at the downstream of phsA in pKD237 vector and sequence verification of the vector with NGS.	198
Figure C-10. Plasmid map of ALP (PhoA) fused with c-micL cloned at the downstream of phsA in pKD237 vector and sequence verification of the vector with NGS.....	198
Figure C-11. Plasmid map of GFP fused with c-micL cloned downstream of pLux in pET22b (+) vector and sequence verification of the vector with NGS.....	198
Figure C-12. Plasmid map of GFP fused with c-micL cloned at the downstream of pLux and constitutively expressed LuxI in pET22b (+) vector and sequence verification of the vector with NGS.....	199
Figure C-13. Plasmid map of IPTG inducible micL sRNA and aTc inducible IL-1Ra gene in pET22b (+) vector and sequence verification of the vector with NGS.	199
Figure C-14. Plasmid map of IPTG inducible micL sRNA and aTc inducible ALP (PhoA) gene in pET22b (+) vector and sequence verification of the vector with NGS.	199
Figure C-15. Plasmid map of aTc inducible ALP (PhoA) gene in pET22b (+) vector and sequence verification of the vector with NGS.	200
Figure C-16. Plasmid map of aTc inducible ALP (PhoA) gene fused with YebF gene in pET22b (+) vector and sequence verification of the vector with NGS.	200
Figure C-17. Plasmid map of IPTG inducible CRISPR Array in pET22b (+) vector and sequence verification of the vector with Sanger.....	200
Figure C-18. Plasmid map of IPTG inducible CRISPR Array and E1 gRNA in pET22b (+) vector and sequence verification of the vector with Sanger.	201
Figure C-19. Plasmid map of IPTG inducible CRISPR Array, and E1 and E2 gRNA in pET22b (+) vector and sequence verification of the vector with Sanger.....	201
Figure C-20. Plasmid map of IPTG inducible CRISPR Array, and E1 and g9 gRNA in pET22b (+) vector and sequence verification of the vector with Sanger.....	201
Figure C-21. Plasmid map of IPTG inducible CRISPR Array, and E1, g9 and g10 gRNA in pET22b (+) vector and sequence verification of the vector with Sanger.	202
Figure C-22. Plasmid map of IPTG inducible CRISPR Array, relB and E1, g9 and g10 gRNA in pET22b (+) vector and sequence verification of the vector with NGS.	202
Figure C-23. Plasmid map of pMUT1 vector with sfGFP expression cassette.	202

Figure C-24. Plasmid map of pMUT2 vector with sfGFP expression cassette.	203
Figure C-25. Plasmid map of IPTG inducible CRISPR Array, and E1 and gLacI gRNA in pET22b (+) vector	203
Figure C-26. Plasmid map of IPTG inducible micL sRNA and aTc inducible 2Rs15d nanobody with DsbA-ss in pET22b (+) vector and sequence verification of the vector with NGS.....	203
Figure C-27. Plasmid map of IPTG inducible micL sRNA and aTc inducible 2Rs15d nanobody with gIII-ss in pET22b (+) vector and sequence verification of the vector with NGS.....	204
Figure C-28. Plasmid map of IPTG inducible micL sRNA and aTc inducible 2Rs15d nanobody with pelB-ss in pET22b (+) vector and sequence verification of the vector with Sanger.	204
Figure C-29. Plasmid map of IPTG inducible micL sRNA and aTc inducible 2Rs15d nanobody with lpp-ss in pET22b (+) vector and sequence verification of the vector with NGS.....	204
Figure C-30. Plasmid map of IPTG inducible micL sRNA and aTc inducible 2Rs15d nanobody with LTB-ss in pET22b (+) vector and sequence verification of the vector with NGS.....	205
Figure C-31. Plasmid map of IPTG inducible micL sRNA and aTc inducible 2Rs15d nanobody with MalE-ss in pET22b (+) vector and sequence verification of the vector with NGS.....	205
Figure C-32. Plasmid map of IPTG inducible micL sRNA and aTc inducible 2Rs15d nanobody with ompA-ss in pET22b (+) vector and sequence verification of the vector with NGS.....	205
Figure C-33. Plasmid map of IPTG inducible micL sRNA and aTc inducible 2Rs15d nanobody with TorA-ss in pET22b (+) vector and sequence verification of the vector with NGS.....	206
Figure C-34. Plasmid map of aTc inducible Nanoluciferase with ssREA-2 signal peptide in Lachnospiraceae shuttle vector and sequence verification of the vector with Sanger.	206
Figure C-35. Plasmid map of aTc inducible Nanoluciferase with ssREA-3 signal peptide in Lachnospiraceae shuttle vector and sequence verification of the vector with Sanger.	206
Figure C-36. Plasmid map of aTc inducible Nanoluciferase with ssREA-9 signal peptide in Lachnospiraceae shuttle vector and sequence verification of the vector with Sanger.	207
Figure C-37. Plasmid map of aTc inducible Nanoluciferase with ssREA-10 signal peptide in Lachnospiraceae shuttle vector and sequence verification of the vector with Sanger.	207

Figure C-38. Plasmid map of aTc inducible Nanoluciferase with ssREA-11 signal peptide in Lachnospiraceae shuttle vector and sequence verification of the vector with Sanger.	207
Figure C-39. Plasmid map of aTc inducible Nanoluciferase with ssREA-12 signal peptide in Lachnospiraceae shuttle vector and sequence verification of the vector with Sanger.	208
Figure C-40. Plasmid map of aTc inducible Nanoluciferase with ssREA-13 signal peptide in Lachnospiraceae shuttle vector and sequence verification of the vector with Sanger.	208
Figure C-41. Plasmid map of aTc inducible Nanoluciferase with ssREA-14 signal peptide in Lachnospiraceae shuttle vector and sequence verification of the vector with Sanger.	208
Figure C-42. Plasmid map of aTc inducible IbpA display shuttle 379-607 variant with ssREA-3 signal peptide and Nanoluciferase in Lachnospiraceae shuttle vector and sequence verification of the vector with Sanger.	209
Figure C-43. Plasmid map of aTc inducible IbpA display shuttle 510-607 variant with ssREA-3 signal peptide and Nanoluciferase in Lachnospiraceae shuttle vector and sequence verification of the vector with Sanger.	209
Figure C-44. Plasmid map of aTc inducible IbpA display shuttle 560-607 variant with ssREA-3 signal peptide and Nanoluciferase in Lachnospiraceae shuttle vector and sequence verification of the vector with Sanger.	209
Figure C-45. Plasmid map of aTc inducible IbpA display shuttle 585-607 variant with ssREA-3 signal peptide and Nanoluciferase in Lachnospiraceae shuttle vector and sequence verification of the vector with Sanger.	210
Figure F-1. Microscope images of holes in the migration assay.	212

LIST OF TABLES

Table 2-1. Blood biochemistry results for main indicators of liver and kidney functions in C57BL/6 mice on the 14th day following intraperitoneal rGFT administration.	52
Table 2-2. Hematologic parameters in C57BL/6 mice on the 14th day following intraperitoneal rGFT administration.	53
Table 5.1. 5000X vitamin solution.....	140
Table 5.2. YPG Media.....	140
Table 5.3. TYG salts solution.....	141
Table 5-1. List of signal peptide candidates selected for secretion efficiency test in Lachnospiraceae species.....	148
Table B-1. Protein or DNA sequences of genes in the thesis.....	189
Table D-1. Q5 polymerase PCR conditions.....	210
Table D-2. Gibson Assembly mix recipe (1.33x).....	210
Table D-3. 5x isothermal buffer recipe	210
Table D-4. T4 ligation reaction recipe	211
Table D-5. Restriction Digestion Reaction.....	211
Table F-1. Proteins detected only in the intracellular soluble fraction	212
Table F-2. Proteins detected only in the cell free supernatant	214
Table F-3. Proteins detected both in the intracellular soluble fraction and cell free supernatant	214

CHAPTER 1

Introduction

Biotechnological drugs, generally used interchangeably with biopharmaceuticals, cover large classes of pharmaceutical modalities whose production involves recombinant living organisms. Owing to their biological origin, biopharmaceuticals can target, activate, or inhibit a variety of biological processes individually and/or sequentially; therefore, these drugs often provide intricate molecular actions to treat diseases whose management is not possible with traditional small molecules. [1] Since the FDA's approval of the first biopharmaceutical product, recombinant insulin for the treatment of type one diabetes, in 1982, biotechnological drugs revolutionized modern medicine by offering therapy options for various diseases, including cancer, auto-inflammatory diseases, rare genetic conditions, and metabolic syndromes. [2] Currently, there are 541 FDA and/or EMA-approved biopharmaceutical products with 435 unique substances encompassing recombinant proteins such as antibodies, hormones, and enzymes, nucleic acid therapeutics, viral vectors, and cell therapy agents. [3] (Figure 1-1)

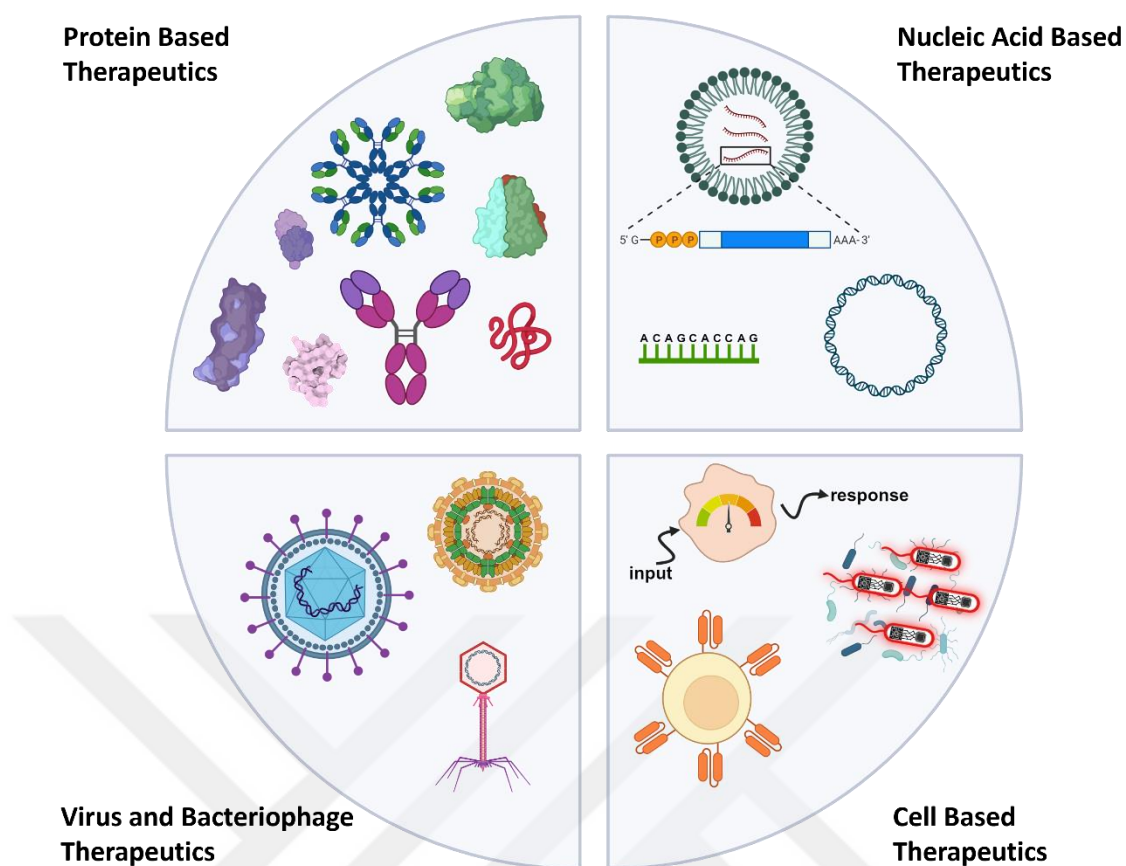


Figure 1-1. Different types of biopharmaceutical substances are currently being developed or approved by regulatory agencies.

Technological advancements in genetic engineering coupled with cumulative knowledge in disease biology fasten the pace for the development of novel biopharmaceuticals. Especially, repurposing the CRISPR-Cas9 system for genome editing and the cost-effectiveness of DNA sequencing and synthesis technologies enable the adaptation of synthetic biology approaches in biopharmaceutical research. [4]–[7] Biomolecules and cells can be programmed to sense a particular signal and articulate a sophisticated response based on the received signal with synthetic biology approaches. These kinds of “smart” drugs have the potential to change current treatment regimens for many diseases and present an opportunity for the treatment of previously incurable diseases.[8]

In the following introductory chapter, I will briefly explain biopharmaceuticals categorized into four different groups in comparison with their pros and cons. Meanwhile, I will also present the efforts to develop “smart” biopharmaceuticals with synthetic biology approaches. This literature review offers a general background on biopharmaceuticals currently in the market as well as efforts to develop new “smart” biotechnological drugs through synthetic biology. At the end of the chapter, I will summarize the scope and the aim of the thesis. The organization scheme of the following chapters will also be given.

1.1. Protein-Based Therapeutics

Proteins are the most diverse and functional biomolecules found in nature. Three-dimensional folding of proteins provides complex interactions to exert desired and specific functions such as inhibiting malfunctioning molecular pathways, activating cascade enzymatic reactions, degrading metabolites, and labeling particular cells for immune recognition. Therefore, numerous proteins are produced as therapeutical substances for the treatment of different diseases. [9], [10]

Due to their intricate chemical and three-dimensional structures, proteins must be produced biologically because the chemical synthesis of these gigantic molecules is not possible with current technologies. The necessity of the use of biological organisms in the production of protein therapeutics makes the manufacturing processes complicated with critical steps exceeding 5,000. [11] In addition, the delicate structure of proteins is prone to denaturation even with small environmental fluctuations which obstruct the long-term storage and transportation of protein therapeutics. Despite these drawbacks of protein therapeutics on production, storage, and transportation, the combination of industrial know-how on protein therapeutics

analysis and production and the versatility of proteins themselves leads to a bloom in the number of approved protein-based drugs. Currently, therapeutical proteins constitute more than 50% of all biopharmaceuticals commercialized as of 2023, whose market size is expected to reach 400\$ billion dollars. [3]

In this section, major classes of protein therapeutics, including antibodies, cytokine/hormones, and enzymes, and studies aiming to improve their effectiveness through molecular engineering will be summarized as well as some miscellaneous proteins, such as lectins and Cas proteins, will be outlined. [10]

1.1.1. Antibodies and Synthetic Antibody-like Scaffolds

Antibodies are the molecular arsenals of the immune system produced exclusively by the B cells against assaulting external substances, e.g., pathogens and their toxins. The three-dimensional structure of antibodies resembles the “Y” letter composed of variable regions at the upper tips of the structure and the crystallizable fragment (Fc) at the bottom tip. The variable regions of antibodies, also called antigen-binding fragments, contain amino acids dictating the antigen(s) that the antibody can interact with. [12] On the other hand, the Fc region contains the necessary molecular clues for interactions with other proteins and cells in the body. Based on the type of the Fc region, antibodies can be categorized into different classes, namely IgG, IgA, IgD, IgE, and IgM. [13] (Figure 1-2a) Upon binding to an antigen, different classes of antibodies trigger distinct downstream events such as targeted cellular toxicity, receptor activation, inhibition or degradation, ligand degradation or blockade, and agglutination of particles. [14] (Figure 1-2b)

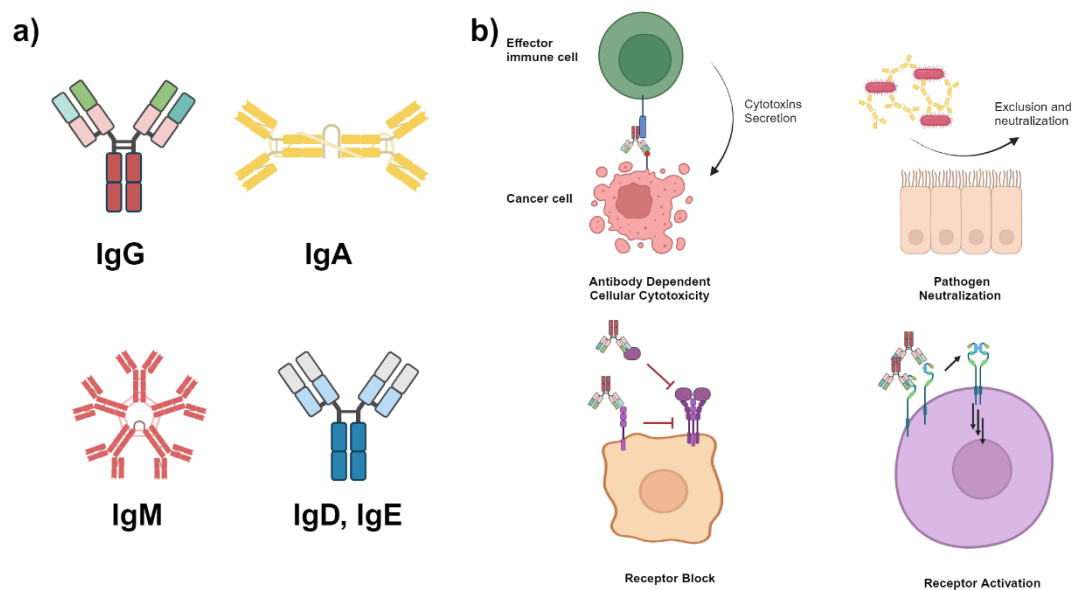


Figure 1-2. Different types of antibodies naturally produced by B cells and their activities. a) Naïve B cells produce IgM and IgD antibodies prior to the antibody class switch. Subsequent to the B cell activation, isotype switch occurs resulting the expression of IgG, and IgE, and IgA. IgG, IgD, and IgE antibodies are dimeric proteins. Meanwhile, IgA and IgM forms multimeric structures with four or more subunits. b) Antibodies offer several unique mechanisms to interfere for disease treat mechanisms. (upper left) IgG antibodies can be used to label cancer cells for antibody dependent cellular cytotoxicity induced by immune cells. (upper right) In the gut, immune cells secrete IgA antibodies to neutralize pathobionts and to prevent their colonization. (lower left) Antibodies can block the activation of receptors by preventing oligomerization of receptor subunits or hindering the binding of ligands to their cognate receptors. (lower right) Antibodies can be used as agonists for receptor to induce oligomerization for receptor activation.

Among all types of antibodies, IgG class antibodies constitute the majority of the approved products in the market for a great variety of indications, but the development of other class antibodies with high valency, e.g., IgM and IgA as

therapeutics, are also emerging. [15], [16] The pipeline for therapeutic antibody discovery consists of several steps, including target selection through basic disease biology, antibody screening to find binders, validation of the antibody-antigen interaction, and validation of downstream events upon antibody binding. There are lots of efforts to improve and fasten each step of the antibody discovery process. For instance, recent advancements in DNA synthesis technologies enable to circumvent *in vivo* immunization of live animals to select optimum binders through the implementation of *in vitro* screening of synthetic fab libraries. Such libraries are used in different studies to select antibodies against SARS-Cov-2 for neutralization [17] and GLP-1R for receptor blockade. [18] In addition, AI technologies can further improve the screening process by enabling the filtering of high variant libraries *in silico* for better antibodies in terms of solubility, stability, and immunogenicity. In a study, it was shown that a trained CNN-based algorithm was able to generate trastuzumab, a clinically used antibody for cancer therapy, variants with higher thermal stability as well as a lower immunogenic profile compared to the original molecule. [19] Cell-free expression technologies can also offer a platform to produce lead candidates in small amounts for further characterization at low cost and less time compared to traditional mammalian expression systems. [20] All these new technologies accelerate antibody discovery and commercialization from target selection to lead characterization by decreasing the costs and time required to screen and select the optimum antibody for the therapeutic hypothesis.

The capabilities, i.e., executed molecular actions, of natural antibodies are limited due to inherent structures. To expand their functionality, antibodies are redesigned through protein engineering and synthetic biology. As an example, bispecific antibodies are constructed rationally to recognize two distinct epitopes. [21] Thus,

these molecules can invoke multiple molecular mechanisms at once, such as labeling a cancerous cell and concurrently activating the immune system against the cancer cell for killing. [22] Aside from multiple antigen recognition, the activity of antibodies can be restricted in specific sites in the body, such as tumor microenvironment, by implementing irreversible or reversible molecular switches. These “smart” antibodies can be used to target common antigens as their activity is controlled with an additional layer. [23], [24]

Apart from limited capabilities, the natural structure of antibodies harbors other disadvantages. For example, the bulky structure of human antibody classes prevents accessing some epitopes as well as deep tissue penetration due to steric hindrance. In addition, post-translational modifications on antibodies necessitate production in mammalian expression systems whose cultivation processes are cumbersome and expensive. Synthetic antibody mimetic proteins are developed to surmount these disadvantages. [25] Structurally, antibody mimetic peptides resemble fab domains by consisting of a constant protein scaffold and distinct variable regions for antigen recognition. [26] Antibody-like proteins are smaller relative to full-length antibodies and lack of any glycosylation; thus, recombinant production of these can be done by using cheaper hosts such as *E. coli*. [27] Small molecular size of these antibody mimetic proteins facilitates deep tissue penetration, which is a desired feature to target dense tumors. [28] Also, multiple epitopes in closed proximity can be targeted by designing antibody mimetics with multiple domains owing to their small size. [29] Although antibody mimetic proteins suffer from low serum half-life and immunogenicity *in vivo*, there are efforts to improve their drawbacks, such as fusing serum albumin binding mimetics to increase their serum half-life. [30] Owing to the ease of production and small size of antibody-like scaffolds, they offer a unique

platform to develop therapeutical agents for several indications, including cancer treatment and viral infections.

Antibodies and antibody-like proteins are marvelous biomolecules that present many opportunities to develop therapeutical modalities for diseases, and engineering tailor-made antibodies and antibody-like proteins with augmented functions can broaden the use of antibodies as therapeutics in different indications.

1.1.2. Cytokines and Protein Hormones

Cytokines and protein hormones are secreted by cells to control metabolic and immunological actions in the body. Cytokines primarily act on cells in the immune system; meanwhile, a great variety of cells can respond to protein hormones whose function is to control important metabolite concentrations such as glucose in the blood as well as to regulate the differentiation and growth of certain cells. [31], [32] Although both protein hormones and cytokines are mainly used as therapeutics to exploit the function of natural proteins expressed in the body, few proteins from both classes are intensely engineered to tune their natural properties, such as serum half-life, acting time, and restriction of their activity on specific cell types. [33], [34]

Amidst the protein hormones, insulin is the most studied and engineered therapeutical hormone since it is a vital drug for the treatment and management of diabetes. Naturally, insulin is composed of two relatively short peptide chains linked with cystine bridges and is secreted by the pancreas to control glucose concentration in the blood. [35] Slow penetration from tissue to blood and short serum half-life of natural insulin leads to life-threatening unwanted effects such as hypoglycemia when administered externally. [36] To improve the pharmacodynamics and pharmacokinetics of natural human insulin, many insulin mutants were engineered

that contain several key mutations. These engineered versions of insulin can linger in the body longer without having a significant peak or act right after injection to control glucose. [37] Currently, the mutant versions of natural insulin provide key features such as long serum half-life or fast response time for the management of diabetes.

Cytokines are vital biomolecules that function as communicators between immune cells. The function of many cytokines is contingent upon both the specific cell types with which they engage and the tissues in which cells are located. For instance, IL-6 elicits pro-inflammatory effects on the skin, whereas it induces anti-inflammatory effects in the gut. [38] Due to their double-edged functions in different tissues, wide distribution in the body of administered causes severe side effects. To improve the clinical effectiveness of cytokines, tremendous efforts are put on the rational design of cytokine therapeutics with biased activity on desired cells. [39] Rational designing IL-2 to diminish its effects on T_{reg} cells fosters a therapeutical opportunity to use mutant IL-2 for cancer therapy. [40] In addition, IL-10 was engineered to decouple its anti-inflammatory and pro-inflammatory activities from each other to create a synthetic version of IL-10 for decreasing inflammatory responses in autoimmune diseases. [41]

As seen, protein hormones and cytokines are useful therapeutical candidates for the treatment of different illnesses, yet some of their intrinsic properties hinder their therapeutical capabilities. Molecular engineering of cytokines and protein hormones instigated life-saving therapies for diabetes and can ameliorate therapeutic strategies pertaining to cancer and autoimmune diseases.

1.1.3. Recombinant Human Enzymes

Enzymes are biocatalysts proteins that convert essential metabolite for the cells to properly function. Enzymes can be used for therapeutical purpose in great variety of indications including are rare metabolic disorders [42]–[45], cancer [46], autoimmune disease [47], sepsis [48] and cocaine overdose [49]. Although, the majority of therapeutic enzymes are the recombinant version of native enzymes, genetically engineered enzymes are emerging for interventions in some indications such as microbial infections with high resistance. [50] In addition, synthetic enzymes with non-biological amino acids can offer additional chemistries to improve capabilities of the natural enzymes. [51] Engineering novel enzymes with better stability and higher specific activity hold substantial promises to improve the remedies for wide range of illnesses.

1.1.4. Miscellaneous Proteins

Protein space, when considering sequence combinations and folding, is staggeringly immense. Organisms found on earth have been adapting to various conditions for billions of years to survive. The adaptation of cells results in the evolution of millions of proteins which confer fitness by their functions to the organisms in the conditions where they reside. The vast sources of proteins found in nature can be explored to discover proteins with therapeutic potential for human diseases, even though the primary function of these proteins might be completely irrelevant in the context of natural organisms in which they are expressed.

For instance, class II type Cas proteins is RNA guided endonucleases encoded in bacteria to fight against bacteriophage infection. [52] Subsequent to their discovery and molecular understanding how they work, Cas proteins are repurposed to edit [53]

and insert DNA sequences [54] in genomes and control the gene expressions *in vivo* [55]. Combined with protein delivery technologies, Cas proteins revolutionized gene therapy. Similarly, many plant and algae species encode lectins proteins which can bind sets of carbohydrate molecules. [56] These lectin proteins can be purposed to develop antiviral as many of viral proteins are heavily glycosylated. [57], [58] In addition, carbohydrate affinity of lectin are exploited to label and kill cancerous cells. [59]

Proteins encoded in various organisms can hold therapeutical values for addressing human diseases. Although mentioned studies above serendipitously discover these proteins, current technologies offer to screen biomes with high-throughput experimental setups. Thus, novel therapeutical proteins can be identified from other species with concrete hypothesis.

1.2. Nucleic Acid Based Therapeutics

Nucleic acid therapeutics are emerging drug modalities encompassing DNA, mRNA, siRNA, microRNA, and synthetic anti-sense oligonucleotides (ASO). As the universal building blocks of all cells, nucleic acids are biomolecules that store and control the expression of genetic information, thereby offering novel mechanisms of actions beyond druggable targets within the cells to develop medications for diseases. [60] This, in turn, holds great promise for the development of effective medications for diseases. For instance, siRNA, microRNA and ASO based therapeutics interfere the translation of mRNA to mature protein, leading to reduction of translated protein levels within the cells without need of permanent genome modification; meanwhile, ASO are also utilized to change splicing patterns of pre-mRNA to produce functional proteins. [61]–[63] (Figure 1-3a, and 1-3b) RNA

aptamers can be designed to bind intracellular proteins which offer to interfere with the formation of certain protein complexes in malfunctioning pathways. (Figure 1-3c) On the other hand, mRNA therapeutics are designed to express proteins once delivered into cells, providing a platform for intervening in intracellular processes. [64] (Figure 1-3d) The novel molecular modes of action harnessed by nucleic acid therapeutics hold the potential to revolutionize disease treatment paradigms. In this section, mRNA therapeutics will be presented in depth as it holds necessary complexity to program “smart” therapeutics.

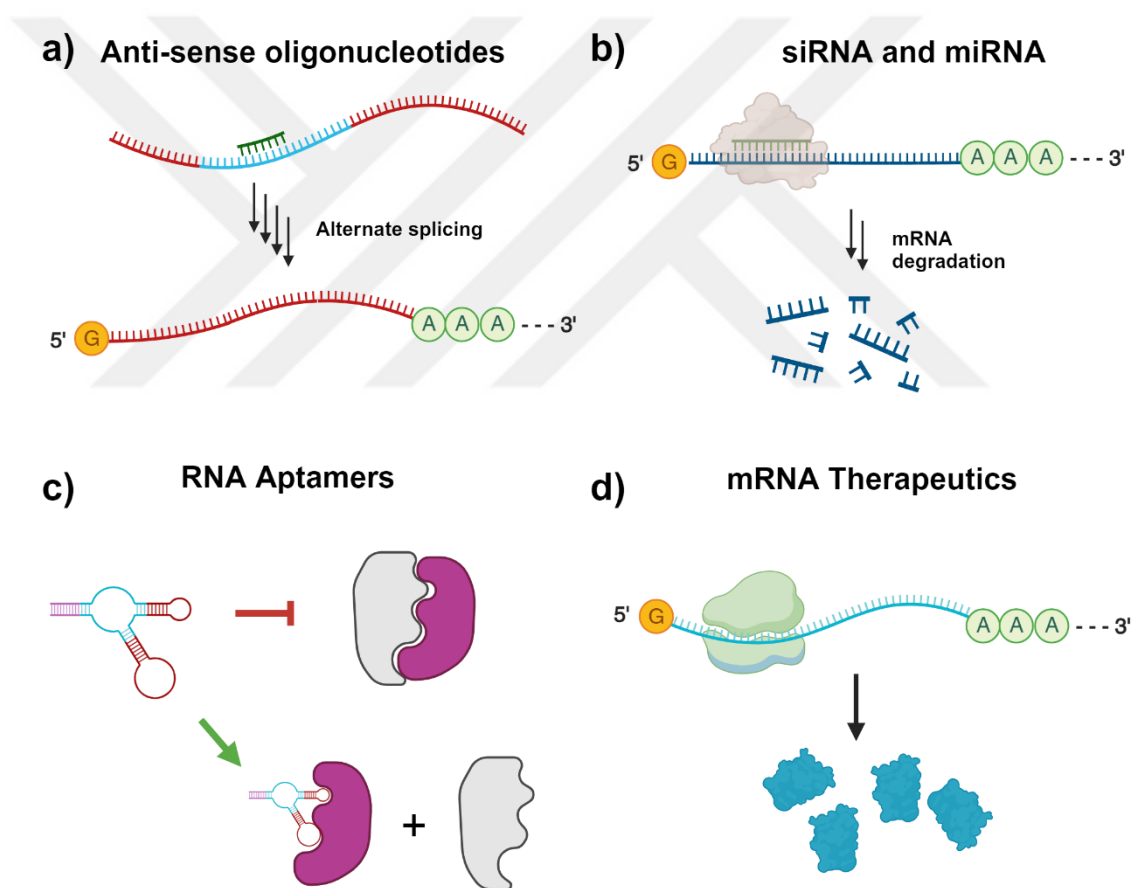


Figure 1-3. Nucleic acid-based therapeutics and their molecular mode of actions. a) Anti-sense oligonucleotides (ASO) are single stranded nucleic acid molecules complementary to certain sequences on disease causing RNAs. Upon binding, ASO can induce alternate splicing products from pre-RNA thereby

preventing formation of mRNA that encodes a toxic or unfunctional form of disease related proteins. b) Similar to ASO, siRNA and miRNA are also designed to be complementary to a particular RNA but instead of interfering splicing process, they are utilized to induce mRNA degradation by intracellular RNA degradation machineries thereby eliminating the disease-causing proteins. c) RNA aptamers are complex RNA molecules with distinct 3-D structures. They are designed to bind particular proteins to block their activity such as complex formation which offers to silence malfunctioning pathways. d) mRNA therapeutics are utilized to express proteins transiently inside cells without any genome modifications. Current applications of mRNA therapeutics are envisioned in vaccination, and immunotherapy.

1.2.1. mRNA Therapeutics

mRNA molecules, acronym for messenger RNA, are composed of a single stranded nucleic acid chain transcribed from DNA to be translated into proteins. In eukaryotes, mRNA molecules are heavily processed prior to cytoplasmic translocation from the nucleus. Eukaryotic mRNA (e-mRNA) molecules contain chemically distinct 5' end, untranslated regions (UTR) at both 5' and 3', the translated region and poly(A) tail at the 3' end. Each part of e-mRNA carries vital importance for protein translation and stability of the mRNA. Modified 5' end of e-mRNA protects mRNA molecules from degradation by nucleases as well as cloaks the molecules from innate immune system sensors whose function is to detect unmodified 5' ends of viral RNAs. Both 5' and 3' UTRs of e-mRNA contain molecular clues that regulate translation rate from the RNA molecules. Poly(A) tails also interact with other proteins to stabilize the mRNA and protect it from nucleases.

The translated region contains codon triplets to tRNA for protein synthesis which have direct effect on translation rate. [65] (Figure 1-4)

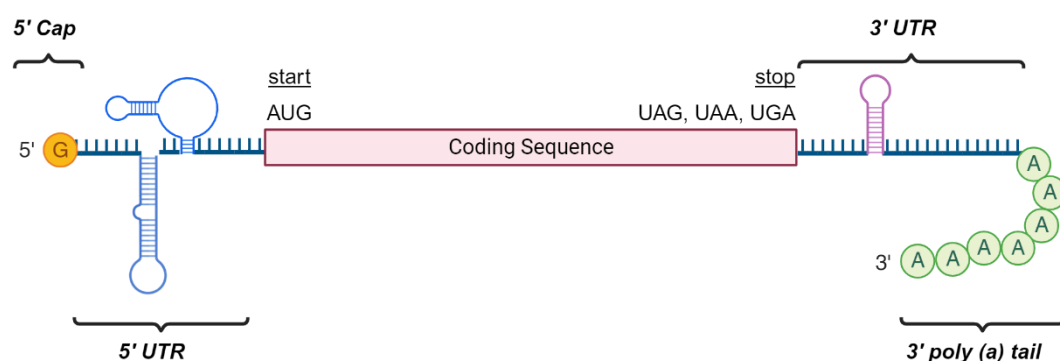


Figure 1-4. Structure of processed eukaryotic mRNA

mRNA therapeutics consists of two components which are therapeutical mRNA molecules and the delivery shuttle. To produce therapeutical mRNA molecules (t-mRNA), the DNA molecule encoding the t-mRNA is constructed by adding the bacteriophage T7 promoter upstream of t-mRNA coding region for *in vitro* transcription. [66] The t-mRNA coding region contains all necessary components of the native e-mRNA including the UTRs and poly(A) tail. The 5' modified end is incorporated into t-mRNA structure during the *in vitro* transcription by using synthetic cap analogs such ARCA [67] and CleanCap [68] that mimics the natural cap structure. In addition to 5' analogs, modified nucleosides such as pseudo-uridine are used to disguise the t-mRNA from innate immune system components. [69] The transcribed t-mRNA molecules are purified to eliminate all IVT reaction components and byproducts such as short double stranded RNAs which can cause unwanted immune activation. [70] The purified t-mRNA are loaded into carrier shuttles for intracellular delivery. To the date, commercialized mRNA therapeutics are mainly utilized lipid nanoparticles (LNP) for delivery of t-mRNA. [71]

Developing 'intelligent' mRNA therapeutics capable of selectively targeting specific cell types and making autonomous decisions represents a highly promising avenue in the advancement of disease treatment. However, t-mRNA therapeutics currently available in the commercial market are hindered by lack of precise targeting and limited durability within cells. [72] Notably, there are significant efforts to program the selective delivery and/or translation of t-mRNA to certain cells and tissues. In one aspect, the translation of t-mRNA molecules is constrained by inserting miRNA binding within the UTR regions. [73] In another approach, lipid nanoparticles laden with the t-mRNA molecules are precisely guided to desired cells by immobilizing antibodies on the surface of the LNP. [74] To control the translation of administrated mRNA *in vivo*, a post-transcriptional switch that is controlled by a small molecule is designed to titrate the protein expression from the t-mRNA. [75] Aside from improving the specificity and tunability of t-mRNA, durability of t-mRNA within cells is enhanced by engineering RNAs to form circular molecule thereby preventing any nuclease attacks at the free ends. [76] As seen from the studies, the going endeavors can address current limitations of the t-mRNA capabilities in terms of specificity and durability. Concurrently, logic gate-based decision making circuits built in the t-mRNA structures are proposed in the literature. [72] Collectively, these innovative approaches pave the way for creating effective “smart” t-mRNAs thereby offering new paths for disease treatment.

1.3. Virus and Bacteriophage Based Therapeutics

Viruses are self-assembled infectious particles that can only replicate within a viable host as they lack molecular machineries and resources such as ribosomes, and tRNAs. Viral particles encapsulate their genetic material, which encodes sufficient

information for successful host invasion and self-replication, in the form of DNA or RNA. Aside from the genetic materials, viral particles are composed of proteinaceous coat that shields the viral genome and, in some cases, surrounding lipid bilayer adorned with proteins. Despite their minimal biochemical activity outside of host cells, nearly all organisms in Earth's biome fall prey to at least one type of virus. As one of the most diverse biological entities within the biome, viruses possess numerous molecular tools, including various proteins, to efficiently invade their respective hosts that have been evolved during the co-adaptation with their host. [77], [78]

Thus, in the therapeutical viewpoint, viruses possess a great variety of molecular instruments and features that can be engineered to develop new-to-nature biopharmaceuticals with unique mechanisms of action. In this section, engineered viral particles will be outlined and discussed in two different categories: human viruses as viral vectors for therapies, and engineered bacteriophages as antimicrobials.

1.3.1. Human Virus as Viral Vectors for Therapies

Human-infecting viruses have long been regarded as formidable pathogens, armed with molecular tools that allow them to assault cells within the body by often evading both innate and adaptive immune responses. Nonetheless, understanding virus biology has facilitated the domestication of the wild-type viruses, leading to the development of safe therapeutical vectors based on these viruses. Aside from classical approach that involves weakening of viruses through physical processes for vaccination, wild-type viruses are genetically engineered for numerous applications. Rationally designed virus-based vectors are utilized in therapies that involve

reprogramming the adaptive immune system, eliminating cancer cells, and delivering nucleic acid payloads for gene therapies. (Figure 1-5)

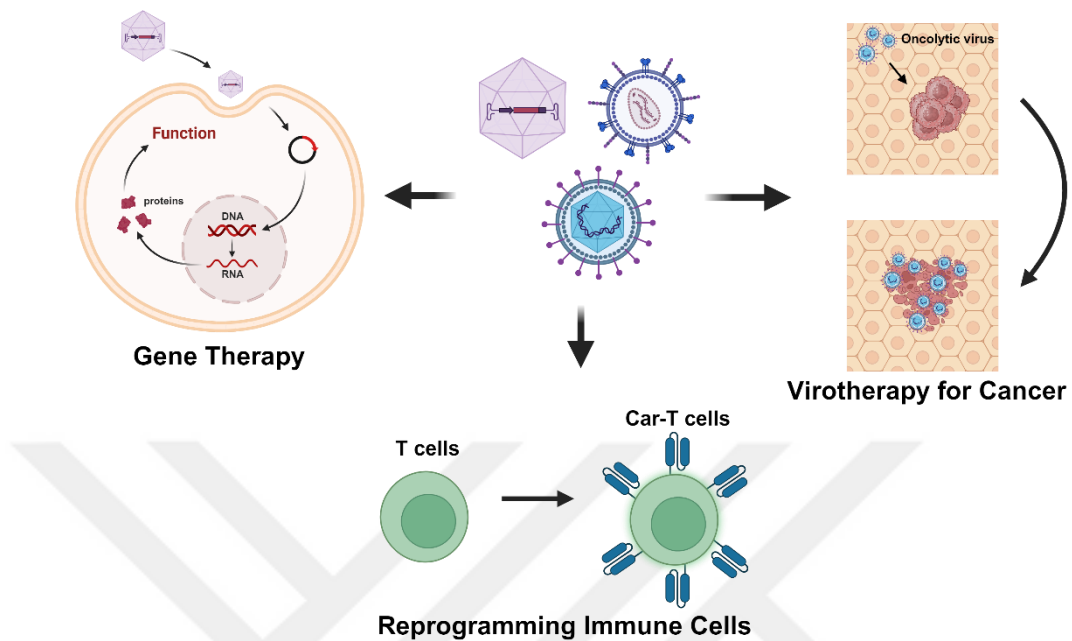


Figure 1-5. Human infecting viruses are utilized as vectors and therapeutic modalities for gene therapy, reprogramming the immune cells and virotherapy.

Adeno-associated viruses and adenoviruses are mainly used to deliver genetic payloads into the cells for expression of therapeutic proteins. Current applications of these viruses as vector encompass gene therapies for hereditary disease such as SMA, and vaccination against pathogens. Lentiviral vectors are utilized to incorporate CAR encoding constructs into the genome of immune cells. The natural ability of oncolytic viruses is harnessed to target cancer cells.

Viruses inherently have high immunogenic activity that activates both antibodies based and cellular based adaptative immune responses without need of an adjuvant molecule. Thus, many viruses including adenoviruses, poxviruses and lentiviruses are engineered to introduce foreign antigens thereby stimulating the adaptive immune responses against them. [79] Yet, the preexisting immunity among the

human population against vector viruses can result in either excessive or null immune responses. To circumvent this, viral vectors are constructed from the virus species that infect the evolutionary close relative of humans such as chimpanzee. [80] In addition, replication of the viral vectors within cells following a successful infection is hindered by deleting necessary elements from genome to ensure the safety of the vectors. [81] Thanks to advances in viral vector technology, viral vectors have been engineered to deliver antigens from lethal infectious agents, like rabies, in order to stimulate an adaptive immune response thereby offering protection against actual rabies infections. [82] Beyond antigen delivery, infectious virus strains can be genetically attenuated by tightly controlling their replication within the production cell line using molecular switches. For instance, a genetically modified influenza virus strain was constructed by switching its native protease linkers with synthetic protease linkers which made the virus incapable of replicating in cell lines other than the designated production cell line. In the study, the engineered virus could induce a potent humoral response without any *in vivo* replication. [83]

Aforementioned viral vectors were also designed to deliver functional proteins either *in vivo* to cells within the body for restoring the malfunctioning pathways due to genetic defects or *ex vivo* to immune cells for cancer therapy. For *in vivo* gene therapies, adenoviruses and adeno-associated viruses are commonly used as vehicles to carry nucleic acids payloads some of which is currently approved for treatment of rare genetic conditions such as type I SMA and Leber's congenital amaurosis. [84], [85] Meanwhile, lentiviruses are utilized for *ex vivo* engineering of T cells for cancer therapy. Integration ability of lentiviruses provide long-lasting expression of the therapeutical cargo protein, the chimeric antigen receptor (CAR), in T-cells which can induce cell death upon recognition by CARs. [86] Both of engineered virus

vectors are tinkered to carry only the payload nucleic acid with minimal additional elements to prevent post-infection replication and to increase the nucleic acid loading capacity by using helper cells. Aside from safety issues, tropism of viral vectors is tuned through engineering viral surface proteins. Notably, capsid protein of AAV based vectors are designed to implement molecular switches to control infections as well as to construct hybrid AAV-bacteriophages vectors. [87]

Oncolytic viruses such as herpes simplex selectively infect and replicate inside the cancerous cells meanwhile non-neoplastic cells are not affected by them. Preferential activity of oncolytic viruses on cancerous cells are harnessed for developing novel therapeutical vectors to treat cancer. In addition, vectors based on the oncolytic viruses can be augmented with genetic circuits to improve their functionalities. For instance, the anti-tumoral activity of these viruses can be amplified by incorporating tailored expression cassettes to produce cytokines for antitumor immune responses against melanomas. [88] To increase the precision of the anti-tumoral response of oncolytic viruses, genetically encoded biosensors to differentiate cancer cells based on the malfunctioning molecular pathways or miRNA profile are deployed within viral vectors. The biosensor for malfunctioning molecular pathways was used to differentiate cancerous cells from healthy cells through sensing the hyperactivity Erb pathways meanwhile miRNA biosensors were used to categorize cells with logic operations based on presence or absence of certain miRNAs [89], [90]

In summary, human viruses are excellent platforms to develop new therapeutical modalities especially for gene therapies and cancer treatments. The inherent ability of viruses to carry nucleic acid payloads into cells can be further exploited to develop advanced therapeutics which can sense and respond to molecular clues.

1.3.2. Engineered Bacteriophages

Bacteriophages are natural bactericidal agents that infect and lyse their cognate bacterial hosts during their replication. Coupled with high replication rate of their hosts, bacteriophages are highly mutable and diverse in the earth biome which provide a vast pool to isolate novel phages. [91] However, the therapeutic use of native bacteriophages as bactericidal agents is often limited to particular strains and are impractical in most of the cases because of high level of escaping mutants. [92] Genetic engineering of bacteriophages opens new avenues to exploit native abilities of phages for developing novel drugs. [93] For instance, host range of bacteriophages can be expanded through protein engineering of receptor binding domains of bacteriophages. Furthermore, bacteriophages can be reprogrammed to deliver nucleic acid payloads into their cognate host bacteria thereby enabling to express and/or knock-out genes inside the host bacteria.

The host range of bacteriophages are dictated by their receptor binding domains which are located either at the baseplate or at the tail fiber tip. The binding of receptor binding domains to their receptor structures, which can be proteins or carbohydrates, on host bacteria surface initiates infection cascade. To tune and expand the phage host-range, the receptor binding domains of phages are diversified by using molecular techniques such as domain shuffling and random mutagenesis. [94] For instance, T3 phage specific to enterobacteria was engineered to expand its host range by targeted mutagenesis of amino acids located in the interface between host receptor and phage host recognition domain. The resulted library of T3 phages prevented to evolution of escape host mutants. [95]

The inherent ability of phages to transduce their host bacteria with foreign DNA provides opportunities to reprogram the host bacteria for therapeutical and diagnostic purposes. The nucleic acids payload can be encoded to express potent toxins and enzymes to eliminate insensitive bacteria within the population. In a study, isolated bacteriophages to target wild type *E. coli* and *K. pneumoniae* obtained from urinary infections were engineered to express anti-microbial toxin which was secreted to extracellular space upon cell lysis. The engineered phages in the study were able to suppress and/or delay the evolution of space mutants. [96] In other study, bacteriophages were utilized as biosensors to sense presence of certain bacteria by delivering an expression cassette for nanoluciferase gene which would be expressed and released to environment following the infection. Authors were able to detect pathogens from urine by using nanoluciferase encoding bacteriophages. [97]

It is evident from studies that bacteriophages as bacteria infecting viruses offer a unique platform to fight against infections caused by multi antibiotic resistance bacteria species as well as to sense particular bacteria strains in complex media.

1.4. Cell Based Therapeutics

Cell based therapeutics are emerging biopharmaceuticals that involve genetically modifying living cells to address unmet clinical needs for intractable diseases. Living cells offer exceptional abilities for intervening disease mechanisms beyond the reach of any pharmaceutical modality. As the smallest functional compartmentalized unit of all living organisms, cells can orchestrate many molecular cascades based on the signal they received in a specific context. These inherent capabilities of cells can be amplified through synthetic biology approaches by implementing recombinant genetic systems into cells. By doing so, cells can be reprogrammed to record

temporal and spatial molecular clues related to disease states which enable cells to decide both the timing and amplitude for required responses to revert pathological conditions. Therefore, genetically programmed cells provide distinctive opportunities to overcome hurdles of other pharmaceuticals including the ability to deliver precise and timely responses, self-regulation, and a high level of specificity. [98]

The efforts to engineer therapeutical living cells fall into two different categories based on the cell chassis that is used: mammalian-based cell therapeutics and microbial cell-based therapeutics. Although the principles that are employed for both the categories are the same, the molecular tools for engineering and inherent characteristics of chassis differ substantially, between mammalian and microbial therapeutics. In this section, efforts, and strategies to engineer living therapeutics will be summarized for both microbial and mammalian-based chassis.

1.4.1. Engineered Human Cells

Cells within the human body can be exploited to perform therapeutical tasks through genetic reprogramming either *in situ* or *ex vivo* via introducing recombinant nucleic acids that encode the molecular circuits. Engineering human cells for therapeutical purposes comes with many benefits since these cells possess many relevant molecular components and biochemical networks which can be rewired to mitigate disease states. Furthermore, engineered human cells compared to others have a higher propensity to acclimate and seamlessly integrate themselves into the body's state of homeostasis, diligently monitoring for any potential resurgence of the disease conditions. [99] The potential benefits associated with human cells have driven the advancement of engineered human cell therapies for addressing conditions like cancer, metabolic disorders, and autoimmune diseases. For this purpose, genetic

circuits can be designed to differentiate certain cells from other cells by performing logic operations. (Figure 1-6a) In addition, prosthetic gene networks can envision to maintain the hemostasis by producing therapeutic metabolites in response to sensed metabolites (Figure 1-6b)

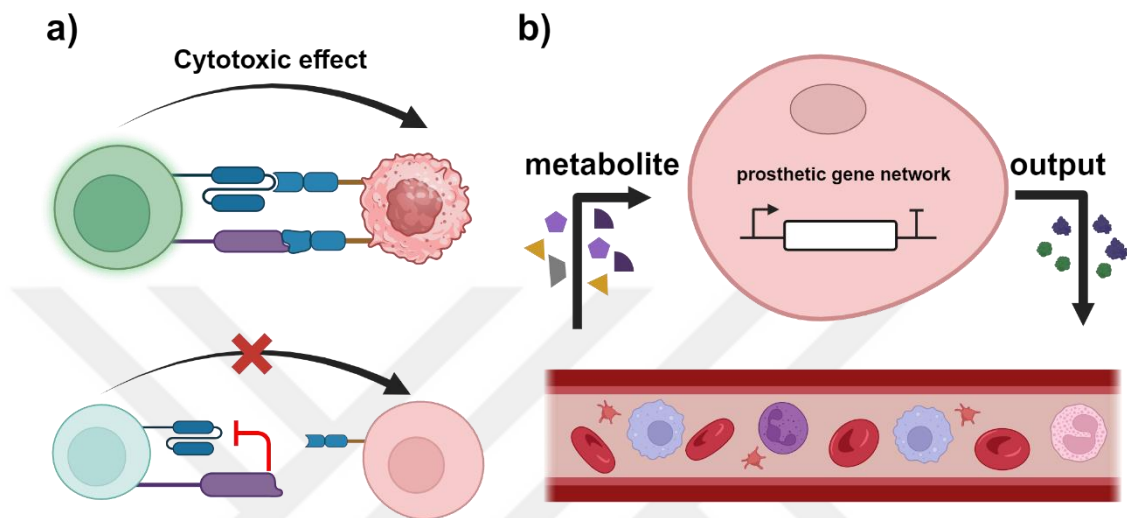


Figure 1-6. Engineered human cells to treat different diseases. (a) The immune cells can be reprogrammed to differentiate certain cells from other cells based on the surface markers. These circuits can be used to eliminate cells that are causing disease conditions such as cancer cells inside tumors or B-cells secreting toxic autoantibodies. b) Prosthetic gene network can be constructed inside somatic cells to maintain homeostasis for excess metabolites in diseases e.g., glucose in diabetes or uric acid in gout disease. The amount of toxic metabolite can be used from blood via biosensor implemented in the engineered cells and therapeutic response such as secretion of therapeutic proteins can be orchestrated upon sensing the excess metabolite.

Successful remission free treatment of hematologic cancers such B-ALL by engineered T-cells expressing chimeric antigen receptors (CAR) have fueled efforts

to develop reprogrammed immune cells for cancer therapy. Thus, immune cells are the key cell chassis for developing treatment of malignancies. However, the use of engineered immune cells in cancer therapy faces several challenges including poor activity of therapeutic cells in solid tumors due to immune suppression, the lack of unique targeting antigens for many types of cancer, and life-threatening side effects of engineered cells. To overcome these hurdles, a range of cell engineering strategies were employed focused on cancer cell targeting and therapeutic payload delivery. For instance, PD-1 locus from T-cells were utilized as the insertion site for CAR expression cassette which resulted in better clinical profile with lack of neurotoxicity compared to random integration. [100] In another study, myeloid cells were engineered to express IL-12 thereby countering the immunosuppression at the tumor environment which induced a potent anti-tumor response. [101] Furthermore, T cells were engineered to perform simple logic operations such as AND or OR functions enable the discriminating of healthy and tumors cells based on more one signal. [102]

Aside from cancer therapy, immune cells can also be engineered to cease acquired self-immunity of the body against to itself. For instance, T cells were engineered to express a variant of CAR which have an autoantigen instead of an antibody binding domain for eliminating the autoreactive B cells in pemphigus vulgaris patients. [103] In addition, regulatory T cells were constructed to surmount the immune rejection of kidney and liver following the transplantation. [104]

In addition to the immune cells, other human cells can be engineered to carry “prosthetic gene networks” to restore metabolic hemostasis of the body which is disturbed in many metabolic diseases. In one approach, a genetic circuit was

constructed to sense elevated uric acid in blood, and then trigger the expression of urate oxidase to reduce the uric acid level. The engineered cells harboring the synthetic circuit successfully lowered excessive uric acid concentrations in the blood when transplanted into mice using alginate gels. [105] Moreover, the same approach was employed to build synthetic circuits for regulating the glucose level in diabetic conditions and controlling thyroid hormone level in hyperthyroidism.[106], [107]

In summary, the utilization of engineered human cells as a foundation for therapeutic development is highly promising, primarily due to their innate capacity to persist within the body and their possession of crucial molecular components. While current research endeavors have concentrated on harnessing engineered cells for the treatment of cancer, metabolic disorders, and autoimmune conditions, these cells hold great potential to address therapeutical gaps for a spectrum of otherwise incurable diseases.

1.4.2. Engineered Microbial Cells

The human body is an ecological niche for a variety of microbes some of which are highly adapted to stably colonize certain organs e.g., gastrointestinal tract. The intricate interactions between microbes and the human body as the host have substantial effects on human health. Many chronic conditions such as diabetes, inflammatory bowel disease, and hypertension are associated with gut microbiota dysbiosis. Due to their ability to interfere with the host metabolism across multiple physiological axes, microbial species offer a novel platform for developing new therapeutic modalities. [108]

In the development of microbial therapeutics, microbial strains are selected based on the safety profile, its inherent capability and availability of genetic tools. Thus, the

efforts for engineering microbial therapeutics are focused on the few microbes which are highly characterized in molecular level and utilized in humans for long periods to ensure their safety. [109] Examples of such microbial chassis employed in therapeutic development include the probiotic *E. coli* strain known as Nissle 1917 (EcN) [110], lactic acid bacteria (LAB) [111], and the yeast strain *Saccharomyces boulardii* (*S. boulardii*) [112]. Although these microbes are tolerated by the immune system, they are lacking the necessary components to exert desired functions for interfering the disease mechanism. The immense diversity of microbes in the biome offers reservoir to mine functional parts for constructing the genetic circuit that can perform the desired action in the therapeutic chassis. With the help of genetic circuits, microbes can be reprogrammed to sense multiple signals at the disease milieu, and orchestrate an output based on the received signals. (Figure 1-7)

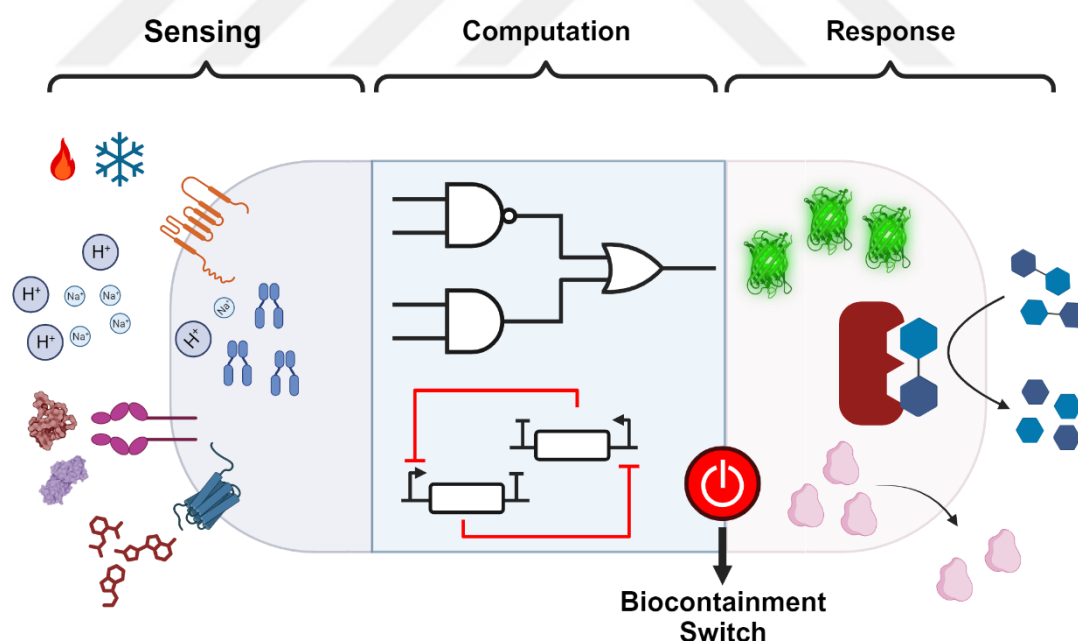


Figure 1-7. Topology of engineered microbes for therapeutical purposes. The microbes carry three modules which are sensing, computation and response modules. In the sensing module, different environmental signals such as temperature, pH, ions,

proteins, and small molecules are received and sensed through cell surface or intracellular sensor proteins. The received signal is transferred to the computation module in which response is determined based on the signals. Logic operations as well as memory recording can be performed with genetic circuits inside the cells. Lastly, cells orchestrate outputs which are gene expression in most cases as a response. In the response module, fluorescent proteins are widely utilized for diagnostics purpose whereas therapeutical proteins are expressed to degrade certain molecules or are secreted from cells to exert their activity. A biocontainment switch is constructed to eliminate engineered cells in case of a need and to prevent their spread to the natural environment.

LAB are often consumed as a part of human diet which involves nutritional products with fermented dairy ingredients. Thus, these microbes are considered generally safe for human use.[113] Owing to available genetic technologies for LAB, *L. lactis* was engineered to secrete an anti-inflammatory cytokine, IL-10, at the gut for mitigating inflammation. The engineered strain showed a positive effect on murine colitis models and was able to decrease the degree of inflammation in the model organism. [114] Aside from delivering effort proteins, the same bacteria strain was modified to sense *Vibrio cholera* by detecting the quorum sensing molecules secreted by *V. cholera* during the early stage of the infection. [115]

EcN strain has been used by humans as a probiotic supplement as well as a pharmaceutical product for certain indications for over 100 years. As a strain of *E. coli* which is common model for microbial synthetic biology research, EcN is adopted as a favored chassis for developing living therapeutics because of available ample information on its metabolism and tools for genetic manipulation. Living

therapeutics research that uses EcN strain as the base chassis covers treatment of a variety of diseases.[110] For instance, EcN was designed to degrade the dietary phenylalanine in the gut. Recently, in phase II of human clinical trials, the optimized version of engineered phenylalanine degrading EcN strain decreased the blood phenylalanine levels meaningfully in PKU patients. [116] Furthermore, there are studies showing that EcN can be engineered for treatment of other chronic metabolic diseases such as enteric hyperoxaluria and hyperammonemia. [117], [118] EcN strain is also engineered to deliver protein based therapeutics to both gut and tumors. [119]–[121]

S. boulardii is probiotic yeast strain as exploit in the treatment of gut intestinal pathologies such as diarrhea and ulcerative colitis.[122] Due to its evolutionary proximity with yeast organism *Saccharomyces cerevisiae* (*S. cerevisiae*), *S. boulardii* can be engineered with genetic tools developed for *S. cerevisiae*. [123] Thus, same as EcN strain, *S. boulardii* is a popular strain to engineer synthetic microbial therapeutics. For instance, an immunotherapy approach was designed to cure *Clostridioides difficile* infection by engineering *S. boulardii* to secrete neutralizing nanobodies against *C. difficile* toxins. [124]

In summary, synthetically designed microorganisms offer new avenues for treating complex diseases by exerting specific and tunable therapeutical activity to manipulate malfunctioning molecular pathways in diseases.

1.5. Thesis Scope, and Structure

In my thesis, I will explain our efforts on the development of protein-based biopharmaceuticals that can be used as prophylactic agents against viruses as well as

genetic technologies for designing microbial therapeutics during my time at Seker lab as PhD student. The goal of this thesis on a broader scale is to discover and/or engineer novel biopharmaceutical modalities in order to address unmet clinical needs for better therapies. I am hoping that the concepts, ideas, and methodologies presented in this thesis will lead to development of unique biotechnological drugs that can improve and widen the treatment options for many diseases in far future.

In chapter 2, I will highlight our work on developing prophylactic antivirals based on lectin proteins for SARS-CoV-2 and its emerging variants. The study was published in *ACS Infectious Diseases* in 2022 and it was a collaborative work of our group, Aykut Ozkul's group at Ankara Biotechnology Institute and Faculty of Medicine at Hacettepe. In the study, recombinant griffithsin (rGRFT), a lectin protein from red algae *Griffithsia*, was purified from *E. coli* to investigate its antiviral potential against ancestral strain (Wuhan) and emerging variants in 2022 (Delta and Omicron) of SARS-CoV-2. The rGRFT showed potent antiviral activity against all strains of SARS-CoV-2 as well as proven to be non-toxic, non-immunogenic and non-carcinogenic. The work demonstrated that lectin-based proteins can be harnessed to develop potent antivirals for current and future viral pandemics.

In chapter 3, I will demonstrate the genetic technology that we optimized for stable gene transfer in EcN probiotic bacterium. In the study, the cryptic plasmids of EcN, known as pMUT1 and pMUT2, were knocked out by using CRISPR-Cas based plasmid curing. Subsequently, both of the cryptic plasmids were engineered as vectors to transfer gene-of-interest stably into EcN as these plasmids can propagate in EcN in absence of any selection. The retention of recombinant plasmids in EcN

was validated *in vitro*, suggesting that cryptic plasmids could be useful tools to transfer genes in EcN without need of cumbersome genome alterations.

In Chapter 4, I will describe a synthetic secretion system developed for EcN to release therapeutical proteins to extracellular medium without need of dedicated secretion machineries. In the synthetic secretion system that we named as inducible Leaky Outer Membrane based Secretion System (iLOM-SS), the outer membrane of EcN was made permeable by knocking-down structural protein(s) at the periplasm via small RNAs. iLOM-SS was employed secrete various proteins ranging from 25kDa to 100kDa. Furthermore, the safety of iLOM-SS when used in EcN was validated on Caco-2 cells *in vitro*. Altogether, the study highlights that iLOM-SS could be a valuable tool to develop EcN therapeutics that can delivery proteins safely.

In Chapter 5, I will explain details of the collaborative study with Mimeo' Lab on engineering commensal Lachnospiraceae species to augment their anti-inflammatory activity in the gut. In this project, I work with Jack Arnold to discover and characterize signal peptides to secrete proteins using Lachnospiraceae species. In addition, I optimized a display shuttle that can be used to express proteins bound to outer membranes of Lachnospiraceae species. These tools could be exploited to design therapeutics based on commensal Lachnospiraceae species which have correlation with a variety of gut pathologies.

In the final chapter, I will summarize the studies given in this thesis and their contribution to current literature. Furthermore, I will discuss future perspectives of the works described herein in context of therapeutical development.

CHAPTER 2

Griffithsin as a Broad-Spectrum Antiviral Agent

The work described in this section is partially published in the following study [125]:

Recep E. Ahan^{1,‡}, Alireza Hanifehnezhad^{2,‡}, Ebru Ş. Kehribar^{1,‡}, Tuba C. Oguzoglu², Katalin Földes², Cemile E. Özçelik¹, Nazlican Filazi², Sıdıka Öztop³, Fahreddin Palaz⁴, Sevgen Önder⁴, Eray U. Bozkurt¹, Koray Ergünay⁵, Aykut Özkul^{*,2,6}, and Urartu Özgür Şafak Şeker^{*,1}

1 UNAM-Institute of Materials Science and Nanotechnology, Bilkent University, Ankara, Turkey

2 Faculty of Veterinary Medicine, Department of Virology, Ankara University, Ankara, Turkey

3 Adana Dr. Turgut Noyan Medical and Research Center, Department of Immunology, Baskent University, Adana, Turkey

4 Faculty of Medicine, Department of Medical Pathology, Hacettepe University, Ankara, Turkey

5 Faculty of Medicine, Department of Medical Microbiology, Virology Unit, Hacettepe University, Ankara, Turkey

6 Biotechnology Institute, Ankara University, Ankara, Turkey

[#]These authors have contributed equally

*Corresponding authors: Aykut Özkul; E-mail: aykut.ozkul@ankara.edu.tr and
Urartu Özgür Şafak Şeker; E-mail: urartu@bilkent.edu.tr

Reprinted with Open Access Licensed under CC-BY 4.0 from following
publications; *ACS Infect. Dis.* 2022, 8, 7, 1253–1264

2.1. Objective of the Study

The global impact of the COVID-19 (coronavirus disease-19) pandemic caused by the novel coronavirus SARS-CoV-2 (severe acute respiratory syndrome coronavirus 2) was substantial, affecting people around the globe for over three years. The rising number of vaccinations is promising, but there remains a significant demand for prophylactic medications to ensure safe social interactions. To address this need, several existing drug compounds have been repurposed to combat SARS-CoV-2, with some demonstrating effectiveness in preventing hospitalizations or ICU admissions. In our study, we showed that griffithsin (GRFT), a lectin protein, can prevent the infection of Vero E6 cell lines and IFNAR^{−/−} mouse models by SARS-CoV-2 and its variants, Delta, and Omicron via binding to the spike protein of SARS-CoV-2. Considering the mutation rate of SARS-CoV-2 and ongoing infection cycles, we believe that the use of GRFT as prophylactic drug against SARS-CoV-2

can prevent transmission of current and emerging variants of the virus such as Delta and Omicron.

2.2. Introduction

SARS-CoV-2, which was first identified in December 2019, causes COVID-19, a disease that evolved into a worldwide pandemic and a significant global health concern until 2023. [126] The airborne transmission of SARS-CoV-2 from person to person led to the virus's rapid dissemination to nearly every country in under a year. [127] Because of the rapid transmission of the virus, many countries have enforced compulsory face mask mandates and social distancing measures. Simultaneously, widespread viral testing was performed to prevent further transmission by identifying and isolating individuals who may be asymptomatic or presymptomatic carriers. [128]

The genetic information of SARS-CoV-2 is encoded in a positive-single-stranded RNA molecule which is composed of six open reading frames for ORF1ab, spike glycoprotein (S), nucleocapsid (N), membrane (M), and envelope (E) proteins. [129] The S-protein of SARS-CoV-2, which exhibits a moderate resemblance to SARS-CoV, encompasses a receptor-binding domain (RBD) responsible for binding with the ACE2 protein which is the main cell surface receptor exploited by the virus for invading the cells. [130], [131] The S protein exists in a trimeric form, displaying two distinct metastable configurations, with the RBD either in the "up" or "down" state. When it interacts with ACE2, surface proteases like TMPRSS2 or cathepsin L trigger the cleavage of the S1 and S2 domains of the S protein. This cleavage event induces irreversible structural alterations and prepares the virus for the fusion process. [132]–[134] Given the strong antigenic properties of S proteins and the

critical role of ACE2 binding in viral entry, numerous preventive vaccines, including those based on mRNA (BioNTech/Pfizer [135] and Moderna [136]), adenoviral vectors (J&J [137] and AstraZeneca/Oxford [138]), and recombinant antigens (Novavax [139]) utilize S proteins to stimulate an immune response against SARS-CoV-2.

Beyond vaccines, the therapeutic options for COVID-19 include the disruption of viral reproduction through the inhibition of viral entry into host cells or the viral replication process. Notably, studies focusing on drug repurposing have effectively pinpointed a range of small molecules, including remdesivir, as antiviral against SARS-CoV-2. Remdesivir, for instance, targets the RNA dependent RNA polymerase of SARS-CoV-2 and has received approval from the FDA for clinical application. [140] Furthermore, emergency-use authorization has also been granted by regulatory agencies for neutralizing antibody (nAb) therapies that cover transfer of convalescent plasma from recovered people and administration of the recombinant monoclonal anti-spike antibodies. [141], [142] In general, neutralizing antibodies (nAbs) impede the interaction between the receptor-binding domain (RBD) and the ACE2 protein, thereby hindering the internalization of the virus. However, certain nAbs bind to regions within the S protein other than the RBD, with the aim of obstructing conformational changes necessary for the fusion state. [143] Despite the evident positive results achieved with neutralizing antibody (nAb) therapies for COVID-19 patients, accessibility to this treatment is constrained by the limited supply of convalescent plasma and the expensive production procedures involved in monoclonal antibody development. [144] Furthermore, the ongoing evolution of SARS-CoV-2 presents a substantial concern, as the accumulation of mutations in the spike (S) protein can result in immune evasion, potentially leading to reinfections

and diminishing the effectiveness of both neutralizing antibody (nAb) therapies and vaccination. Throughout the course of the COVID-19 pandemic, a multitude of SARS-CoV-2 variants have emerged and became the dominant circulating strain before being supplanted by others. [145]

The Alpha/B.1.1.7 variant had been the dominant strain circulating around the globe in 2021, owing to its heightened transmissibility, until emergence of Delta/B1.617.2 variant in 2022 which had higher immune evasion compared to other strains of SARS-COV-2. [146], [147] Subsequently, Delta/B1.617.2 was replaced by emergence of Omicron/B1.1.529 which have at least 15 amino acids mutation at its RBD thereby enabling the virus to escape antibody responses acquired from both natural infection and vaccination. [147] The complete evolutionary potential of the S protein is yet to be determined, but it is highly probable that SARS-CoV-2 will accumulate further mutations in response to the pressure exerted by vaccine-induced immunity. This evolution of SARS-CoV-2 is further accelerated because of the uneven distribution of vaccines. [148] Consequently, the development of novel therapeutic agents with distinct inhibitory mechanisms is imperative for coping with viral evolution until the complete eradication is achieved.

Lectin proteins derived from seaweeds have demonstrated remarkable efficacy as antiviral agents against enveloped viruses, including HIV-1 [149], herpes virus[150], and highly pathogenic human coronaviruses, SARS-CoV [151] and MERS-CoV [152]. The antiviral properties of seaweed lectins stem from their strong affinity for surface glycoproteins present on viruses, such as the gp-120 protein of HIV-1 and the spike proteins of SARS-CoV and MERS-CoV. Through their binding to these viral surface proteins, lectins typically obstruct the initial viral internalization phase, thus

impeding viral infection. [153] Therefore, lectin-based proteins can offer distinct mechanism of action than antibodies from humoral immune responses to hinder SARS-CoV-2 replication. In this chapter, I will present the study in which we have showed that griffithsin protein from griffithsia algae have potent antiviral activity against both ancestral and evolved strains of SARS-CoV-2.

2.3. Materials and Methods

2.3.1. Cloning, Expression, and Purification of Griffithsin Protein

The griffithsin (GRFT) nucleotide sequence (accession number AY744144) with N-terminal 6× His-tags and flexible GS linkers was codon-optimized for *E. coli* and synthesized (Genewiz, NJ, USA). PCR reaction was performed with pREA144 and pREA145 to add suitable overhang sequences homologous to the pET22b cloning vector, without the *pelB* leader sequence. The GRFT sequences are listed in Table B1. The *grft* gene was cloned into the *pet22b* vector without a *pelB* leader sequence, using Gibson assembly. The reaction was transformed into a chemically competent strain of *E. coli* DH5α PRO by heat shock transformation. Selected colonies were verified by Sanger sequencing (Genewiz, USA).

For expression, the construct was transformed to the *E. coli* BL21 (DE3) strain. The overnight culture of cells was diluted to 1:100 in LB medium with appropriate antibiotics and was grown until the OD₆₀₀ was 0.4–0.6. Then, the culture was induced with 1 mM IPTG for 24 h at 16 °C, 200 rpm. The cells were harvested by centrifugation and resuspended in the lysis buffer (20 mM NaH₂PO₄, 500 mM NaCl, 20 mM imidazole pH 7.4). The suspension was sonicated at 30% power for 10 cycles of 15 s on/45 s off. The samples were centrifuged at 21,500 g for 1 h and the

supernatant was filtered with a 0.45 µm filter. The filtered lysate was loaded on a pre-equilibrated HisTrap nickel column (GE Life Sciences 17524701) using an ÄKTA start protein purification system. The column was washed with 10 column volumes of lysis buffer (20 mM imidazole) and 5 column volumes of lysis buffer with an imidazole gradient from 20 to 40 mM. Finally, proteins were eluted with 5 volumes of elution buffer (20 mM NaH₂PO₄, 500 mM NaCl, 500 mM imidazole pH 7.4). Purified griffithsin was desalted into 1× PBS (pH 7.4) using a HiTrap desalting column (GE Life Sciences) in an ÄKTA start protein purification system. Protein concentration is calculated via BCA colorimetric assay with BSA standards (Thermo Fisher Scientific).

2.3.2. SDS-PAGE and Western Blotting

Samples were boiled at 95 °C for 5 min with 1× SDS loading dye and electrophoresed on 15% SDS-polyacrylamide gel. The gel was placed into Coomassie blue staining solution for ~1 h with shaking and incubated in destaining solution (60% ddH₂O, 30% methanol, and 10% acetic acid) until the bands were clearly visible. For the western blot, the gel was transferred to the PVDF membrane using Trans-Blot Turbo (Bio-Rad). The membrane was blocked with 5% milk in TBS-T for 1 h at room temperature. Then, the membrane was incubated in 5% milk in TBS-T containing 1:10,000 primary anti-His mouse antibodies at 4 °C, overnight. The membrane was washed in TBS-T, then incubated in 5% milk in TBS-T containing 1:10,000 horseradish peroxidase (HRP)-conjugated goat anti-mouse secondary antibodies (Abcam ab6789-1 MG) for 1 h at room temperature. After washing in TBS-T, the membrane was incubated in ECL substrates (Bio-Rad 170-5060) and visualized using Vilber Fusion Solo S.

2.3.3. Circular Dichroism

A concentration of 7.5 μM rGRFT was prepared in 2 mM PBS buffer, pH = 7.4. Mannose concentrate solution was prepared with ddH₂O and diluted in GRFT solution at 7.5, 15, 30, 60, and 120 μM final concentrations. The mannose protein mixture was incubated for 30 min at room temperature. The CD spectra of GRFT/mannose mixtures were measured from 240 to 200 nm with three repeats (Jasco J-815) at room temperature with a 300 s delay time and a 1 mm bandwidth. The secondary structure composition was calculated using the BeStSel tool developed by Micsonai et al. [154]

2.3.4. ITC Assay

ITC assays were performed using MicroCal ITC200 (Malvern Panalytical). Spike protein was purchased from Synbiotik Biotechnology LLC (Ankara, Turkey); 1 μM spike protein solution and 50 μM rGRFT protein solution were prepared in 1 $\times$ PBS, pH = 7.4. Two microliters of rGRFT solution was injected with 120 s intervals into the calorimetric cell containing 280 μL of spike protein solution. Buffer and dilution effects were corrected using Origin MicroCal Analysis software. Graphs were generated via GraphPad Prism software.

2.3.5. QCM Assay

The QCM-D gold sensor (Biolin Scientific QSense QSX 301 Gold) surface was cleaned with piranha solution (H₂O₂/H₂SO₄ in 1:3 ratio) for 30 min at 80 °C to remove any contaminants. After piranha cleaning, the gold sensor chip was immersed in ddH₂O for 5 min twice.

The cleaned gold sensor chip was immersed in 20 mM 11-mercaptoundecanoic acid (11-MUA) solution and incubated overnight. The gold sensor was rinsed with first 100% ethanol then ddH₂O. The gold sensor surface coated with 11-MUA was functionalized by a QSense chamber with 400 mM EDC, which was followed by 100 mM NHS with a flow rate of 20 μ L/min. The chip was rinsed with 1 $\times$ PBS buffer to remove any residual EDC and NHS. Then, 150 μ L of spike protein with a concentration of 400 μ g/mL was introduced into the chamber with a 6.49 μ L/min flow rate to make the spike protein attached to the surface. After coating the surface with spike protein, the surface was deactivated by 1 M ethanolamine HCl with a flow rate of 20 μ L/min to avoid further attachment to the functionalized surface rather than spike protein. Deactivation was followed by the 1 $\times$ PBS washing step to equilibrate the chamber at 20 μ L/min. Afterward, 200 μ L of 0.5, 1, 2, 3, 4, 5, 7.5, and 10 μ M rGRFT was introduced sequentially at a 20 μ L/min flow rate. Between each concentration, the chip was washed with 1 $\times$ PBS at a 20 μ L/min flow rate to remove any unbound molecule. First, third, fifth, seventh, and ninth overtones for frequency and dissipation values were recorded simultaneously.

The binding affinity of rGRFT toward SARS-CoV-2 virus was tested with QCM-D. The surface of the gold sensors was prepared as described above. The viral particles were attached to the surface of QCM-D via amine coupling. After the attachment of the viral particles, the activated surface as mentioned above was blocked with ethanolamine (1 M) in PBS. Final surface topographies were analyzed with atomic force microscopy as demonstrated elsewhere.[155] Afterward, rGRFT proteins at varying concentrations of 1.7, 2.7, 5.4, 6.8, and 13.6 μ M were flown on top of the inactivated and immobilized SARS-CoV-2 virus on sensor chips.

2.3.6. rGRFT Binding on SARS-CoV-2 by ELISA

Ninety-six-well plates (353916, Corning) were coated with 50 μ L of 106 particle per mL SARS-CoV-2 isolated from Vero E6 cells by adding 150 μ L of 100 mM bicarbonate/carbonate coating buffer (pH 9.6) and incubated O/N at 4 °C. The wells were washed three times with 200 μ L of PBS containing 0.1% Tween-20 (Merck) (PBS-T) Two hundred microliters of blocking buffer (1% BSA in PBS-T) were added to wells and incubated for 1 h. After incubation, the blocking buffer was discarded and 100 μ L of 1 μ g/mL rGRFT was added to each well and incubated for 2 h. The wells are washed with 200 μ L of PBS-T three times. Then, each well was incubated with a PBS-T solution containing 1:3000 primary anti-His mouse antibodies (PTGLAB 66005) for 1 h. Washing was performed with 200 μ L of PBS-T. Secondary antibody solution containing 1:3000 HRP-conjugated goat anti-mouse secondary antibodies (Abcam ab6789-1 MG) in PBS-T was added to each well and incubated for 1 h. Washing steps were performed with PBS-T and wells were incubated with 100 μ L of TMB substrate solution for 10 min in the dark. TMB stop solution was added and absorbance at 450 and 650 nm was measured with a microplate reader (SpectraMax M5, Molecular Devices).

2.3.7. Virus, Cells, Animals, and Ethical Approval

SARS-CoV-2 local isolates Ank2 (Wuhan like; GenBank Acc. No: MT478019), Ank-DLT-1 (Delta variant; GenBank Acc. No: OM295705) and Ank-OmicGKS (Omicron variant) were used in the experiments requiring live virus. [155] African green monkey kidney (Vero E6, ATCC: CRL-1586) cells, obtained from the cell culture collection of the Department of Virology, Ankara University Faculty of Veterinary Medicine, were also utilized in the study. The experiments with C57BL/6

and IFNAR^{-/-} mice were performed with permission of the Ankara University Ethical Committee for Animal Experiments (06 May 2020, 20120-8-66) in a high containment animal facility (ABSL3+), conducted according to the national regulations on the operation and procedure of animal experiments' ethics committees (regulation no. 26220, 9 September 2006). The 14-day nonlethal SARS-CoV-2 infection model in IFNAR^{-/-} mice was previously reported. [155]

Tissue culture infective dose 50% (TCID₅₀) was used to assess in vitro virus infectivity, performed as described previously. [155], [156] Quantitative genome detection was carried out by real-time reverse transcription polymerase chain reaction.

2.3.8. Toxicity and Immunogenicity

Vero E6 cells were grown in Dulbecco's Modified Eagle's Medium (DMEM) supplemented with 10% heat-inactivated fetal bovine serum, 2 mM l-glutamine, 100 units/mL penicillin, and 100 µg/mL streptomycin at 37 °C in a 5% CO₂ humidified incubator. To determine toxicity of rGRFT on Vero E6 cells, MTT assay was used; 3 × 10⁵ cells were seeded onto 96-well plates and incubated at 37 °C in a 5% CO₂ humidified incubator for 24 h. Then, different concentrations of rGRFT in PBS were added on cells and incubated for another 48 h. Cell viability was assessed using the MTT Cell Proliferation Assay Kit (Trevigen, 4890-25-K), according to the manufacturer's instructions.

For innocuity testing, four C57BL/6 mice were intraperitoneally inoculated with 100 nM (250 µL) rGRFT three times with 24 h intervals and observed for 14 days with daily measurements of body temperature and weight. They were sacrificed at the end of the period with specimens for biochemical, hematological, and histopathological

assessment. Mouse sera at day 14 was screened for rGRFT-specific antibodies. For this purpose, ELISA plates were coated overnight with rGRFT in bicarbonate buffer and the sera in 1/100 dilution were inoculated for an hour at room temperature. The assay was evaluated using anti-mouse-IgG-HRP conjugates and TMB substrates. The 6×-His peptide present in the recombinant protein was targeted as the control, detected using mouse anti-His antibodies, anti-mouse-IgG-HRP conjugates, and TMB substrates. The same serum samples were used to determine concentration of IFN γ , TNF α , and IL-2 using commercially available ELISA kits (BioLegend).

2.3.9. In Vitro Testing of rGRFT

In vitro virus inhibition was monitored by two approaches. Initially, neutralizing activity of rGRFT at different incubation periods was determined. For this purpose, rGRFT (6 μ M) was fivefold diluted in high glucose DMEM, combined with an equal volume (100 μ L) of 100 TCID₅₀ of Ank2 (Wuhan like), Ank-DTL-1 (Delta variant), and Ank-OmicGKS (Omicron variant) strains of SARS-CoV-2, and the mixture was incubated at room temperature for 15 min. A 150 μ L of the mixture was then inoculated on Vero E6 cells grown in a 96-well flat-bottomed tissue culture plate, with highest concentration (6 μ M) of rGRFT without the virus was used as control, serum-free DMEM as cell, and 75 μ L of 100 TCID₅₀ virus as virus controls in each plate. The plates were incubated at 37 °C in a 5% CO₂ atmosphere and evaluated when the virus controls showed 100% CPE. The 50% inhibitory dose (IC₅₀) was calculated as described elsewhere. [157]

In the second approach, Vero E6 cells grown in 24-well tissue culture plates were infected with the SARS-CoV-2 Ank2 isolate at 1 MOI and rGRFT (in IC₅₀ concentration) was added at 1 h post infection. The test was incubated until virus

controls showed 90% CPE. TCID₅₀ values and virus genome copies were determined in culture supernatants by commercially available q-PCR and cell layers at termination. The test was performed in quadruplicate for all time points.

2.3.10. In Vivo Testing of rGRFT

rGRFT was delivered intranasally in 100 nmol per nostril in 50 µL volume of carrier solution including 2.5% (w/v) polyethylene glycol 1450 (PEG 1450), 0.5% benzyl alcohol, 1.5% (w/v) glycerine, 0.02% (w/v) h EDTA-diNa, and 0.02% (w/v) benzalkonium chloride. The solution was sterilized by membrane filtration (0.2 nm, Millipore, USA) and stored under chilled conditions. Intranasal virus inoculations were carried out as 10³TCID₅₀ in 50 µL per nostril.

A total of 28 IFNAR^{-/-} mice (8–12 weeks) were randomly divided into six groups for prophylaxis and contact experiments as well as control individuals in the prophylaxis group received intranasal rGRFT and were inoculated with SARS-CoV-2 Ank2 after 4 h. Each animal further received the same amount and route of rGRFT following 4 subsequent days. The protective effect of rGRFT was tested with a contact scenario in pairs housed in the same cage. One individual in each cage was intranasally inoculated with the virus and the accompanying mouse received rGRFT (protected contact) or null carrier solution (unprotected contact) (Figure 2). The rGRFT and carrier treatments were repeated at the same time for 6 additional days. Animals were monitored clinically for 14 days with tail vein blood specimens collected on days 0 and 7. At the end of the period, the animals were humanely euthanized and tissue specimens were collected for histopathological evaluation and immunohistochemistry, performed as previously described. [155]

2.4. Results

2.4.1. Cloning of rGRFT into pET22b

Gene encoding the optimized version of rGRFT for *E. coli* expression was synthesized from Genewizz and subsequently amplified with pREA1 and pREA2 to add suitable overhangs for cloning into pET22b vector with isothermal assembly reaction. (Figure 2-1a) PCR products of rGRFT amplification that appeared at 473bp was isolated and cloned into KpnI-XhoI linearized pET22b. Following the transformation of circularized plasmid into DH5 α cells, colony PCR was performed to verify the cloning. Plasmid isolation was performed from the colonies that gave the correct band in colony PCR. The plasmids were verified with NGS. (Figure 2-1b)

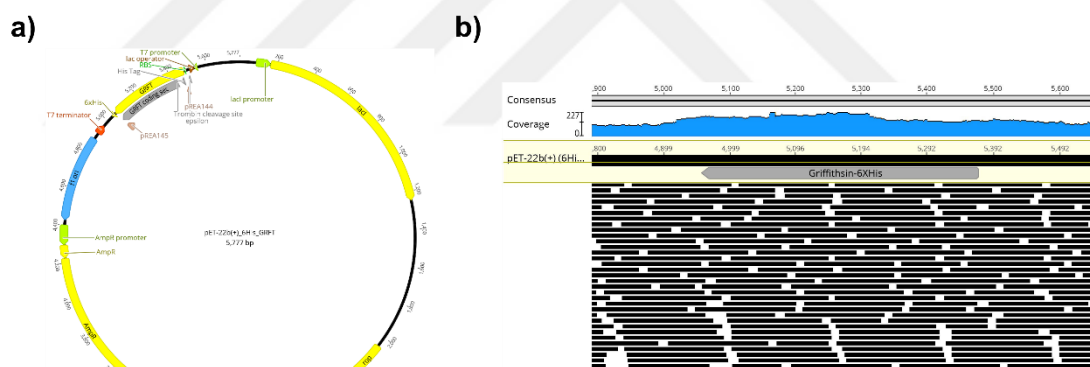


Figure 2-1. Expression plasmid for rGRFT expression in *E. coli*. a) Constructed plasmid map for recombinant GRFT production. b) The alignment of NGS reads to the constructed map.

2.4.2. Expression and Purification of rGRFT

The verified plasmid was transformed into BL21 (DE3) via chemical transformation for protein expression. After overnight growth of transformants on the agar plate, a

single colony was picked and inoculated into liquid media. In the next day, grown cells were diluted 100-fold into fresh media and induced with 1mM IPTG when the OD₆₀₀ of the media reached 0.4. Subsequently, cells were collected, lysed, and purified with Ni-NTA column using Akta Start FPLC system. A sharp peak was observed in the elution step, suggesting that the rGRFT was successfully expressed in soluble form and purified. (Figure 2-2a) To validate purification, peaks were collected and run in the SDS-PAGE gel. rGRFT protein band was seen at the 13kDa which was further confirmed with western blot by labelling the proteins with anti-histidine antibody. (Figure 2-2b) Image analysis in ImageJ revealed that the purified rGRFT had at least 95% purity.

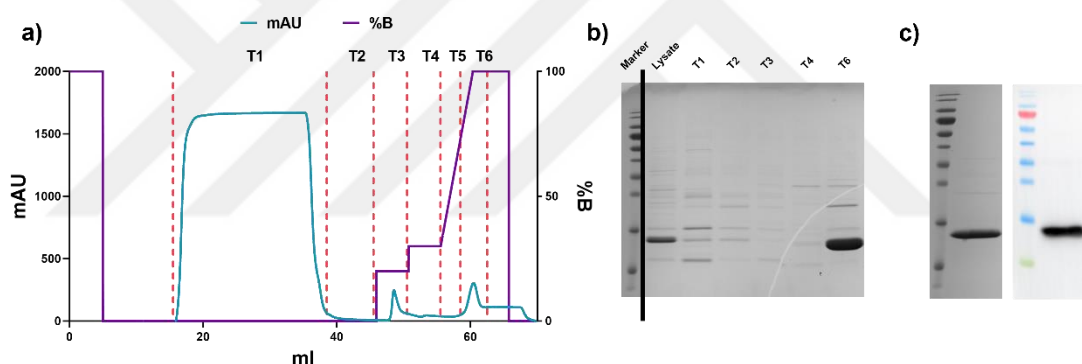


Figure 2-2. Purification of rGRFT. a) FPLC chromatogram. b) SDS-PAGE analysis of peak obtained from FPLC. c) Western blot and SDS-PAGE analysis of purified rGRFT with anti-histidine antibody.

2.4.3. Structural Changes of rGRFT upon Binding with Mannose Sugar

The secondary structures of rGRFT were assessed using circular dichroism (CD) spectroscopy in both its unbound (apo) state and when bound to mannose sugar. The CD spectrum of apo-rGRFT exhibited a minimum at approximately 218 nm, aligning with its predominantly beta-sheet-rich 3D structure. Notably, there were no

discernible alterations in the overall secondary structure elements upon the addition of mannose, even when present in up to a 16-fold molar excess over rGRFT. (Figure 2-3)

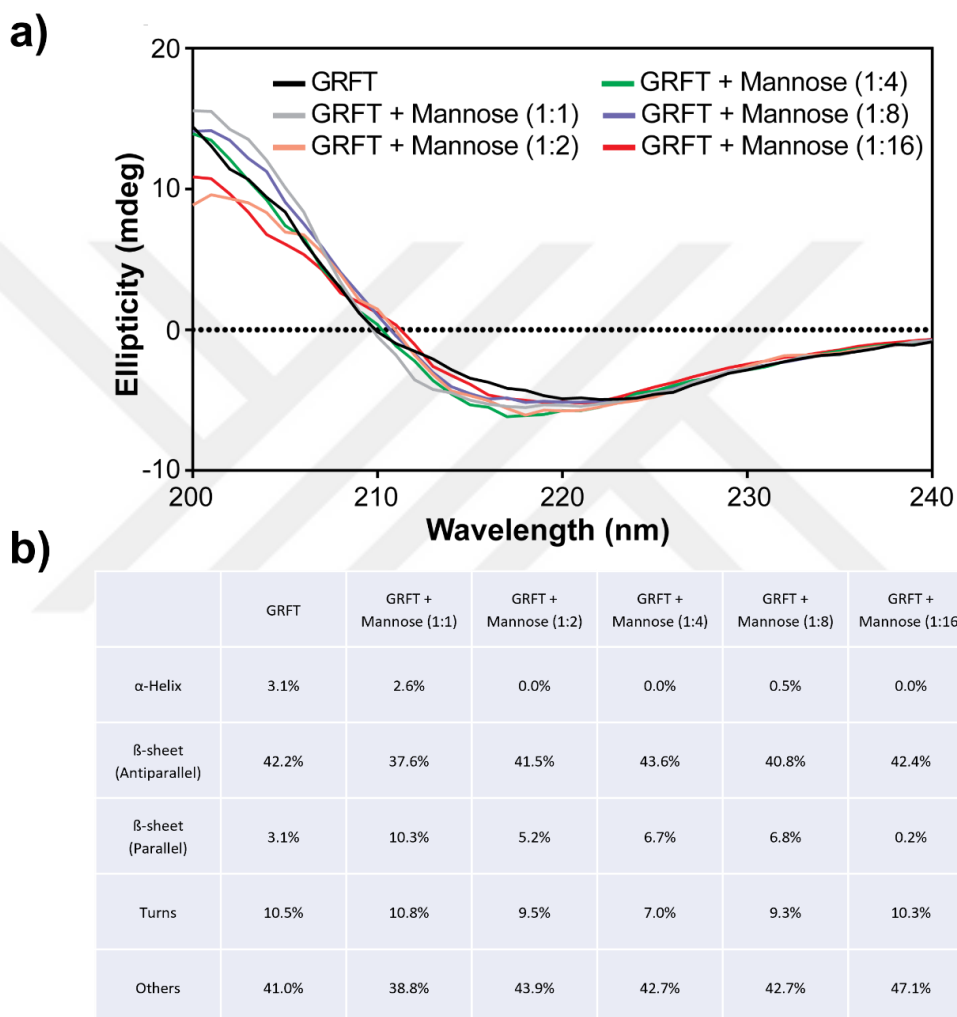


Figure 2-3. CD analysis of apo-rGRFT and holo-rGRFT with mannose. CD spectrum (up), and numeric analysis of CD spectrums to calculate % of structures with BeStSel (bottom)

2.4.3. Binding of rGRFT to SARS-CoV-2 Spike Protein

The interaction between purified rGRFT and the recombinant SARS-CoV-2 spike protein was probed utilizing both ITC and QCM techniques. The dissociation constant (K_d) for the rGRFT-spike protein complex was determined to be 9.9 μM via ITC and 1.6 μM through QCM-D measurements. Notably, ITC results revealed a binding stoichiometry of 6:1 between rGRFT and the spike protein. (Figure 2-4a and 2-4b)

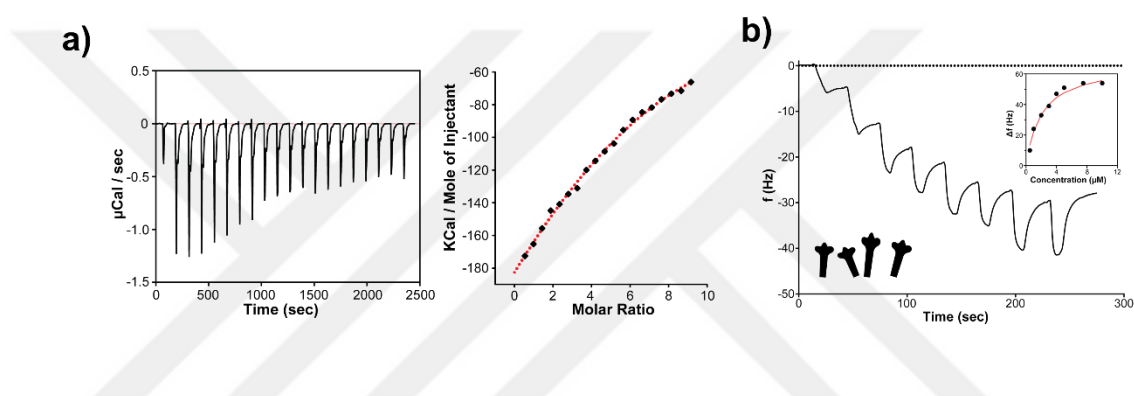


Figure 2-4. Validation of rGRFT binding to SARS-CoV-2 S protein. a) ITC titration of rGRFT onto S protein. b) QCM-D measurement for rGRFT binding to immobilized S protein onto surface of carboxylated gold chip.

2.4.4. Binding of rGRFT to Whole SARS-CoV-2

In addition to rGRFT binding to S protein, the interaction between rGRFT and the whole SARS-CoV-2 virus was examined using an in-house developed ELISA. In this assay, heat-inactivated SARS-CoV-2 particles were immobilized on the ELISA wells to capture added free rGRFT. In the presence of SARS-CoV-2 particles, rGRFT remained firmly attached to the wells even after washing, while significantly less rGRFT was observed on the wells in the absence of SARS-CoV-2 particles. (Figure 2-5a) To determine the binding affinity of rGRFT to the viral particles, QCM-D

analysis was employed, revealing that rGRFT could bind to heat-inactivated SARS-CoV-2 viral particles with a K_d value of 4.1 μM . (Figure 2-5b)

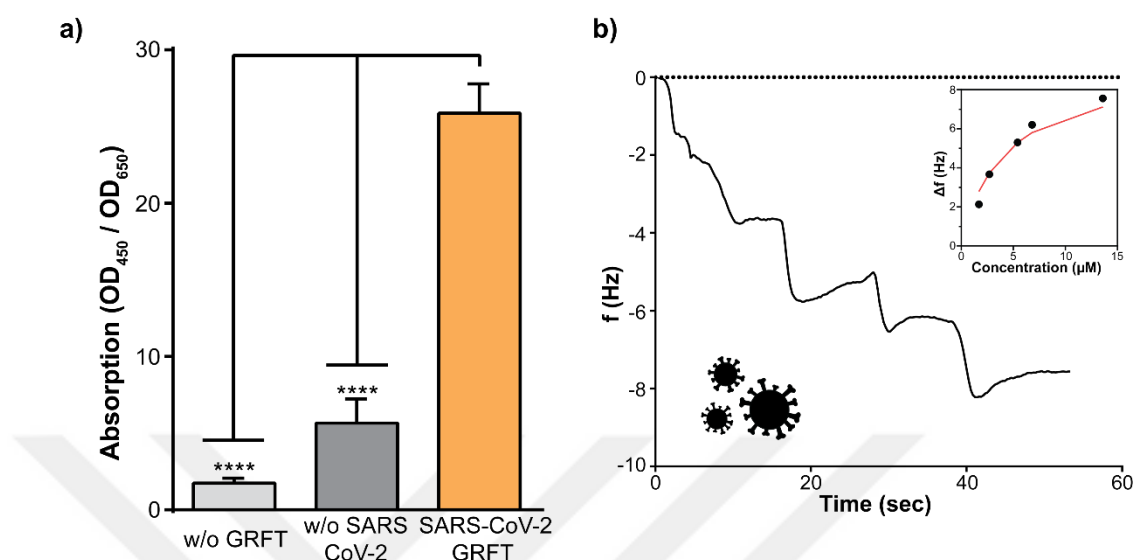


Figure 2-5. Validation of rGRFT binding to whole SARS-CoV-2 virus (Ank2 strain). a) In-house whole SARS-CoV-2 ELISA result to validate qualitatively the binding of free rGRFT to surface absorbed SARS-CoV-2 viruses. b) QCM-D measurement for rGRFT binding to immobilized inactivated SARS-CoV-2 viruses onto surface of carboxylated gold chip.

2.4.5. Antiviral Activity of rGRFT against SARS-CoV-2 *in vitro*

The antiviral activity of rGRFT against SARS-CoV-2 infection was evaluated *in vitro* using the Vero E6 model cell line. (Figure 2-6a) Serially diluted rGRFT solutions, spanning concentrations from 6 μM to 15.4 pM, were mixed with infectious SARS-CoV-2 particles. Subsequently, cell infections were carried out with these mixtures and incubated until the cytopathic effect (CPE) reading reached 100% in the sample that did not contain rGRFT. The infection results indicated that the IC_{50} of rGRFT against SARS-CoV-2 infection was determined to be 33.2 nM.

(Figure 2-6b) Furthermore, the impact of rGRFT at the IC₅₀ concentration on SARS-CoV-2 replication was examined. Vero E6 cells were infected with SARS-CoV-2 in the presence or absence of rGRFT, and the quantity of total viral particles in the cell supernatant was determined through qPCR. It was observed that rGRFT could effectively inhibit SARS-CoV-2 infection within a maximum of 24 hours following inoculation. (Figure 2-6c)

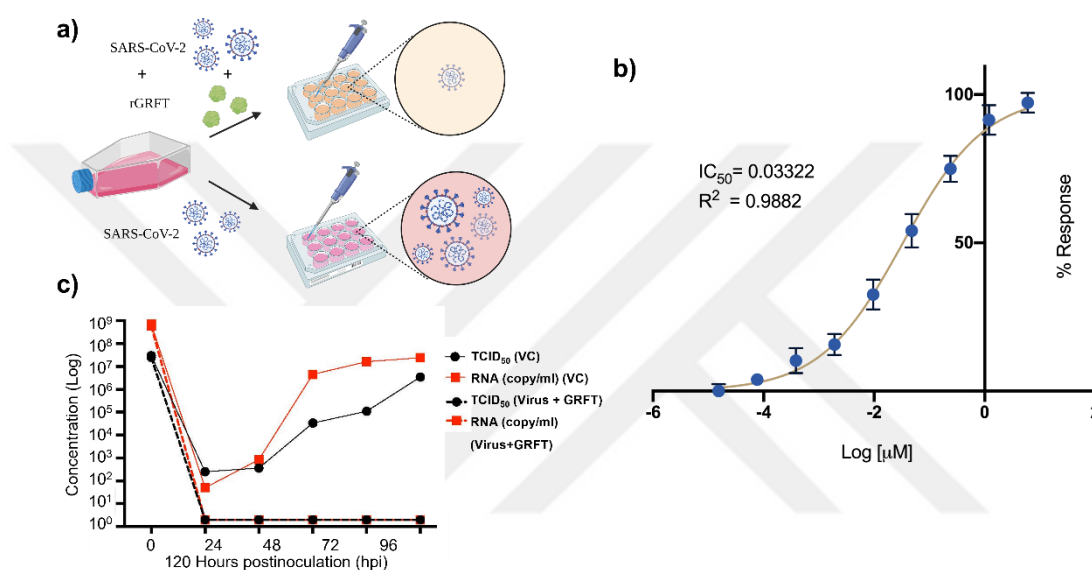


Figure 2-6. *In vitro* assessment of antiviral activity of rGRFT against SARS-CoV-2 infection. (a) Illustration depicting the inhibition of SARS-CoV-2 infection in Vero E6 cells by rGRFT, generated using BioRender.com. (b) Calculation of the IC₅₀ value of rGRFT for SARS-CoV-2 in Vero E6 cells. (c) Evaluation of the effect of rGRFT on SARS-CoV-2 replication by Q-PCR in Vero E6 cells.

2.4.6. Toxicity and Immunogenicity of rGRFT

Before starting the efficacy experiments, toxicity, carcinogenicity, and immunogenicity was investigated *in vitro* with veroE6 cells and *in vivo* with C57BL/6 mice. To test cytotoxic and carcinogenic effects of rGRFT on Vero E6

cells, different amounts of rGRFT ranging from 8.56 nM to 137 nM was added onto cells and cells were incubated for 24 hours. Next day, viable cells were quantified via MTT assay. According to MTT assay results, rGRFT doesn't have any cytotoxic and carcinogenic effects Vero E6 cells. (Figure 2-7)

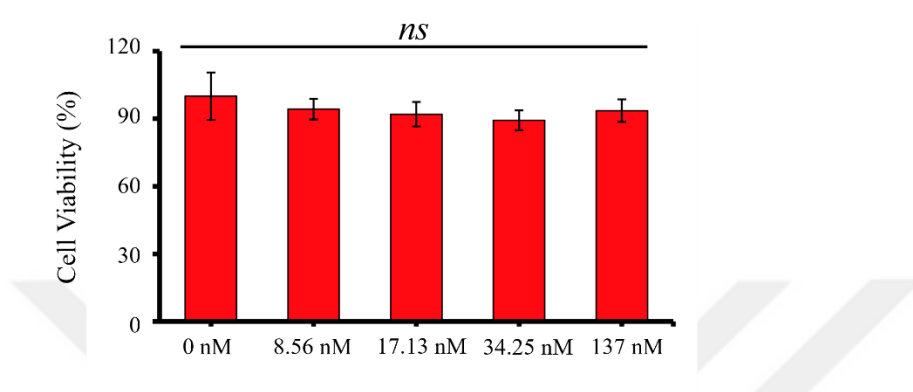


Figure 2-7. Cytotoxic and carcinogenic effects of rGRFT on Vero E6 cells.

To test the toxicity and immunogenicity of rGRFT *in vivo*, rGRFT was administered into C57BL/6 mice intraperitoneally and treated mice were checked for 14 days. Injection of rGRFT did not result in any substantial alterations in the body temperature or weight of the mice (Figure 2-8a and Figure 2-8b). Throughout the 14-day observation period, the mice exhibited no indications of local or systemic toxicity. Tissue samples extracted from the liver, lung, kidney, and spleen of sacrificed animals were meticulously examined under a microscope to assess any potential histopathological effects. Notably, no discernible histopathological changes were observed in the respective tissue samples (Figure 2-8d). The immunogenicity of rGRFT was evaluated using an in-house developed ELISA designed to capture anti-rGRFT antibodies from serum via surface-immobilized rGRFT. The ELISA results conclusively demonstrated that rGRFT did not elicit a humoral immune response (Figure 2-8c). Additionally, serum concentrations of IL-2, TNF α , and INF γ were

quantified using commercially available ELISA assays. These analyses revealed no significant alterations in serum cytokine levels between the control group and the mice that received rGRFT (Figures 2-9a, 2-9b and 2-9c).

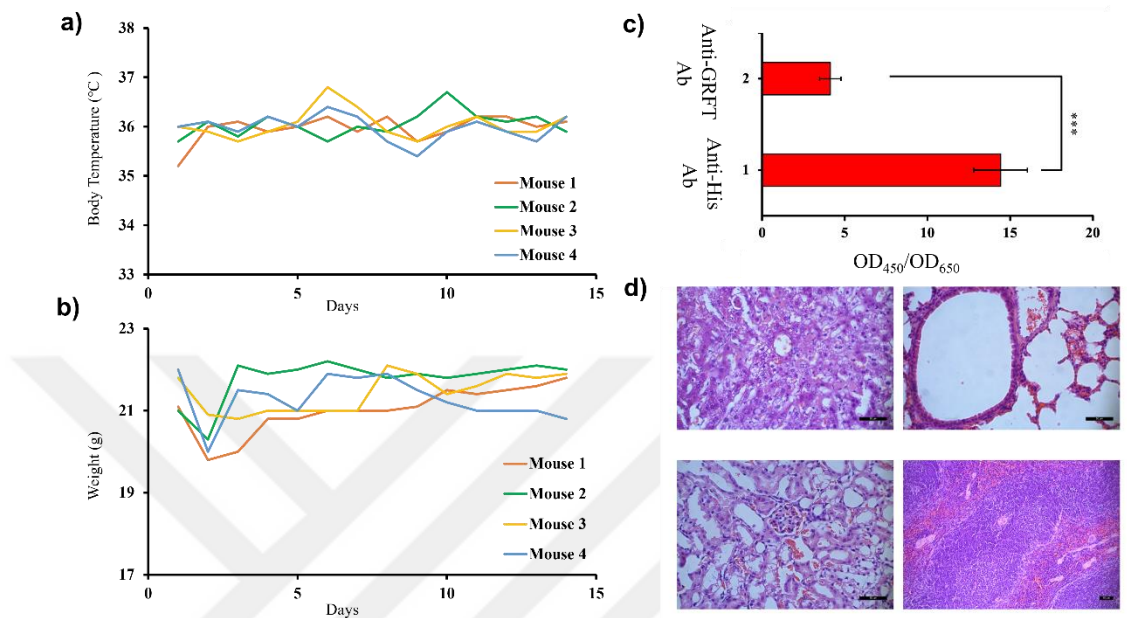
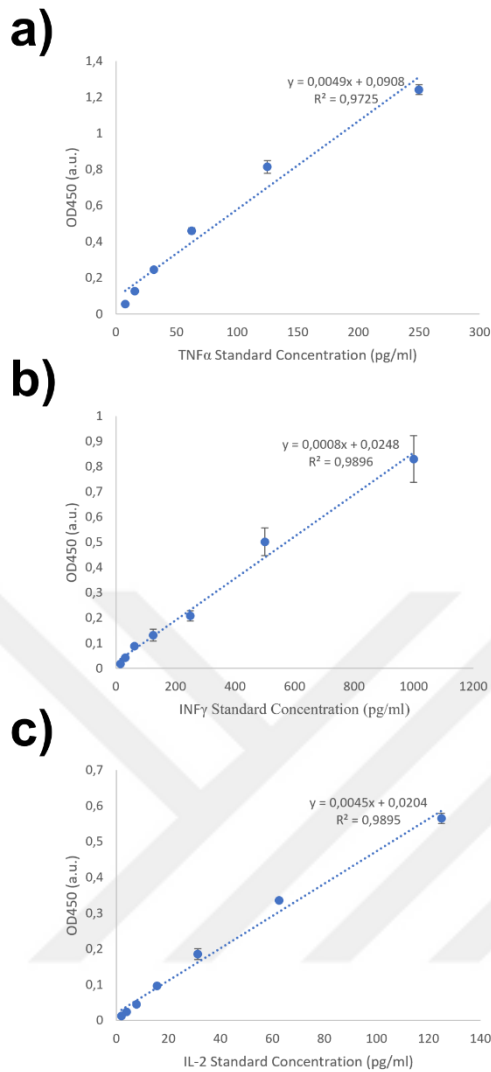


Figure 2-8. Assessment of rGRFT safety in vivo with C57BL/6 mice model. a, b) Body temperature (a) and weight (b) of mice that were treated with rGRFT. c) Humoral antibody response to intraperitoneal injected rGRFT. As control, commercially available anti-histidine antibody was used. ($p < 0.05$) d) Tissue sections from the liver (upper left), lung (upper right), kidney (lower left), and spleen (lower right) of C57BL/6 mice were examined on the 14th day after receiving intraperitoneal rGRFT administration. The sections were stained with hematoxylin and eosin, and the scale bar represents 50 μm.



Sample Name	Absorption	Calculated concentration*
rGRFT mouse 1	-0,0021	ND
rGRFT mouse 2	0,0005	ND
rGRFT mouse 3	0,0503	ND
rGRFT mouse 4	0,0004	ND
Control 1	-0,001	ND
Control 2	0,0024	ND
Control 3	0,0192	ND

Sample Name	Absorption	Calculated concentration*
rGRFT mouse 1	0,0048	ND
rGRFT mouse 2	0,0077	ND
rGRFT mouse 3	0,0365	ND
rGRFT mouse 4	0,0087	ND
Control 1	-0,0020	ND
Control 2	0,0024	ND
Control 3	0,02855	ND

Sample Name	Absorption	Calculated concentration*
rGRFT mouse 1	-0,0002	ND
rGRFT mouse 2	0	ND
rGRFT mouse 3	0,0002	ND
rGRFT mouse 4	0,0015	ND
Control 1	0,0043	ND
Control 2	0,006	ND
Control 3	-0,0008	ND

Figure 2-9. Analysis of serum cytokine levels of mice following rGRFT treatment. a) ELISA results for TNF α . b) ELISA results for INF γ . c) ELISA results for IL-2.

In addition, the biochemical parameters related to liver and kidney function, as well as blood counts, all remained within the healthy range (Table 2-1 and Table 2-2).

Table 2-1. Blood biochemistry results for main indicators of liver and kidney functions in C57BL/6 mice on the 14th day following intraperitoneal rGRFT administration. The experiment was done in four replicates. (ALB: albumin, ALT:

alanine aminotransferase, ALP: alkaline phosphatase, AST: aspartate aminotransferase, BUN: blood urea nitrogen, CRE: creatinine, GLU: glucose).

Parameters							
	ALB (g/dL)	ALT (U/L)	ALP (U/L)	AST (U/L)	BUN (mg/dL)	CRE (mg/dL)	GLU (mg/dL)
Average	2,95	66,5	126,75	133,25	16,75	0,275	202,75
	± 0,29	± 6,56	±14,39	±17,69	±3,40	±0,05	±27,46
Range	2,6 - 3,3	58 - 73	109-144	116-158	12 - 20	0,2 – 0,3	174-240

Table 2-2. Hematologic parameters in C57BL/6 mice on the 14th day following intraperitoneal rGRFT administration. The experiment was done in four replicates. (WBC; white blood cells, RBC: erythrocytes, HGB: hemoglobin, PLT: platelets, NEUT: neutrophils, LYMP: lymphocytes, MON: monocytes, EOS: eosinophiles, BASO: basophiles, HCT: hematocrit). For EOS, one mouse was excluded as outlier (0,71)

Parameters										
	WBC (K/ μ L)	RBC (M/ μ L)	HGB (g/dL)	PLT (K/ μ L)	NEUT (K/ μ L)	LYMP (K/ μ L)	MON (K/ μ L)	EOS (K/ μ L)	BASO (K/ μ L)	HCT (%)
Averg.	7,52	9,29	12,8	857	1,68	8,39	0,75	0,07	0,04	53,5
	±2,68	±1,88	±3,07	±257,08	±0,94	±3,82	±0,33	±0,05	±0,06	±9,71
Range	4,57- 10,9	7,57- 11,2	9,78- 16,3	568- 1179	0,94- 3,04	3,51- 12,3	0,38- 1,01	0,01- 0,12	0- 0,12	41- 63

2.4.7. Anti-viral Activity of rGRFT on SARS-CoV-2 *in vivo*

The *in vivo* efficacy of rGRFT was evaluated in two different experimental setups. In the first setup, rGRFT was administered intranasally to mice prior to direct virus inoculation through the nose. In the second setup, mice were divided into three groups: the first group received rGRFT, the second group remained untreated, and the third group was intranasally infected with virus particles (Figure 2-10a). Subsequently, each member of the third group was housed in a cage with either a member of the first group or a member of the second group. Treatments were administered to each individual mouse over a 6-day period. The mice were closely monitored for 14 days to evaluate the protective efficacy of rGRFT. In both scenarios, the administration of rGRFT led to a reduction in the number of viable infectious viral particles as well as a decrease in viral load, as determined by qPCR. (Figure 2-10b)

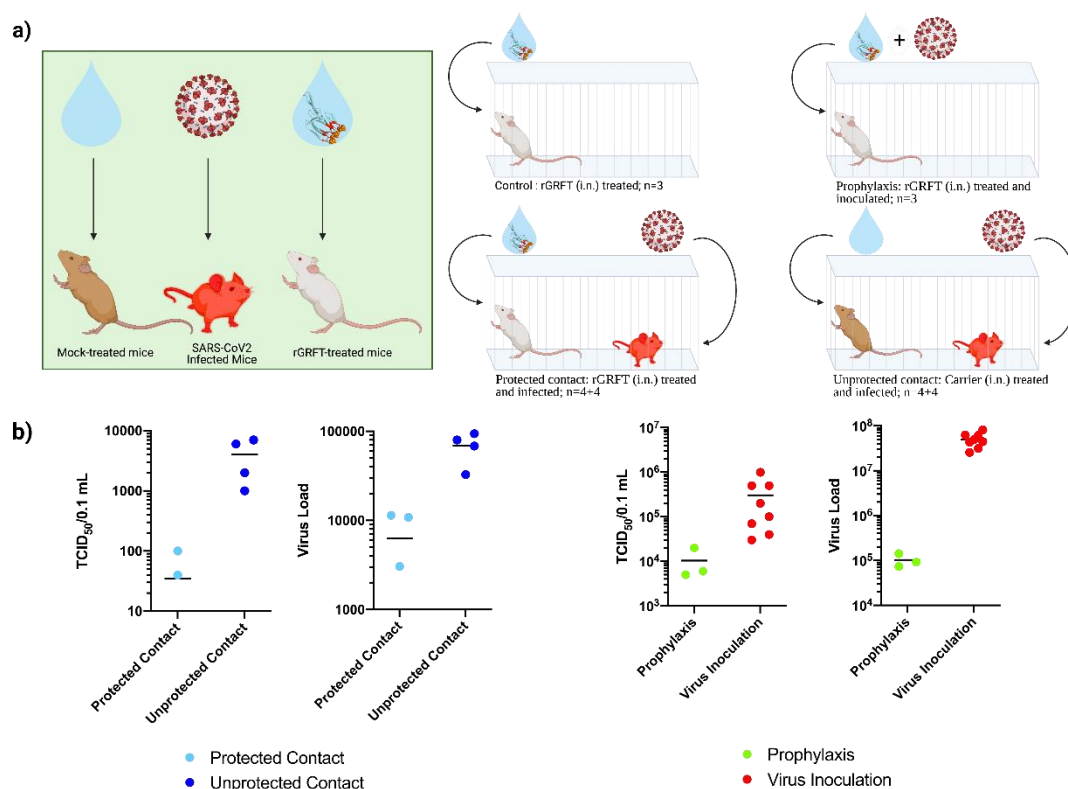


Figure 2-10. *In vivo* antiviral activity of rGRFT against SARS-CoV-2 infection shown for direct or contact-based infection models. (a) IFNAR^{-/-} mice assigned to various study groups. The control group is mice treated with only rGRFT (i.n.) in a cage (n = 3), the prophylaxis group is rGRFT (i.n.)-treated and SARS-CoV-2-inoculated mice in a cage (white mice; n = 3), the protected contact group is consisted of rGRFT-treated mice and SARS-CoV-2-infected mice (white mice and red mice, respectively; n = 4 + 4), and the unprotected contact group is mock (i.n.)-treated mice and SARS-CoV-2-infected mice (brown mice and red mice, respectively; n = 4 + 4). (b) SARS-CoV-2 loads and infective virus titers in lung tissues of individual IFNAR^{-/-} mice assigned to various study groups. rGRFT-treated mice from the protected contact group and mock-treated mice from the unprotected contact group are compared to SARS-CoV-2 loads and infective virus titers; the prophylaxis group (rGRFT-treated and SARS-CoV-2-infected mice) and all SARS-CoV-2-infected mouse are compared to SARS-CoV-2 loads and infective virus titers. The protected contact group values are represented with red color; the unprotected contact group values are represented with blue color; the prophylaxis group values are represented with green color; and the virus inoculation group is represented in orange color.

On the day when the experiment concluded, the presence of SARS-CoV-2 infection in lung tissues was assessed using the *in situ* hybridization test (ISH). In unprotected contact animals, the viral genomic RNA was observed in a more widespread distribution within the lungs compared to the protected group. Conversely, in protected contact animals, the pattern of infection appeared more localized or patchy. The SARS-CoV-2 RNA was detected in various cell types, including bronchial

epithelial cells, alveolar epithelial cells of both type I and type II, as well as in macrophages in all groups. (Figure 2-11)

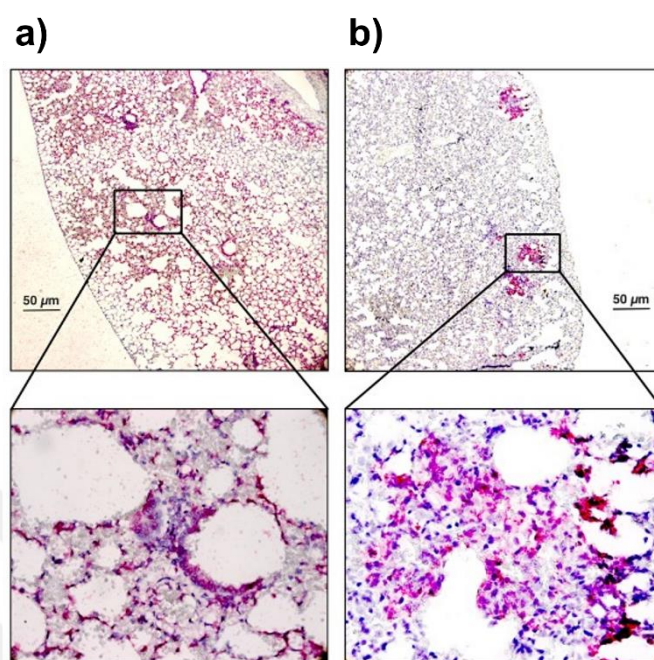


Figure 2-11. *In situ* hybridization of viral RNA in unprotected (b) and protected (a) animals following SARS-CoV-2 infection.

2.4.8. Anti-viral Activity of rGRFT on SARS-CoV-2 Variants

The inhibitory potency of rGRFT was determined against the emerged SARS-CoV-2 variants, specifically Delta and Omicron, *in vitro* with Vero E6 infection model. In order to calculate the IC₅₀ values, the infection experiment was performed using the same procedure as previously detailed for the Wuhan strain. Based on the cytopathic effect (CPE) readings, the IC₅₀ values for rGRFT against the Delta and Omicron variants were determined to be 34.0 and 5.4 nM, respectively. (Figure 2-12) The infection results conclusively demonstrate that rGRFT can effectively hinder the replication of both Delta and Omicron variants of SARS-CoV-2 in Vero E6 cells.

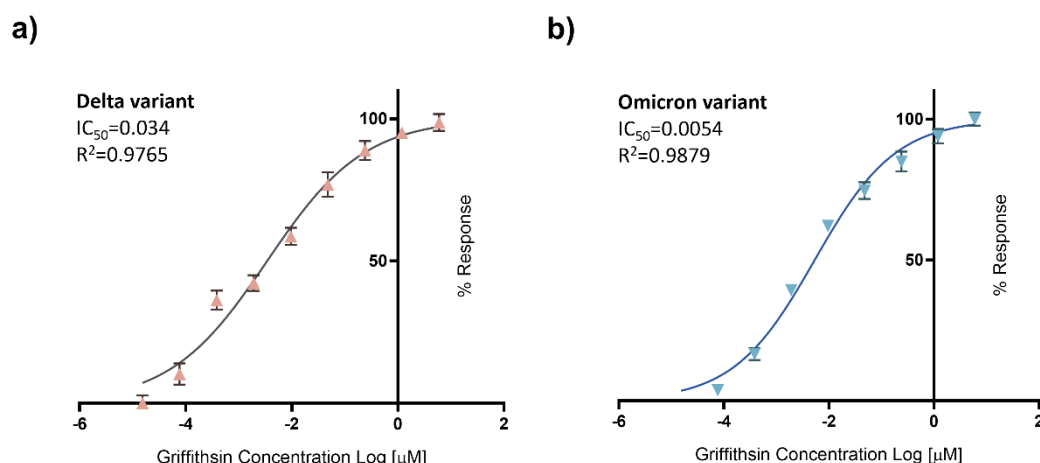


Figure 2-12. rGRFT inhibitory activity against SARS-CoV-2 variants *in vitro* with the Vero E6 infection model. The calculated IC_{50} value of rGRFT for Delta (a) and Omicron (b) variants.

2.5. Discussion

Although, the COVID-19 pandemics is officially ended thanks to effective vaccines that have been developed in the virtue of the relentless efforts of researchers, governments, and pharmaceutical companies, the ongoing evolution of SARS-CoV-2, particularly the mutations occurring in the spike protein region, presents a risk that could lead to reduced efficacy of both vaccines and neutralizing antibodies. [158]–[163] For instance, several virus strains, such as P.1 and B.1.351, have shown the ability to partially evade neutralizing antibodies, as determined through *in vitro* pseudovirus assays. [164] Moreover, the low homology of SARS-CoV-2 Spike protein with other ACE2-recognizing coronaviruses, such as HCoV-NL63, suggests that the S protein can adapt to various mutations without losing its functionality. [148] This concerning scenario indicates the urgent need for the development of broad-spectrum "pan-coronavirus" inhibitory agents to mitigate transmission,

especially in the presence of emerging variants, until a viable vaccine becomes available for them as well.

The spike glycoprotein is extensively glycosylated, featuring a combination of oligomannose and complex-type sugars across its estimated 22 N-linked and 2 O-linked potential glycosylation sites per protomer. In fact, the total glycan structures, as observed when expressed recombinantly in HEK293 cells, contribute to approximately one-third of the molecular weight of the spike protein and cover roughly 40% of the protein's surface (Figure 2-13a). [165], [166] Considering the conservation of sequons among SARS-CoV and SARS-CoV-2 spike proteins[167], we hypothesized that the GRFT protein could potentially serve as an antiviral agent against SARS-CoV-2, building on its previously documented antiviral effects against SARS-CoV. GRFT is characterized as a domain-swapped dimer, adopting a beta prism fold and having the capacity to bind three sugar molecules per protomer (Figure 2-13b). [168]

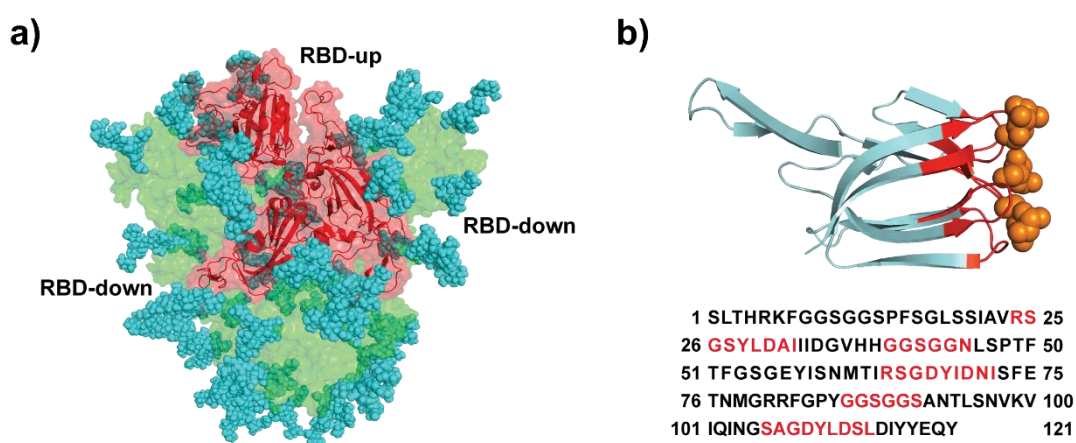


Figure 2-13. Three dimensional structures of SARS-CoV-2 spike protein and griffithsin protein. (a) Structure of HEK293 produced recombinant SARS-CoV-2 spike protein, and the RBD with attached glycan structures colored blue (side view)

Structure and amino acid sequence of rGRFT complexed with mannose sugar which is depicted as orange spheres.

To this hypothesis, we produced and purified the recombinant rGRFT protein in *E. coli*. we showed that the griffithsin protein can bind to the S protein of SARS-CoV-2 and whole viral particle *in vitro*. Furthermore, when employed as a prophylactic agent, we proved that rGRFT effectively inhibits SARS-CoV-2 infection in both the *in vitro* Vero E6 cell line and an *in vivo* mouse model. Toxicity assessments of rGRFT using mouse models have provided reassuring results, indicating that it is a well-tolerated agent even at concentrations exceeding its therapeutic concentration range.

Prior studies have demonstrated that both MERS and SARS infections are hindered by rGRFT administration. [156], [157] Although the S protein of SARS-CoV-2 shares only low-to-moderate similarity with MERS-CoV and SARS-CoV [169], it is theorized that the inhibitory mechanism of rGRFT is primarily rooted in its binding to glycan structures rather than specific amino acid sequence motifs on S proteins. [157] Notably, these glycan structures, unlike conventional amino acid sequences, are conserved between SARS-CoV and SARS-CoV-2. Furthermore, reports have indicated that the glycan sequons of SARS-CoV-2 remained relatively stable during the early stages of the COVID-19 pandemic, spanning from December 2019 to April 2020 [170] This observation aligns with another study revealing that SARS-CoV-2 mutants lacking glycans, particularly at N165 and N234 sequons, displayed reduced affinity for the ACE2 protein. In this study, the authors suggested that the loss of glycans at these specific sites shifted the equilibrium of the receptor-binding domain (RBD) to a "down" position, potentially leading to the development of less infectious

viral particles. [171] Hence, there appears to be a natural selective pressure on SARS-CoV-2 to conserve N-glycans which could be leveraged through rGRFT treatment for hindering the viral transmission.

It has been postulated that rGRFT exerts its antiviral activity against wide spectrum of viruses by Binding to N-linked glycans on viral receptors. Nevertheless, the presence of N-linked glycans on human cells has raised safety concerns regarding the use of GRFT as therapeutics. In contrast to other lectins, such as banana lectin, concanavalin A, and cyanovirin-N (CV-N), which have exhibited unwanted side effects, GRFT has consistently proven to be a safe microbicide in numerous studies. These studies have encompassed *in vitro* cell culture experiments [156], [172], *in vivo* experiments with macaques [173], and a human phase I trial. [174] Moreover, while there is evidence of GRFT binding to human cells, including peripheral blood mononuclear cells (PBMC) and human epithelial cells, this binding has not elicited mitogenic, immunogenic, or toxic effects on the cells [172] Additionally, *in vivo* experiments conducted on nonhuman primates and clinical trials involving human subjects have revealed that the administration of GRFT through the vaginal route did not induce any proinflammatory or toxic responses. [173], [174] These findings collectively reinforce the safety profile of rGRFT as a therapeutic modality.

Beyond its notable efficacy and safety profile, it's worth noting that the production cost of rGRFT is reported to be exceptionally low, estimated at \$3500 per kilogram, with a yield of approximately 2.5 grams per liter in *E. coli* at a scale of tons. [175] Furthermore, rGRFT exhibits remarkable stability, with the ability to maintain its integrity at room temperature in phosphate-buffered saline (PBS) for up to two years without forming aggregates or undergoing degradation. [176] Considering the board

activity of rGRFT on different coronaviruses, the evolutionary cost of glycan loss on SARS-CoV-2, the available high-yield recombinant production strategies for rGRFT, and the good stability of rGRFT at room temperature, rGRFT may provide a possible solution for the emerging variants.



CHAPTER 3

Genetic Technologies to Modify Probiotic *E. coli* Nissle Strain

3.1. Objective of the Study

E. coli Nissle (EcN) is a widely used probiotic bacteria for treatment of gastrointestinal pathologies with record long safety history in humans. In the last decade, with the emerging of synthetic biology, EcN strain also is one of the most engineered bacteria for developing living therapeutics for diverse indications, including rare metabolic diseases and cancer. Yet, stable recombinant DNA transfer in EcN involves homology direct insertion of foreign DNA into genome through recombineering which is a laborious and time-consuming method. Here, we explored the use of cryptic plasmids of EcN, pMUT1 and pMUT2, as vectors to deliver recombinant DNA stably. For this aim, the cryptic plasmids were removed from the EcN strain by using CRISPR-Cas system meanwhile cryptic plasmids were redesigned as DNA vectors for the cloning of foreign DNA in order to express recombinant proteins.

3.2. Introduction

EcN is a probiotic strain of *E. coli* species isolated from a German soldier who do not develop diarrhea in Shigella contaminant area during World War I. The early use of EcN as a remedy encompasses intestinal infections because of the strong antagonism of EcN to pathogenic enterobacteria strains. [177] Till then, EcN is widely used in

the treatment of several gastrointestinal pathologies because the natural capabilities of EcN provide therapeutical actions required to interfere disease mechanism and ensure the safety in human use. EcN secretes siderophores, microcin and polyketides whose natural function is to provide competitive advantage for EcN to survive in the gut however antagonist effects against common pathogens are stemmed from this ability [178] Taxonomically, although EcN strain is related to UPEC strains of *E. coli*, yet it is lack of any known virulence factor. Also, there is no reported pathogenesis or infection by EcN at the urinary tract or other sites of the body. [179] EcN strains carry a O6:K5:H1 serotype but O6 antigen is truncated due to detrimental mutation O6 antigen polymerase gene, and truncated O6 is hypothesized to be non-immunotoxic. [180] Despite the molecular pathway for colibactin, genotoxic DNA alkylating agent, is present in EcN genome, recent report showed that EcN do not possess such activity. [181] In addition, EcN strain is sensitive to mammalian serum non-specific complement system. [178]

The inherent capabilities and safety profile recorded over hundred years in human use make the EcN strain the excellent candidate as a living chassis for designing sentinel cells genetically with desired therapeutical functions. Extensive research on EcN have demonstrated that they possess many other therapeutical activities beyond their antagonist effects against intestinal pathogens. [182]–[185] Aside from the use of WT EcN as therapeutics, many studies have aimed to extend capabilities of EcN by genetically designing the bacteria through synthetic biology approaches to perform desired tasks. For instance, a biosensor was designed and implemented in the EcN to detect presence of cancer by releasing color substrates in urine as sensor output. [186] In addition, EcN was reprogrammed genetically to release synthetic antigens synchronically at tumor cores for guiding CAR-T cells to the tumors. [187]

Aside from academic works, EcN-based therapeutics are also attracted companies to develop engineered microbes for variety of conditions. Currently, the company Synlogic have invested heavily on the development of genetically engineered EcN strains for various indications including, phenylketonuria, homocystinuria, enteric hyperoxaluria, cancer, and undisclosed immunological diseases. [188]–[191]

Stable genetic material transfer in EcN currently was accomplished by insertion of recombinant DNA into genome of the bacteria via recombineering with λ phages enzymes. As a process, recombineering is labor intensive and time consuming that requires many cell passages to obtain positive colonies. To circumvent the genomic insertion for stable gene transfer, additive plasmid vectors and auxotroph strains have been developed that enables to maintain the transferred recombinant genetic materials over generation. [192], [193] However, these strategies can disturb the metabolism of the organisms by putting a significant burden on cells. Also, fitness of the engineered cells might negatively be affected in the gut because of the burden created by synthetic plasmids.

Wild type EcN strains contain two cryptic plasmids, namely pMUT1 and pMUT2, which are stably transferred to daughter cells due to the presence of ColE-like origin of replication in both plasmids. [194] In addition, these plasmids have been suggested not be conjugated to other cells as they are only detected in EcN. [178] Previous studies have shown that both of the cryptic plasmids are redundant for EcN cells thereby indicating that plasmids can be removed without affecting the phenotype of cells. [195] Therefore, cryptic plasmids of EcN, pMUT1 and pMUT2, provide an excellent platform to transfer recombinant genetic materials into EcN.

In this study, we aimed to explore the utilization of pMUT1 and pMUT2 for stably transfer recombinant genes into EcN strains. For this purpose, both pMUT1 and pMUT2 were cured from EcN strain by using CRISPR-Cas. Subsequently, both of these plasmids were designed to express a recombinant protein, sfGFP protein, in the Δ MUT1 Δ MUT2 EcN strain.

3.3. Materials and Methods

3.3.1. Construction of Plasmids Used in the Study

All plasmids used in this study were constructed with Gibson Assembly. Plasmid sequences were verified with either Sanger sequencing or NGS. All enzymes used in this for plasmid construction were purchased from New England Biolabs (NEB). Primers were ordered from Oligomer Inc. and BiOligo Inc. If needed, gene synthesis service of Genewiz Inc was utilized. All plasmid maps, primers, and sequencing results were given in Appendixes. LB media and LB agar with appropriate antibiotics were used for bacteria cultivation unless otherwise noted. Bacteria cells were grown in a shaking incubator with liquid media if not otherwise indicated.

To construct pET22b CmR pLacO mock CRISPR array, pET22b with chloramphenicol resistance was amplified with pREA196 and TR-PK-REV meanwhile the CRISPR array was constructed by amplifying the pLacO promoter via pREA140 and pREA195. Both of the PCR products carried compatible overhangs for Gibson Assembly. The desired plasmid was constructed by assembling PCR products with Gibson Assembly.

To clone the gRNAs into CRISPR array, the pET22b CmR pLacO mock CRISPR array plasmid was digested with BsaIV2-HF enzyme. The linearized plasmids were

further treated QuickCIP to prevent self-ligation. CIP was heat activated at the end of the reaction. The gRNAs were synthesized as complementary primers and hybridized in 100mM NaCl supplement TE buffer by heating the solution to 95 °C then gradually cooling the room temperature. The hybridized primers were treated polynucleotide kinase to phosphorylate their 5' for ligation. The BsaI linearized plasmid and the hybridized primers coding the gRNA were ligated with T4 DNA ligase. The colony PCR with pREA161 and SEQ22R were formed from the transformants to verify the ligation. For pET22b targeting E1 gRNA, pREA197 and pREA198 primers were used. For pMUT1 targeting E2 gRNA, pREA199 and pREA200 primers were used. For pMUT2 targeting g9 gRNA, pREA211 and pREA212 primers were used. For pMUT2 targeting g10 gRNA, pREA213 and pREA214 primers were used.

To clone the relB antitoxin into CRISPR Array plasmid targeting the pMUT2, the RelB gene was amplified from *E. coli* genome with pREA215 and pREA216, and the transcription terminator was amplified with pREA217 and pREA218. The array plasmid containing the g9 and g10 gRNA was digested with AatII for linearization. The linearized plasmid and PCR products were assembled with Gibson Assembly.

To build the pMUT1 with sfGFP expression cassette, pMUT1 was amplified from EcN with pREA221 and pREA222; meanwhile, ampicillin resistance cassette and sfGFP expression cassette were amplified from an AmpR containing plasmid and the plasmid library containing the constitutive active fluorescent expression cassette, respectively. pREA203 and pREA204 were used to amplify the AmpR cassette, and pREA205 and pREA206 were used for sfGFP expression cassette. The PCR products were assembled by using Gibson Assembly.

To assemble the pMUT2 with sfGFP expression cassette, linearized pMUT2 was obtained with PCR using pREA219 and pREA220 from EcN. The linearized pMUT2 and the same expression cassettes for AmpR and sfGFP from pMUT1 cloning were ligated to each other using Gibson assembly.

3.3.2. Curing of pMUT1 and pMUT2 from EcN strain

WT EcN cells were prepared for electroporation. To do so, cells at mid-log phase were collected and washed with two times with cold pure water, two times cold 50% glycerol, and after washing, cells were dissolved in 50% glycerol by concentrating cells 100x fold. pCas plasmid[196] was transformed into WT EcN strains via electroporation. Built-in protocol in Bio-rad Pulser X-system was utilized. Following electroporation, cells were recovered in SOC media and spread on agar plate with appropriate antibiotics. pCas carrying EcN cells were transformed with either CRISPR array plasmid with E1 and E2 or CRISPR array plasmid with E1, g9 and g10 for curing of pMUT1, and pMUT2. During the recover step, cells were induced with IPTG for expression of gRNAs. Cells were spread on agar plate with kanamycin and incubated at 30°C overnight. Next day, the colony PCR performed the single colonies for screening the curing of plasmids. The confirmed colonies were grown at 42 °C overnight to cure pCas plasmid.

3.3.3. Colony PCR for pMUT1 and pMUT2

Single colonies were picked with 10 µl pipette tip. To amplifying regions from plasmids, pmut5 and pmut6 primers were used for pMUT1, and pmut7 and pmut8 were used for pMUT2. [194] PCR products were analyzed with agarose gel electrophoresis.

3.3.4. Stability of pMUT1 and pMUT2 in Δ pMUT1 Δ pMUT2 EcN Strain

pMUT1-sfGFP and pMUT2-sfGFP were transformed separately into Δ pMUT1 Δ pMUT2 EcN Strain via chemical transformation. Single colonies were picked for each transformation and grown overnight in the presence of ampicillin. Next day, cells were diluted 10000-fold into fresh media without antibiotics and grown for 24 hours. This step was repeated for 5 consecutive days. In each day, whole cell fluorescence was measured and normalized by OD₆₀₀. M5 spectrophotometer was used for fluorescence measurements.

3.4. Results

3.4.1. Construction of gRNA Expression Plasmid

gRNAs were expressed from pLacO controlled CRISPR array that contained a nontargeting mock gRNA with BsaI site. (Figure 3-1a) To construct the plasmids, pET22b vector and pLacO promoter were amplified with PCR. Primers were designed to add CRISPR array as homolog overhang to both of the fragments for assembly. Expected bands were observed in the agarose gel electrophoresis which are purified and then assembled with Gibson Assembly. (Figure 3-2b)

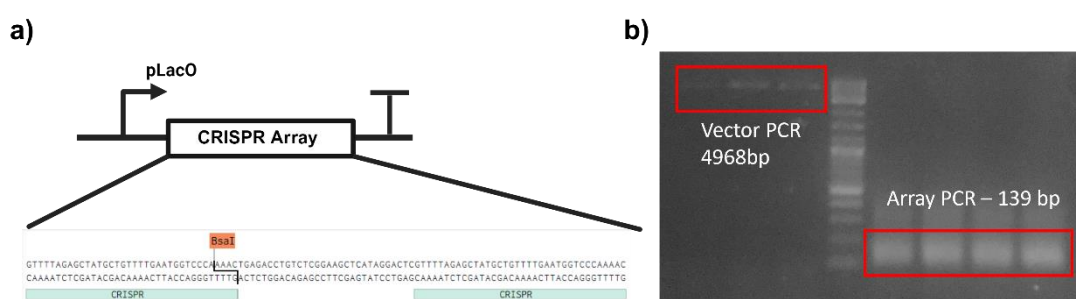


Figure 3-1. Construction of pLacO mock CRISPR Array plasmid for expression of gRNA in the E. coli cells. a) Depiction of expression module for transcription of

CRISPR array. The BsaI restriction cut site was highlighted in the mock gRNA. b) PCR results of pET22b amplification and pLacO amplification which have compatible ends for Gibson Assembly. All DNA bands appeared at their expected sizes.

3.4.2. Cloning of gRNAs targeting the pMUT1 and pMUT2 into CRISPR Array Plasmid

gRNAs were cloned into mock CRISPR array under control of pLacO. Similar cloning strategy described in [197] was employed. The plasmid was cut with BsaI to linearize the vector and to create sticky overhangs. CIP was used to remove the 5' phosphates thereby preventing self-ligation of digested vectors. Meanwhile, gRNA sequences with same sticky ends were generated with in vitro hybridization. 5' phosphates were added to hybridized primers with PKN reaction for ligation with the vector. As the BsaI recognition site remained untouched in the mock gRNA sequence, after the successful ligation, same workflow could be performed to clone next gRNA (Figure 3-2a). The strategy was successfully conducted to construct gRNA plasmids that would be used for curing experiments. Each addition of gRNA sequence to CRISPR array resulted in the extension of array sequence by 66 base pair which was detected via colony PCR. (Figure 3-2b, 3-2c, 3-2d, and 3-2e) In addition to gRNA cloning, relB antitoxin was cloned to g9/g10 gRNA CRISPR array vector since it was recommended in a previous study for curing of pMUT2 which encodes a relB like toxin system. [198]

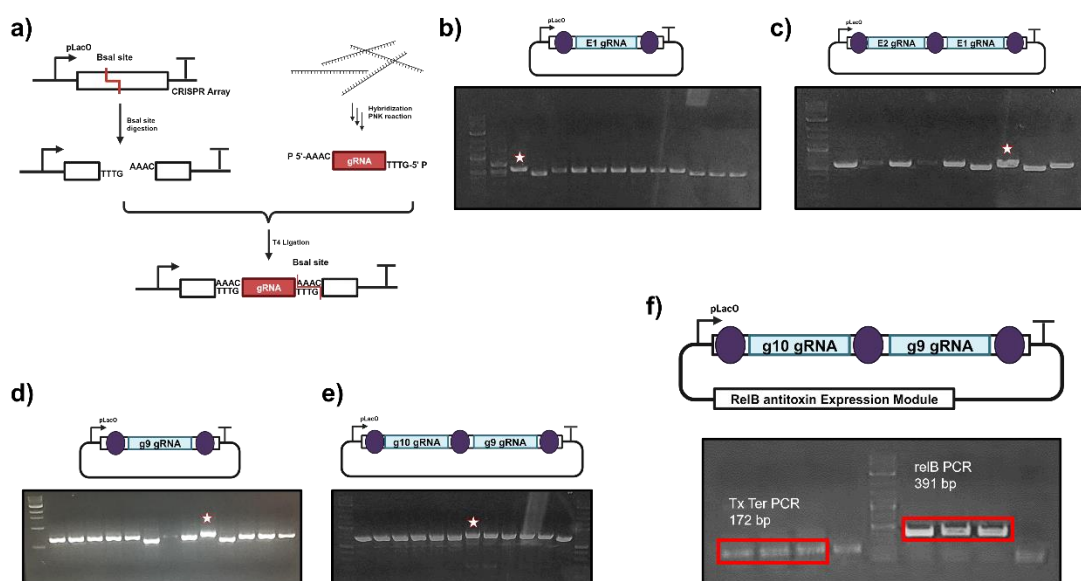


Figure 3-2. Insertion of gRNAs into CRISPR array. a) Depiction for gRNA cloning into CRISPR array plasmid. Colony PCR results verify cloning of E1 gRNA into mock CRISPR array (b), E2 into E1 CRISPR array (c), g9 gRNA into E1 CRISPR array (d), and g10 into g9/E1 CRISPR array (e). Positive colonies were marked with white stars f) Cloning of relB expression construct to g9/g10/E1 CRISPR array vector. PCR products of relB with its native promoter and a transcription terminator, both carrying homolog overhangs to each other and the vector for cloning via Gibson Assembly. All DNA bands appeared at their expected sizes.

3.4.3. Curing of pMUT1 in WT EcN strain

CRISPR-Cas system was used to eliminate the plasmids inside WT EcN. Briefly, cells were transformed with two plasmids pCas and CRISPR Array. pCas encodes for Cas9 protein and trans-activating RNA under control of its native operon promoter. In addition, the origin of replication of pCas is heat sensitive which enables the elimination of pCas plasmids from cells by growing at temperatures. The CRISPR array plasmid encodes the gRNAs between CRISPR sequences whose

expression controlled by pLacO. CRISPR array plasmids are used to express gRNAs specific to CRISPR array plasmids and either pMUT1 or pMUT2. Thus, cells carrying both plasmids, when induced by IPTG, are expected to express gRNAs that guide the Cas9 the CRISPR array plasmid, and either pMUT1 or pMUT2 based on the second gRNA encoded in the CRISPR array plasmid for cutting. Following the restriction of plasmids by Cas9, plasmids will be degraded by intracellular exonuclease which causes the loss of plasmids. (Figure 3-3) Subsequently, pCas can be cured from cells by growing cells at higher temperatures.

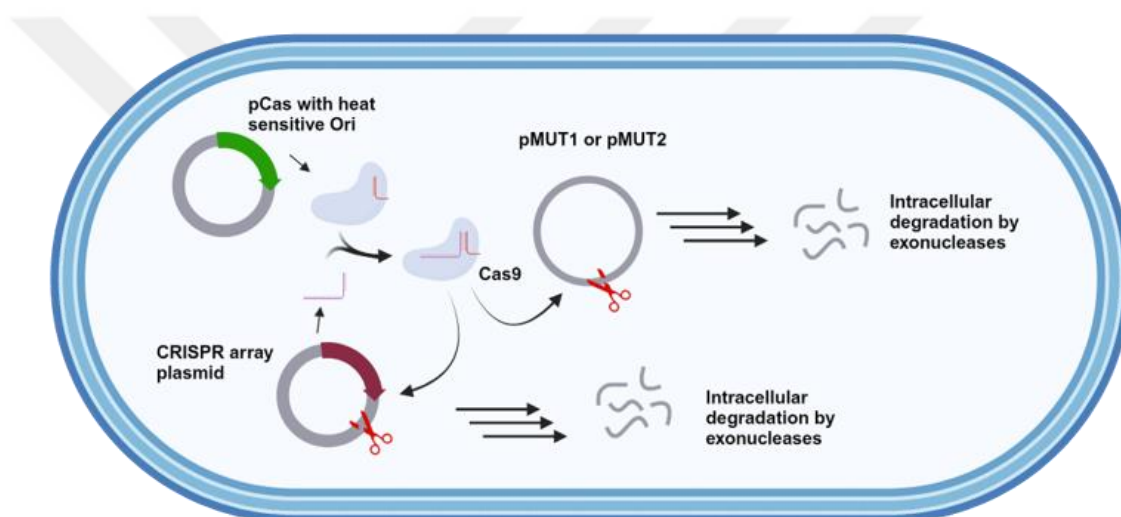


Figure 3-3. Illustration of plasmid curing with pCas. The Cas9 protein loaded with either pMUT1 or pMUT2 targeting gRNA cuts the corresponding plasmid based on the loaded gRNA. The digested plasmids are degraded by intracellular exonuclease thereby resulting in the loss of the plasmid.

The strategy was employed to cure pMUT1 from WT EcN cells. Both pCas and CRISPR array plasmid with E1 and E2 gRNAs were transformed into cells. (Figure 3-4) Subsequently, transformant cells were induced with IPTG, and resulting colonies after induction were screened with colony PCR for presence of pMUT1 and pMUT2. All the colonies were carrying the pMUT2 based on PCR results as

expected whereas there was no band for pMUT1 in one colony thereby indicating that the plasmid was lost. (Figure 3-4) The single colony lacking the pMUT1 was picked then grown at 42 °C to eliminate pCas. Finally, the isolated colony was checked where it was growing in the presence of chloramphenicol or kanamycin. As expected, cells failed to grow in the media with chloramphenicol or kanamycin, showing that cells were lack of CRISPR array plasmid and pCas.

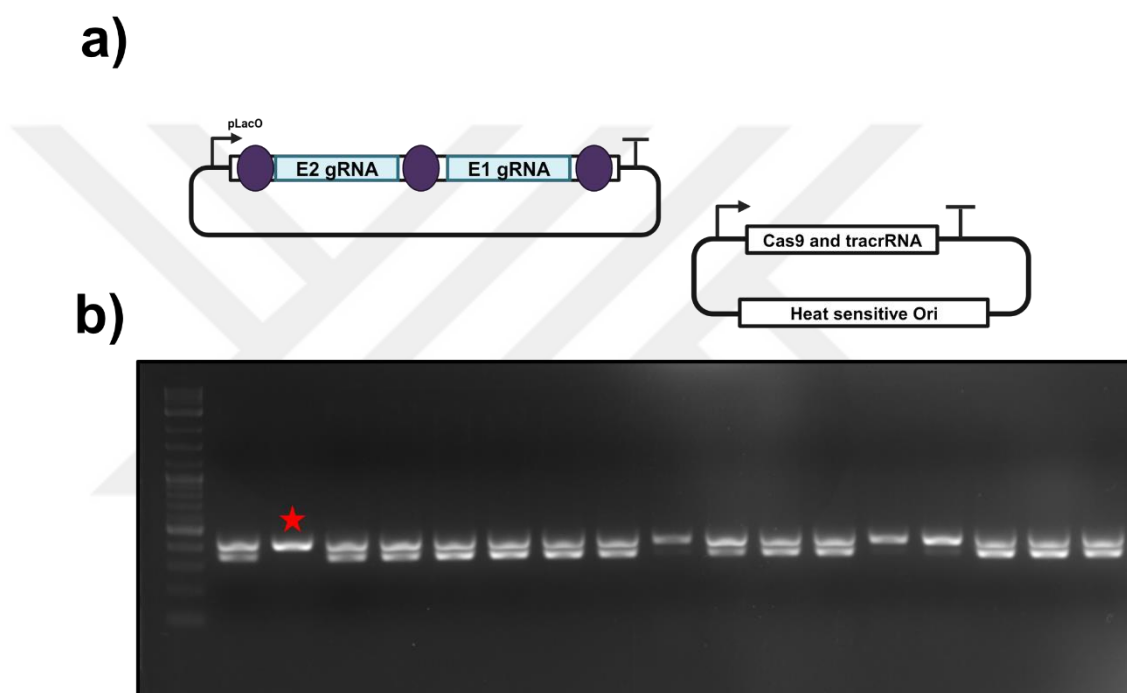


Figure 3-4. Curing of pMUT1 from the WT EcN strain. a) Depiction of plasmids transformed into the cells and b) Colony PCR results for amplification of pMUT1 (361bp) and pMUT2 (427bp). All DNA bands appeared at their expected sizes.

3.4.4. Curing of pMUT2 in Δ pMUT1 EcN strain

Δ pMUT1 EcN strain was co-transformed with CRISPR array plasmid with g9/g10/E1 and pCas plasmid. (Figure 3-5a) Cells were cured from pMUT2 by inducing gRNA expression. The grown single colonies were screened with colony

PCR to find the colony which was lack of pMUT2 however all colonies were positive for pMUT2. (Figure 3-5b) A previous study was suggested that pMUT2 might contain a relB-like toxin-antitoxin system and expression of relB toxin during the curing process resulted in successful elimination of pMUT2. [198] Therefore, relB antitoxin was expressed from CRISPR array plasmid. (Figure 3-5c) Following the curing experiments, all the resulting colonies were negative for pMUT2 suggesting the loss of pMUT2. (Figure 3-5d) Negative colonies were grown and cured from pCas by using the same procedure described above.

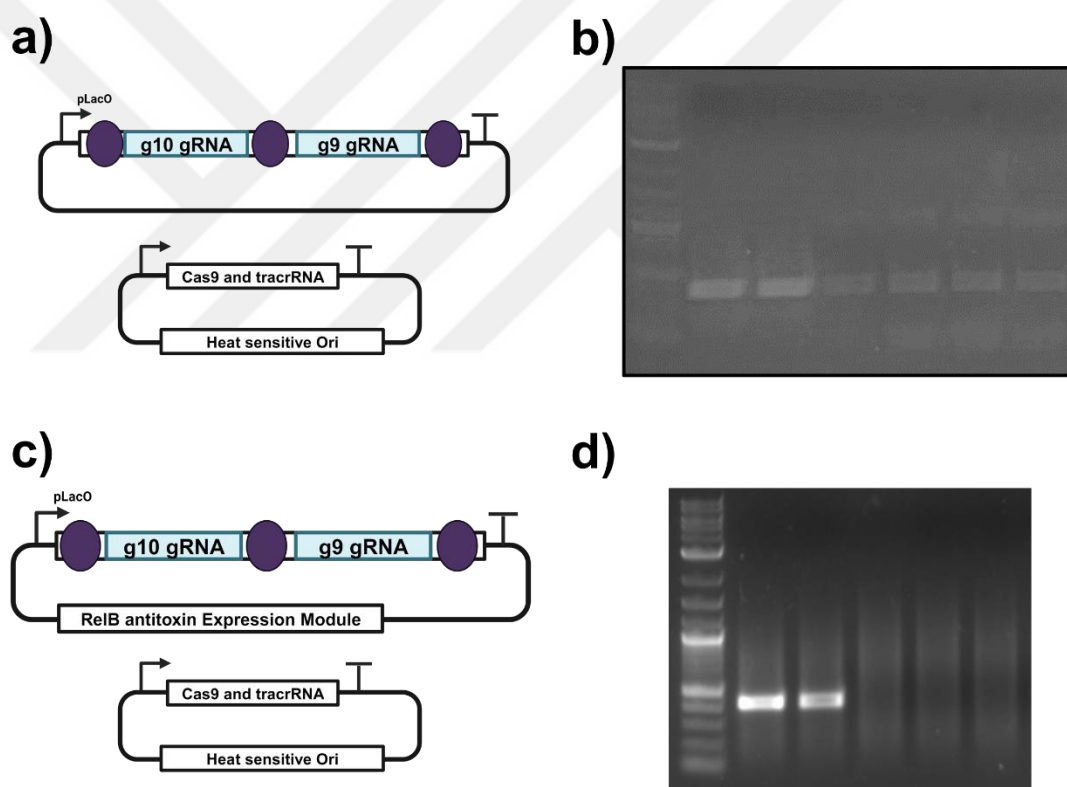


Figure 3-5. Curing of pMUT2 from the Δ pMUT1 EcN strain. a, b) Plasmid curing trial with g10/g9 gRNA plasmid without relB antitoxin. Transformed plasmids were illustrated at the (a) while the colony PCR results to screen presence of pMUT2 is shown at the (b). c, d) Deletion of pMUT2 using g10/g9 gRNA plasmid

with relB antitoxin. The plasmids that are transformed into cells were depicted in the (c) and colony PCR results for pMUT2 were shown in the (d). In both of the agarose gel results, second and third lanes were PCR results of WT EcN and Δ pMUT1 EcN strain, respectively, as controls. All DNA bands appeared at their expected sizes.

3.4.5. Construction of Recombinant pMUT1 and pMUT2 with Constitutively Active sfGFP Expression Cassette (pMUT1-sfGFP and pMUT2-sfGFP)

Recombinant versions of pMUT1 and pMUT2 were designed and constructed. To select the plasmids following the transformation, ampR resistance cassette was placed in the vector whereas expression profile of the vector was tested by cloning a constitutively active sfGFP expression module under control of proD promoter. (Figure 3-6a) The pMUT2 and pMUT1 plasmids were amplified with PCR with Gibson overhangs to clone the cassette in the vector. (Figure 3-6b and 3-6c) AmpR and sfGFP cassette were also obtained with PCR by adding homologous regions for Gibson Assembly. (Figure 3-6d and 3-6e) The expression cassette was cloned into both pMUT1 and pMUT2 via Gibson Assembly.

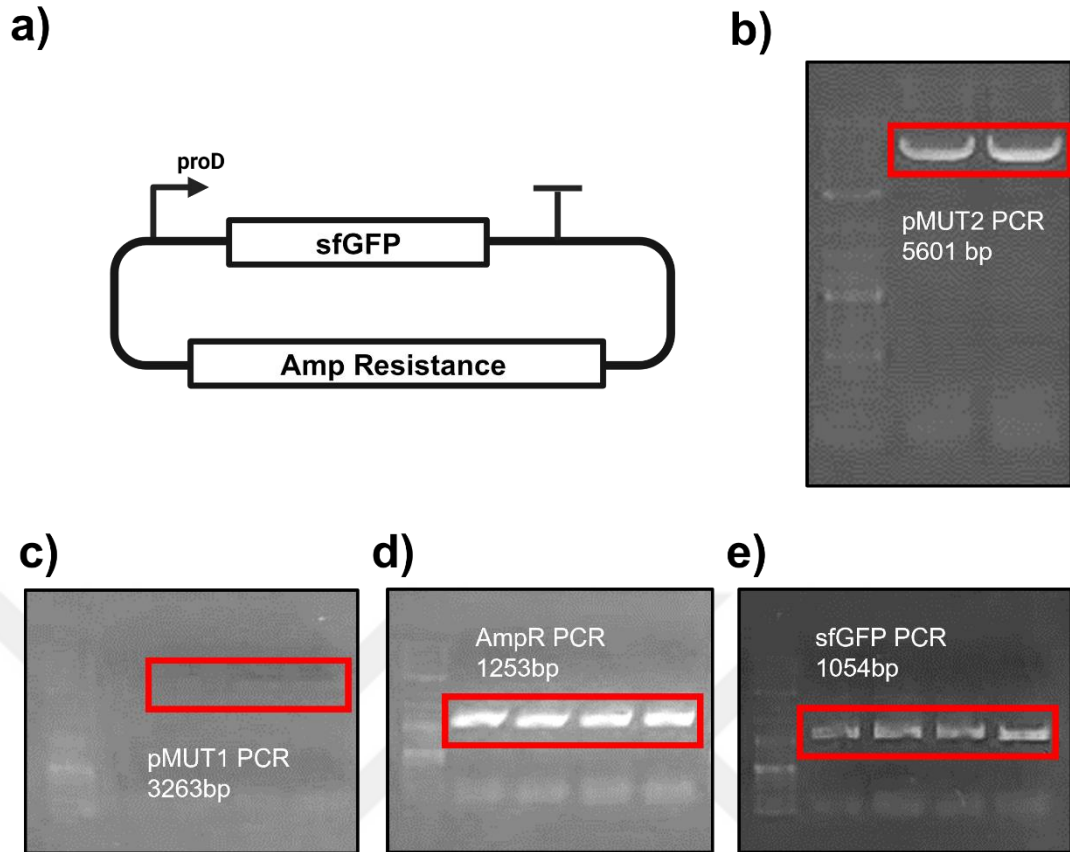


Figure 3-6. Recombinant cryptic EcN plasmids with the constitutively active sfGFP expression construct. a) Schematic representation of recombinant pMUT1 and pMUT2 plasmids with ampicillin resistance and sfGFP expression cassette. b, c) PCR results for amplification of pMUT2 (b) and pMUT1 (c) to construct recombinant plasmids with Gibson assembly. d) Amplification of AmpR cassette via PCR. e) PCR products of sfGFP amplification. All DNA bands appeared at their expected sizes.

3.4.6. Stability of pMUT1-sfGFP and pMUT2-sfGFP in Δ pMUT1 Δ pMUT2 EcN Strain

The pMUT1-sfGFP and pMUT2-sfGFP plasmids were transformed into Δ pMUT1 Δ pMUT2 EcN Strain individually to test their stability *in vitro*. Single colonies were picked from the agar plate and grown overnight in the presence of ampicillin. Next day, the cells were rediluted into fresh media without any antibiotic by 10.000-fold. The diluted cells were grown overnight again. This step was repeated for five days. In each, green fluorescence of cells bearing either pMUT1 or pMUT2 was measured and normalized with OD₆₀₀. The whole cell fluorescence of cells was revealed that both pMUT1 and pMUT2 were stably inherent to daughter cells in absence of any selection pressure. (Figure 3-7)

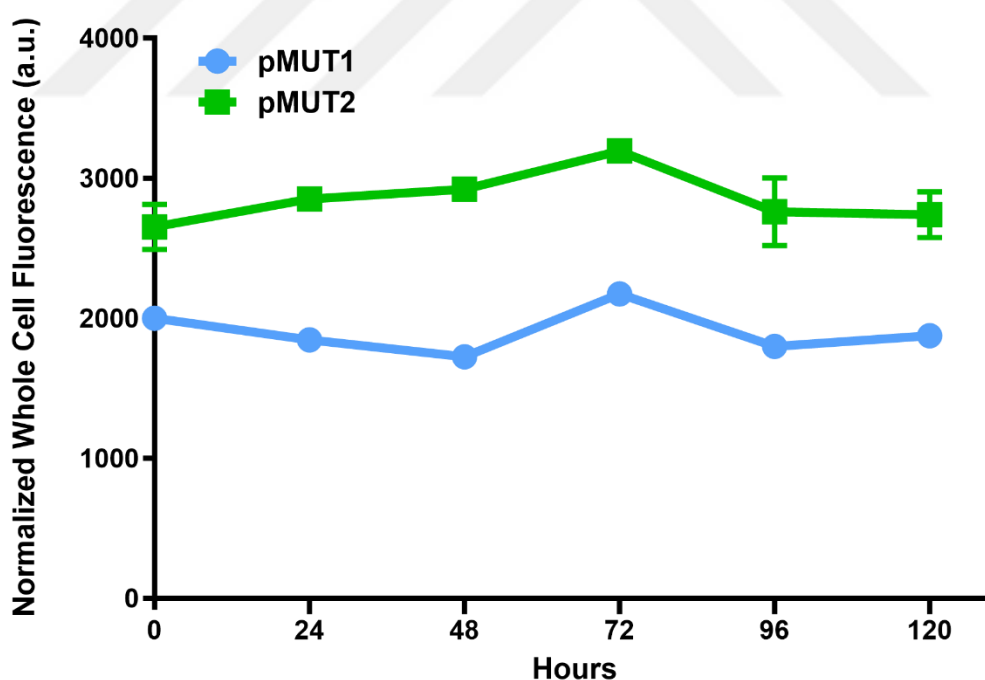


Figure 3-7. Stability of recombinant pMUT1 and pMUT2 in vitro in the Δ pMUT1 Δ pMUT2 EcN Strain. Cells were transformed with either of the plasmids and grown for five days by consecutive dilution into fresh media without any

selection pressure. The whole cell fluorescence of cells was shown. Experiments were performed with triplicates.

3.5. Discussions

Here, we have demonstrated that cryptic plasmids can be used as vector to transfer foreign DNA into EcN stably for expression of recombinant proteins. For this purpose, the Δ pMUT1 Δ pMUT2 EcN strain was created by curing the cryptic plasmid with CRISPR-Cas system. Meanwhile, both pMUT1 and pMUT2 were developed as DNA vectors by cloning of the AmpR resistance cassette for selecting transformant colonies as well as a constitutively active full gene expression unit with compatible restriction sites for modularity. When transformed into EcN, both pMUT1 and pMUT2 were proven to provide stable expression of a reporter protein, sfGFP, over 5 days without any selection pressure for the vector.

The CRISPR-Cas system is shown to be an effective tool for genome modifications in both bacteria and mammalian cells. Here, it is demonstrated that the system can be exploited to eliminate plasmids from bacterial cells by inducing double strand break on the plasmid thereby causing the intracellular degradation. The native CRISPR array is used to drive the expression of synthetic gRNAs rather than the synthetic single guide. The streamlined cloning protocol enables cloning multiple gRNA into the same transcriptional unit in the same plasmid. This provides ease of the burden on gRNA cloning for curing because the need of cloning additional promoters and transcriptional terminators to drive multiple sgRNAs are circumvented. In addition, both tracrRNA and Cas9 protein are expressed from the native promoter in order to prevent off-target Cas9 toxicity. [199]

The stable genetic transfer into EcN is required for clinical applications to sustain the therapeutical actions following administration yet common plasmid vectors which are sustained in the cells with antibiotic pressure are rapidly cured from cells in the absence of the antibiotics. [200] Thus, insertion of foreign DNA directly to the genome are preferred for the clinical applications of engineered EcN. [201] Yet, recombineering which is the genetic technology generally used to alter the bacterial genome, is cumbersome technique with low efficiency. [202] Moreover, in many cases more than one copies of the recombinant DNA are inserted into genome for preventing event of accidental loss through natural evolution as well as for increasing the overall expression of the recombinant payload. [201] The use of cryptic plasmids offers an opportunity to bypass many drawbacks of preferred methodologies for stable gene transfer. Firstly, both of the cryptic plasmids are inherently stably transferred to the new cells even in the absence of any selection pressure. In addition, Multicopy nature of both plasmids enable to maximize the expression of payload. Lastly, recombinant plasmids with new payloads can be assembled *in vitro* and transformed into EcN with few steps thereby providing a fast and reliable process to construct and optimize recombinant EcN strains. In conclusion, cryptic plasmids of EcN can be used vector to transfer recombinant DNA stable into the bacteria thereby easing the genetical engineering of EcN for therapeutical applications.

CHAPTER 4

Synthetic Secretion Systems for *E. coli* Nissle

The work described in this section is partially published in the following study [203]:

Recep Erdem Ahan¹, Cemile E. Özçelik¹, İrem Niran Çağıl¹, Urartu Özgür Şafak Şeker^{*,1}

1 UNAM-Institute of Materials Science and Nanotechnology, Bilkent University, Ankara, Turkey

*Corresponding authors: Urartu Özgür Şafak Şeker; E-mail: urartu@bilkent.edu.tr

Reprinted with permission of authors from following publications;

<https://doi.org/10.1101/2023.06.14.544856>

4.1. Objective of the Study

Various species of bacteria can thrive and establish colonies in different regions of the human body. Bacteria naturally residing in disease sites, such as tumors and the gut, can be engineered to facilitate targeted protein delivery, potentially enhancing the effectiveness of protein-based therapies. Therefore, there is a need for a versatile,

efficient, and safe system for protein secretion, especially in gram-negative bacteria (GNB) with inherent therapeutical potential, which possess a second cell wall. For this, we propose an approach called iLOM-SS, short for "inducible Leaky Outer Membrane-based Secretion System," designed for the secretion of proteins in gram-negative bacteria (GNB). In iLOM-SS, the outer membrane of GNB is made permeable by conditional and reversible suppressing of structural proteins thereby allowing the unrestricted diffusion of cargo proteins expressed in the periplasm. *In vitro* experiments demonstrated the successful secretion of various proteins, including enzymes and a human cytokine, by EcN using iLOM-SS. Further characterization of iLOM-SS in ECN revealed its capacity for rapid and controllable secretion, conditional and reversible activation by small molecules, and the functional integration into different genetic circuit architectures. We anticipate that this system will open new avenues to the design of gram-negative bacteria for protein secretion, enabling a wide range of applications, including the development of sentinel cells for therapeutic purposes.

4.2. Introduction

The human body is inhabited by a diverse community of microbes from a very young age, and the composition and abundance of these microbes play pivotal roles in human health. [204] Human-associated microbes which are constantly interacting with the body, are being engineered using synthetic biology techniques to detect disease-related pathological conditions and respond to the detected signals in order to ameliorate the disease. [205]–[209] Engineered microbes with embedded genetic logic circuits can be used to deliver multiple therapeutic protein-based payloads to

specific regions in the body at specific times. This makes them a powerful platform for developing new treatments for a variety of diseases and conditions. [210]–[212]

Protein-based therapeutic payloads need to be secreted across the cell membrane in order to exert their effects, but this is a critical challenge for GNB such as EcN strain due to their outer membrane. GNB possess a variety of protein secretion systems that can transport proteins across the membrane in one or two steps. In Type I, III, IV, and VI secretion systems, proteins are directly secreted from the cytoplasm to the extracellular matrix. In Type II, V secretion systems, and the curli biogenesis system, proteins are first translocated to the periplasm by the Sec or Tat translocon before being secreted to the extracellular matrix. [213], [214]

Using the natural secretion system of bacteria to move proteins out of the cell has many problems that make it difficult to develop protein-secreting GNB-based living bacterial therapeutics. First, all systems require the recombinant expression of the secretion system itself, and many of these systems are associated with pathogenicity. [215]–[218] Also, the construction and use of these nanomachines impose a significant burden on the metabolism of the cells. [219]–[221] In addition to the burden on cellular metabolism, repurposing native secretion systems for extracellular protein translocation has other drawbacks. For example, these systems have narrow substrate specificity, meaning that only a small subset of proteins can be secreted using these systems. [222]–[224] Furthermore, secreted proteins will carry a secretion tag that can affect their function or cause unwanted inflammatory responses. [225] Finally, deploying an optimized secretion system from one species to another is generally not possible, as these systems require many additional elements, such as chaperone proteins, for proper function. [222], [224] As a result,

producing secretory proteins in non-model Gram-negative bacteria with characterized secretion systems often requires time-consuming optimizations that rely heavily on trial and error.

To overcome the limitations of native protein secretion systems in Gram-negative bacteria, researchers have developed alternative methods. Din et al. showed that intracellularly expressed recombinant proteins can be released into the extracellular environment by expressing a bacteriophage lysis gene.[226] The approach was utilized in different applications many of which are related to cancer therapy. For instance, nanobodies were secreted to block the immunosuppression at vicinity of tumors to induce strong anti-tumor response. [227] However, this approach may not be suitable for all applications, such as those requiring anti-inflammatory activity, as it releases dsDNA and other intracellular components that can cause inflammation. Additionally, the sudden release of payload protein can lead to local drug toxicity and prevent sustained therapeutic levels. In addition, we also developed a controlled protein release system in GNB that relies on enzymatic cleavage of surface-displayed proteins with a site-specific protease. This system has a titratable protein secretion profile, meaning the amount of secreted protein can be tuned with a small molecule, such as anhydrotetracycline (aTc). This system is simple to implement and can be functionally integrated into complex genetic circuits. However, the main disadvantage of our system is the use of TEV protease, whose activity is affected by pH, temperature, and other environmental factors. [210] While these efforts have improved some pitfalls of native systems, proposed secretion systems and their use are still limited to certain scenarios. Therefore, there is a need to develop an efficient protein secretion system that can be utilized for a wide range of applications.

In this chapter, I will present a new synthetic protein secretion system called iLOM-SS, which stands for inducible Leaky Outer Membrane based Secretion System. In iLOM-SS, small RNAs repress structural proteins at the periplasm to slightly destabilize the outer membrane integrity, thereby releasing proteins translocated to the periplasm into the extracellular milieu via free diffusion. We constructed the iLOM-SS in *E. coli* strains by repressing the *lpp* gene (Figure 4-1). We showed that iLOM-SS can secrete proteins up to 100kDa in detectable amounts within minutes of activation. Additionally, protein secretion with iLOM-SS can be titrated with a small molecule, and the protein secretion can be halted by relieving the repression of the *lpp* gene. Furthermore, a compact version of iLOM-SS was constructed and validated its functionality in different genetic circuit architectures. As a proof-of-concept, iLOM-SS was utilized to secrete the human cytokine IL-1Ra and the 2Rs15d nanobody.

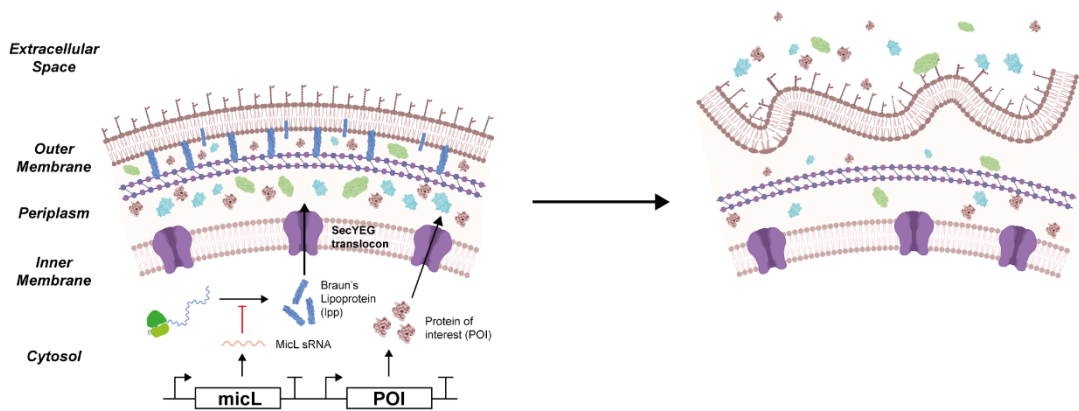


Figure 4-1. Depiction of iLOM-SS constructed in *E. coli*. In the iLOM-SS system, the protein of interest (PoI) is transported to the periplasmic space using either the TAT or Sec-dependent secretion signal peptides (SPs), such as *pelB* SP, *dsbA* SP, or *TorA* SP. A leaky outer membrane (LOM) is then induced by expressing the *micL* sRNA, which specifically inhibits the translation of Braun's lipoprotein (Lpp). This

leads to bleb formation on the outer membrane, which allows periplasmic content to diffuse out from periplasmic space to extracellular milieu.

4.3. Materials and Methods

4.3.1. Construction of Plasmids Used in the Study

All plasmids used in this study were constructed with Gibson Assembly. Plasmid sequences were verified with either Sanger sequencing or NGS. All enzymes used in this for plasmid construction were purchased from New England Biolabs (NEB). Primers were ordered from Oligomer Inc. and BiOligo Inc. If needed, the gene synthesis service of Genewiz Inc was utilized. All plasmid maps, primers, and sequencing results were given in Appendixes. LB media and LB agar with appropriate antibiotics were used for bacteria cultivation unless otherwise noted. DH5 α with Pro cassette was used for experiments explained in results secretion of 4.4.2, 4.4.3, 4.4.4, 4.4.7, 4.4.8, 4.4.9, and 4.4.10. BW25113 Δ LacI was used for experiments in 4.4.6. For all other experiments, EcN 1917 strain was used. Bacteria cells were grown in a shaking incubator with liquid media if not otherwise indicated.

To construct pET22 pLacO micL, the pET22b vector was amplified with OApETF and TR-PK-REV primers. micL coding region was amplified from *E. coli* genome by using pREA138 and pREA139 primers. The LacO coding region was added to micL region with PCR by using pREA138 and pREA140. The linearized pET vector and micL region with pLacO were assembled.

The regions encoding dsbA, bALP, MalS and ChiA proteins were cloned into pET22b pLacO micL vector. Native Tet-inducible promoter was placed upstream of the protein encoding regions. To construct the vectors, pET22b pLacO micL T7

mScarlet vector were amplified with pREA161 and pREA190; meanwhile, tet inducible promoter was amplified with pREA191 and JO22, and bALP was amplified with pREA119 and pREA192. The PCR products were assembled. The constructed plasmid was digested with HindIII and KpnI to remove the bALP coding region. DsbA coding region was amplified with Gibson overhangs by using pREA236 and pREA237. The MalS coding region was amplified with Gibson overhangs by using pREA238 and pREA239. ChiA coding region was amplified with Gibson overhangs by using pREA240 and pREA241.

To construct the toggle switch plasmid for controlling iLOM-SS, pECJ plasmid was amplified with pREA229 and pREA230 while micL region region was amplified with pREA227 and pREA228. pTetO was added upstream of micL region by PCR using pREA227 and pREA231. The linearized vector and micL with pTetO were assembled.

gRNA targeting the LacI region was cloned into CRISPR array plasmid same as described previously. pREA265 and pREA264 were used in the cloning.

To construct the compact micL plasmid, HH and HDV ribozymes were added to the micL region by PCR. pREA223 and pREA224 were used for the first amplification. The product of the first amplification was used as template for second PCR in which pREA225 and pREA226 were used. The micL/bALP plasmid was linearized by restriction digestion by MluI and HindIII enzymes. The linearized plasmid and the product of the second PCR were assembled.

For plasmids used in the memory based recombinant circuit, the inverted proD flanked with BxbI attb and attp sites was amplified with pREA262 and pREA261. The bALP coding region was amplified via using JO17 and pBS216. The pZA CmR

vector was linearized with MluI and AatII digestion. The fragments were assembled to obtain registry plasmid with bALP. For the registry containing the bALP fused with compact-micL, PCR was performed with JO17 and pREA263 to amplify bALP fused with compact-micL encoding region, and it was used instead of bALP fragment in the plasmid assembly.

For the plasmids used in the thiosulfate sensing circuit, pKD237 plasmid was amplified using pREA284 and pREA285 while pREA286 and pREA287 primers were used to amplify both bALP and bALP fused with compact-micL. The fragments were assembled to obtain the pKD287-bALP and pKD287-bALP fused with compact-micL.

For the plasmids used in the quorum sensing circuits, pLux promoter and sfGFP fused with compact micL coding regions were assembled in the pET22b vector. The vector was linearized with PCR using OApETF and TR-PK-REV primers. The pLux promoter and sfGFP were amplified by pREA280 and pREA281 whereas compact-micL was amplified with pREA282 and pREA283. The fragments were assembled. Subsequently, the LuxI with constitutive active promoter was cloned into the same plasmid. To do so, LuxI was amplified with pREA293 and pREA292 meanwhile the vector was linearized by PCR with TR-REV-PK and pREA291 primers. The fragments were assembled.

For the nanobody secretion experiments, pET placO micL/ pNTetO bALP plasmid was linearized with digestion by KpnI and HindIII, subsequently pelB tagged 2Rs15d nanobody was cloned. The 2Rs15d coding region was amplified with pREA119 and pREA242. The pelB signal peptide was changed to OmpA, DsbA, Lpp, MalE, TorA, and gIII signal peptides with the strategy that described for gRNA

cloning. pREA253 and pREA254 were used for the cloning of OmpA signal peptide. pREA244 and pREA243 were used for the cloning of DsbA signal peptide. pREA248 and pREA247 were used for the cloning of lpp signal peptide. pREA252 and pREA251 were used for the cloning of MalE signal peptide. pREA246 and pREA245 were used for the cloning of gIII signal peptide. For TorA signal peptide cloning, pREA255 and pREA256 were used to amplify the signal peptide from genome for *in vitro* assembly.

For IL-1Ra expression, IL-1Ra coding region was cloned into pET placO micL/pNTetO bALP. The vector was digested with BamHI and HindIII whereas IL-1Ra coding region was amplified with pREA294 and pREA295. The fragments were assembled.

4.3.2. Scanning Electron Microscopy

Bacteria cells bearing the plasmid that expresses micL sRNA under pLacO promoter were induced overnight in the presence of 1mM IPTG at 37 °C in LB liquid media with shaking. After the induction, cells were collected by centrifugation and washed two times with PBS. Cell pellet was resuspended in PBS at 10⁹/ml concentration. A small piece of isopropanol washed silica wafer was placed at the bottom of the 96 well plate and 10µl of PBS suspended bacteria was put on the silica wafer. The droplet was incubated for 10 minutes; 200µl of 2.5% glutaraldehyde (diluted in PBS) solution subsequently was added into well. Samples were fixed in glutaraldehyde solution on top of silica plate overnight at 4 °C. Next morning, the silica plate was washed 3 times with PBS followed by 3 times with ddH₂O and once with ethanol solutions (25%, 50%, 75%) for 5 minutes each. Lastly, silica plate was washed three times pure ethanol and dried with critical point dryer. After drying, wafer piece with

bacteria was mounted on the SEM stub and coated with Au at 8nm thickness. For visualization, Environmental SEM (Tecnica) was utilized.

4.3.3. Outer Membrane Permeability Assay with FTIC-4kDa

Cells were induced with IPTG for expression of micL sRNA, same as described in the main text. Following the induction, cells were collected and centrifuged at maximum speed. Cell supernatants were discarded; obtained pellets were washed twice with PBS to remove any residues of LB. Then, cells were resuspended in 10mg/ml FTIC-4kDa Dextran dissolved in PBS. Suspended cells were incubated at room temperature for 2 hours with shaking. After the incubation, FTIC-4kDa Dextran solution was removed via centrifugation. Cell pellets were washed two with PBS after the treatment. Washed pellets were resuspended in PBS to measure whole cell green fluorescence and to visualize under a fluorescence microscope.

4.3.4. Determination of Growth Curve and Cell Motility

To determine the growth curve, cells picked from single colonies on an agar plate were grown overnight. The next day, cells were rediluted into fresh media with 1:100 dilution factor and induced with 1mM IPTG. Optical density of cells at 600nm was measured every 20 minutes for 4 hours.

Cell motility was determined using the soft agar method. LB agar plates that contain 0.3% agar were supplemented with 1mM IPTG and proper antibiotics if needed. 10 μ l of overnight grown culture was dropped at the center of plate. Plates were incubated at 37 °C for overnight. Plate images were taken with Vilber Imaging System.

4.3.5. Western blot analysis for protein secretion with iLOM-SS

Cells carrying the plasmids for N-terminal histidine tagged recombinant versions of DsbA, ALP, MalS and ChiA was induced with 1mM IPTG to express micL sRNA and 250ng/μl aTc to express the proteins for overnight at 37°C in shaking incubator. Following the induction, cells were centrifuged at maximum speed for 5 minutes to collect cell-free supernatant. 20μl of cell supernatant for each protein was mixed with Laemmli SDS-PAGE loading dye and loaded into wells of SDS-PAGE (12%) after boiling for at minutes at 95 °C. Proteins were transferred to nitrocellulose membrane with Bio-Rad's Trans-Turbo Blot System using the Turbo-Blot RTA kit. Anti-histidine mouse IgG at 1:15000 dilution and anti-mouse IgG goat IgG conjugated with HRP at 1:10000 were utilized to obtain the western blot image. Vilber Gel Imaging System was used for imaging.

4.3.6. PnPP Assay for ALP Activity

Following the induction of cells, cell-free supernatant was collected as described above. 50μl of cell-free supernatant was mixed with 200μl of 5mM PnPP solution (PnPP dissolved in 1 mm MgCl₂, 1 mm ZnCl₂, 0.1 m glycine, pH 10.4). The solution was incubated at 37 °C for 1 hour. Absorption at 405nm was measured to assess the formed PnP. To normalize the data to cell number, absorption values obtained from the activity assay were divided by cell optical density at 600nm. The M5 spectrophotometer (Molecular Devices) was used for all measurements.

4.3.7. iLOM-SS Characterizations

To characterize the iLOM-SS in different E. coli strains, bacterial cells were transformed with the plasmid that encodes IPTG inducible micL and aTc inducible

phoA. Single colony transformants on selective agar were picked and grown overnight in the presence of appropriate antibiotics. The next day, cells were rediluted 1:100 into fresh media, grown until the OD600 of bacteria culture reaches 0.1, and induced 1mM IPTG and 250ng/μl aTc for 4 hours. After the induction, cell supernatants were collected with centrifugation at maximum speed for 5 minutes. ALP activity of cell supernatants was determined as described above.

In inducer titration experiments, EcN cells carrying the phoA/micL expression plasmid were grown the same as above. At the induction step, the concentration of one of the inducers was kept at the maximum induction concentration which was 1mM for IPTG and 250ng/μl for aTc, meanwhile, concentration of the other inducer was titrated. The titration range for IPTG was 0.125mM, 0.25mM, 0.5mM, and 1mM; on the other hand, the range for aTc was 31.25ng/μl, 62.5ng/μl, 125ng/μl, and 250ng/μl. Following the induction, ALP activity of cell supernatants was determined same as above.

For comparison of iLOM-SS with YebF and the mock control experiments, phoA/micL expression plasmid, phoA expression plasmid without micL, and only phoA expression plasmid were transformed into EcN strain separately. Cells were grown and induced with IPTG and aTc their maximum induction concentration. Following the induction, Cell supernatants were collected in each 30 minutes for 4 hours. ALP activity of collected cell supernatants was determined.

To find the time required for iLOM-SS activation, EcN cells that were transformed with phoA/micL expression plasmid was grown overnight in presence of 250ng/μl aTc. Cells were rediluted at OD600 of 0.8 into fresh media supplemented with antibiotics and both of the inducers. After the dilution, cell supernatants were

collected every 20 minutes for 100 minutes and ALP activity of supernatants was determined.

4.3.8. Reversible Activation of iLOM-SS with Toggle Switch

Previously deposited pECJ3 plasmid was used to construct the toggle switch used in the study.[228] Δ LacI was knock-out from BW25113 by using pCas stem. gRNA targeting the LacI region was designed using RGEN software tool [229] and cloned into CRISPR array. A repair DNA was constructed by assembling the left arm region of LacI, Spectinomycin resistance cassette and right arm region of LacI. For amplification of the regions, pREA266 and pREA267 were used for left arm region of LacI, pREA268 and pREA269 were used for spectinomycin resistance cassette, and pREA270 and pREA271 were used for right arm region of LacI. The BW25113 were co-transformed with pCas, LacI targeting CRISPR array plasmid and the repair DNA. Cells were induced with IPTG and arabinose during the recovery and spread on the agar plate with kanamycin and spectinomycin. The grown cells were screened colony PCR to validate the insertion of the repair into LacI region by using pREA266 and pREA296. After validation, the pCas plasmid was cured by growing cells at 42 °C overnight.

Cells were co-transformed with the toggle switch plasmid that encodes mCherry and micL sRNA, and the plasmid that encodes arabinose inducible ALP. Transformed cells were selected on LB agar plate supplemented with antibiotics. A single colony was picked and was grown overnight in M9 minimal media supplemented with 0.2% glycerol, 0.05% casamino acids, and antibiotics at 37°C in the shaking incubator. Cells in grown culture were diluted 1:10000 into fresh media supplemented with antibiotics and 1mM IPTG. Cells were induced for 24 hours at 37°C in the shaking

incubator. The next day, cells were diluted into two fresh media in different tubes with 1:10000 and 1:100 dilution factors. 1:10000 diluted cells were induced with 250ng/μl aTc in the same condition as IPTG to change the state of the cells. Meanwhile, arabinose was added into the tube that contained 1:100 diluted cells to express ALP. Arabinose induction was carried out at 37°C for 4 hours in the shaking incubator. Following the arabinose induction, the ALP activity of the cell-free supernatant was determined as described above. In addition to that, whole cell red fluorescence was measured, and cells were visualized under fluorescence microscopy. The cell state was changed six times with consecutive induction of cells with IPTG and aTc. In each state change was validated with measuring the ALP activity of cell-free supernatants, and the whole cell red fluorescence, and visualizing cell under fluorescence microscope. Zeiss fluorescence microscope and M5 spectrophotometer (Molecular Devices) were utilized.

4.3.9. Functional Assessments of Compact iLOM-SS in Different Genetic Circuits

To validate the functionality of the compact iLOM-SS, plasmids that encode pTetO controlled either PhoA with c-micL or mock control PhoA were transformed into the strain. Single colonies were picked from agar plate and induced with 250ng/μl aTc for overnight in shaking incubator at 37 °C. The next day, supernatant ALP activities of cells were determined same as described above.

For functionality assay in recombinase-based memory circuits, cells were co-transformed with the plasmid that encodes aTc inducible BxbI recombinase enzyme [230], and the plasmid that encodes either ALP with c-micL or ALP at the downstream of inverted proD promoter flanked by anti-parallelly aligned attP and

attB sites of BxbI. Single colonies were picked agar plate and grown in LB liquid media with 250ng/μl aTc overnight at 37°C in the shaking incubator to express BxbI. Subsequent to overnight induction of BxbI, cells were diluted into fresh media at the 1:10000 dilution factor. Diluted cells were grown for 24 hours in liquid LB media supplemented only with antibiotics. The next day, ALP activity of cell-free supernatant of grown cells was measured same as described above.

To test the functionality of the compact iLOM-SS in thiosulfate sensing TCS sensor, cells were co-transformed with pKD236 and modified pKD237 [231] that expresses either ALP or ALP with c-micL under control of phsA promoter. Single colonies were picked from the agar plate and induced with 5mM sodium thiosulfate in liquid LB media overnight at 37 °C. ALP activity of cell-free supernatants obtained from induced and uninduced cells was determined described as above.

For functional validation of the compact iLOM-SS in LuxI/LuxR quorum sensing circuit [232], cells carrying the plasmids that encode arabinose inducible PhoA, AHL inducible GFP fused with c-micL and LuxI with constitutively expressed promoter was grown for overnight in liquid LB media at 37 °C. The next day, cells were diluted with 1:100 dilution factor into a 1:1 mixture of fresh media and spent media that obtained from overnight culture. Arabinose is added into the media to induce ALP expression at 0.2% final concentration. The next day, ALP activity of cell-free supernatant of grown cells were measured same as described above. As controls, plasmids that either express GFP without c-micL or lack of LuxI expression cassette were used.

4.3.10. Determination of Effects of Leaked Periplasmic Content (LPC) on Caco-2 Cell Viability, Monolayer Permeability, and Cell Migration

For mammalian culture experiments, Caco-2 cells were expanded in high glucose DMEM supplemented with 10% FBS and 4mM L-Glutamine. Cells were grown in humidified incubator with 5% CO₂ at 37 °C for all experiments.

Caco-2 cells were treated with LPC containing supplemented DMEM to test safety parameters. To obtain the LPC-DMEM, EcN cells bearing the micL expression plasmid were diluted 1:100 into fresh DMEM and grown for 37 °C in shaking incubator for 3 hours in the presence of 1mM IPTG. Cells were centrifuged at full speed for 5 minutes at 4 °C to obtain the cell-free LPC-DMEM. To remove remaining cells after centrifugation, the media was filtered with 0.2µm filter. Subsequently, 10% FBS, and 4mM L-glutamine were supplemented to LPC-DMEM to obtain full growth media for Caco-2. To ensure the sterility, media was filtered again with 0.2µm filter. As a control, mock DMEM was obtained similarly by using WT EcN.

For cytotoxicity assay, 100µl of suspended Caco-2 cells were seeded into 96 well plate at cell concentration of 10⁵ per well. Seeded cells were incubated for 24 hours before the LPC treatment. 100µl of supplemented LPC-DMEM was added into well on the day of experiment. As negative controls, mock DMEM control and untreated DMEM were used. Cells were treated for 3 days with LPC. At the end of experiments, MTT assay was conducted to compare the viable cell numbers in the treatment and control groups.

For permeability testing, Caco-2 monolayers were formed onto cell culture insert (Greiner-Bio, 662640). To do so, 10⁶ Caco-2 cells that were dissolved in 250µl

growth media per insert were seeded onto apical side of inserts. Inserts were placed into 24-well plate and 750 μ l of growth media was added into basolateral side. Cells were grown for 21 days to induce spontaneous differentiation for monolayer formation. During the growth, media in both apical and basolateral sides were changed in two-day intervals. On day 21, transepithelial electrical resistance was measured using Minicell ERS-2 Volt ohmmeter (Millipore) complying with manufacturer's instructions. Monolayers that had TEER value over 250 Ω .cm² were selected for LPC treatment. Growth media at apical side was changed to mixture 1:1 of supplemented DMEM and LPC-DMEM while standard growth media was used for basolateral side. For control groups, normal growth media or DMEM obtained from WT EcN was used instead of LPC-DMEM. Cells were treated for three days with LPC. TEER values for all groups were determined in each day during the treatment. At the end of the treatment, FTIC-Dextran diffusion assay was conducted with the monolayers. Growth media at the apical and basolateral sides were collected for IL-8 ELISA. 250 μ l and 750 μ l of HBSS buffer supplement with 1% glucose was added into apical and basolateral sides, respectively. 4kDa-dextran conjugated with FTIC was added to apical side at final concentration of 10 μ g/ml. Cells were incubated for 1 hour in the incubator. Following to the incubation, 200 μ l of solution from both sides were collected; fluorescence intensity of solutions was measured with spectrophotometer. As a diffusion positive control, empty cell insert was used.

For migration assay, Caco-2 cells were seeded into 96 well plate and grown till they reached the full confluency. A small hole was created with pipette tip attached to vacuum after discarding the spent media. 100 μ l of fresh growth media was added into wells. Cells were incubated for 24 hours to recover at the hole edges. Subsequently, 100 μ l of LPC-DMEM was added into wells for treatment. The same

controls used in previous experiments were also included in the migration assay. Cells were incubated for 3 days to observe the cell migration. In each day, picture of the holes was taken with a light microscope (Zeiss). Migration was calculated based on the closure percentage of hole compared initial hole dimensions. ImageJ software was used for calculations.

4.3.11. Determination of Inflammatory Response of Treated Caco-2 Cells with LPC

IL-8 secretion was determined from LPC treated differentiated and undifferentiated Caco-2 cells. Collected growth media from the differentiated cells used for permeability was used to determine secreted IL-8. For undifferentiated cells, 100µl of Caco-2 cell suspension were seeded into 96-well plate at 10^5 cells per well seeding density and incubated for 24 hours for proper attachment. LPC treatment was conducted same as described in cytotoxicity assay. At the end of incubation, cell media was collected. IL-8 specific ELISA (R&D Systems) was conducted to determine secreted IL-8 amount following the LPC treatment.

4.3.12. Caco-2 Invasion Assay

Caco-2 cells were grown until the full confluency, same as described in the main text. On the day of the experiment, the spent growth media was discarded, and cells were washed 2 times with PBS gently. 200µl of a bacterial cell suspended in PBS whose an OD600 of 0.5 was added onto Caco-2 cells; then they were incubated at 37 °C for 4 hours. After the incubation, the bacterial cells were discarded, and Caco-2 cells were washed twice with PBS. 100µl of 0.2% Tween-20 dissolved in PBS was added onto cells to collect intracellular bacteria. Following a one hour, incubation,

solutions were collected, serially diluted, and plated on agar plates to calculate CFU for each well.

4.3.13. ELISA, Western blot analysis and Functional Assay for IL-1Ra

Cells bearing the expression cassette for IL-1Ra and micL sRNA were induced with aTc and IPTG for overnight at 37 °C. Next day, cell-free supernatant was collected same as described above. ELISA (R&D Systems) was performed to determine secreted IL-1RA in the control and sample supernatants without any enrichment method e.g., ultrafiltration. 50x fold concentrated supernatants were used in western blot analysis and activity assays. Western Blot was performed same as described above except the IL-1Ra specific antibody was used instead of the anti-his tag antibody.

To validate the activity of secreted IL-1Ra, cell-free supernatant that contains secreted IL-1Ra was obtained same as described above. The supernatant was further concentrated by 50x fold with ultrafiltration. The desalting column was used to exchange the buffer of supernatant to PBS. The obtained solution was used in activity assays. A375.S2 cell line was used to determine the activity. 2000 cells per well were seeded into 96-well plate. Cells were incubated for a day prior to the treatment. On the day of treatment, 5ng/ml IL-1 α was added to each well. Concentrated supernatant was added into wells in different amounts ranging from 0.1 μ l to 10 μ l. Following the three-day incubation, MTT assay was conducted to measure cell viability. Pure IL-1Ra, the supernatant obtained from micL sRNA expressing EcN and PBS were included in the assay as controls.

4.3.14. Nanobody Secretion with iLOM-SS

Cells were transformed with plasmids separately encoding the micL under control of pLacO and 2Rs15d nanobody [233] with different signal peptides at the downstream of pNtetO. Cells were induced overnight in presence of aTc (250ng/μl) and IPTG (1mM). Following the induction, cell free supernatants were collected via centrifugation. The supernatants were analyzed with western blot to assess the secretion efficiency of signal peptides.

4.3.15. Statistical Analysis

GraphPad Prism was used to visualize data and calculate p-values. $0.05 < p\text{-value}$ was taken as significant for all data obtained in the study. To compare values, appropriate statistical tests were conducted such as Student's t-test, one-way ANOVA, and two-way ANOVA. Each measurement was taken independently. At least, three replicates were used for all experiments. Symbols of *, **, ***, and **** were used for p values that were smaller than 0.05, 0.01, 0.001 and 0.0001, respectively.

4.4. Results

4.4.1. Cloning of micL into IPTG Inducible pET22b

IPTG inducible micL sRNA expression module was constructed in the pET22b vector. To clone, pET22b vector was amplified with T7 terminator and without T7 promoter. LacI expression cassette was not destroyed. micL sRNA was amplified with primers and micL sRNA was added to pLacO promoter with two successive

PCRs. (Figure 4-2a, 4-2b, 4-2c) Vector and expression cassette were assembled and transformed.

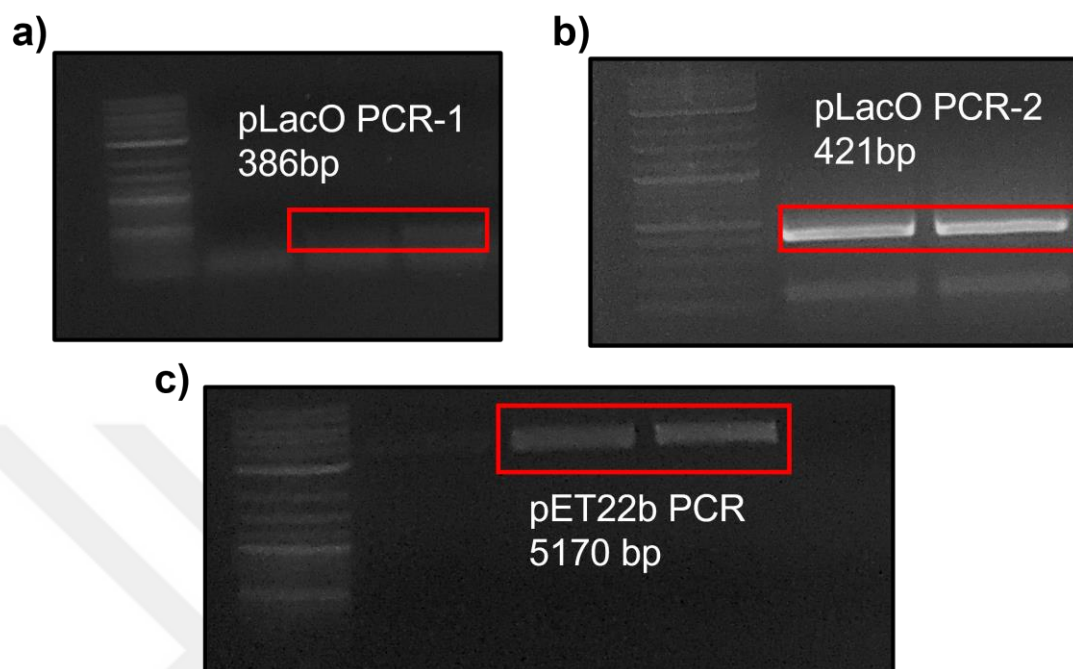


Figure 4-2. Construction of IPTG inducible micL expression plasmid. a) Addition of micL at the downstream of pLacO via PCR. b) Amplification of pLacO-micL for assembly. c) Linearization of pET22b with PCR for assembly. All DNA bands appeared at their expected sizes.

4.4.2. ESEM Analysis of micL Expressing Cells

Cells carrying the micL expressing the plasmid were induced to check whether micL expression is sufficient to induce leaky outer membrane (LOM) phenotype. Following the induction, cells were fixed, dried and spotted on silica. The gold coated cells were visualized under ESEM. The LOM phenotype was determined based on blebbing of the membrane which was shown to be an indicator of the phenotype. Blebs were observed on the cells expressing the micL sRNA, indicating that the outer membrane became leaky.

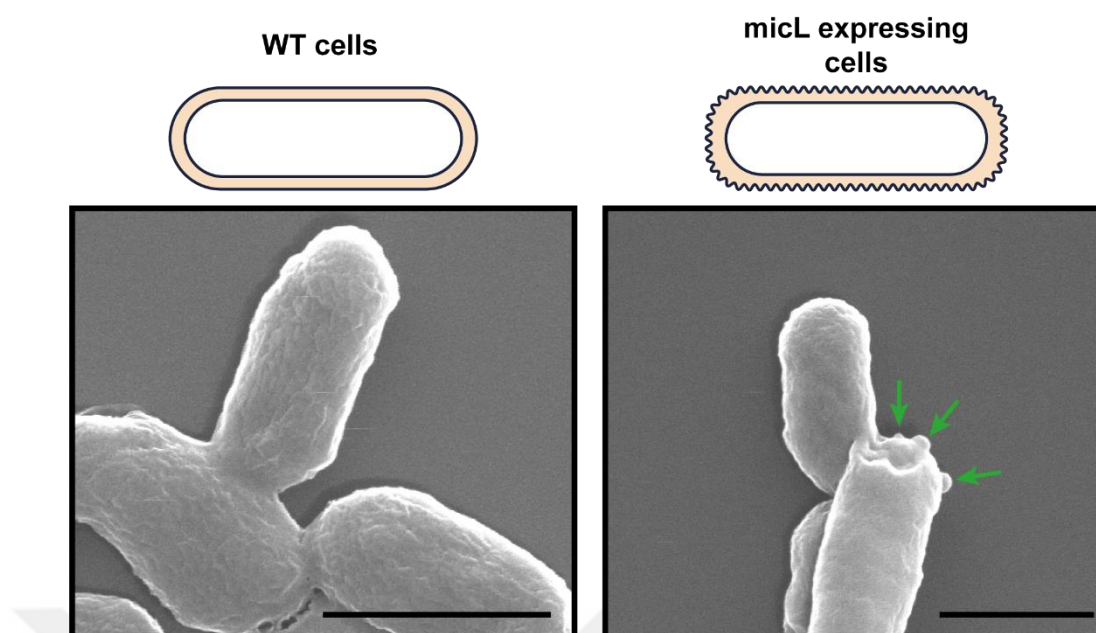


Figure 4-3. ESEM images of WT and micL expressing *E. coli* cells. Blebings on outer membrane were pointed by green arrows. Black bars are 1 μm .

4.4.3. FTIC-4kDa Dextran Diffusion Assay with micL Expressing Cells

LOM phenotype induction with micL expression was further confirmed with free diffusion of FTIC conjugated 4 kDa dextran molecules into periplasm. Since the free diffusion of macromolecules is not possible through pores at the outer membrane, only cells with LOM phenotype can take up macromolecules from the extracellular milieu. (Figure 4-4a) The induced and uninduced cells were incubated with FTIC conjugated 4 kDa dextran for hour. Following the incubation, cells were visualized under fluorescence microscope, at the same time, whole cell fluorescence was also measured. Both microscopy images and fluorescence measurements indicated that micL expressing were able to retain the macromolecules, showing that micL expression induced the LOM phenotype. (Figure 4-4b, and Figure 4-4c)

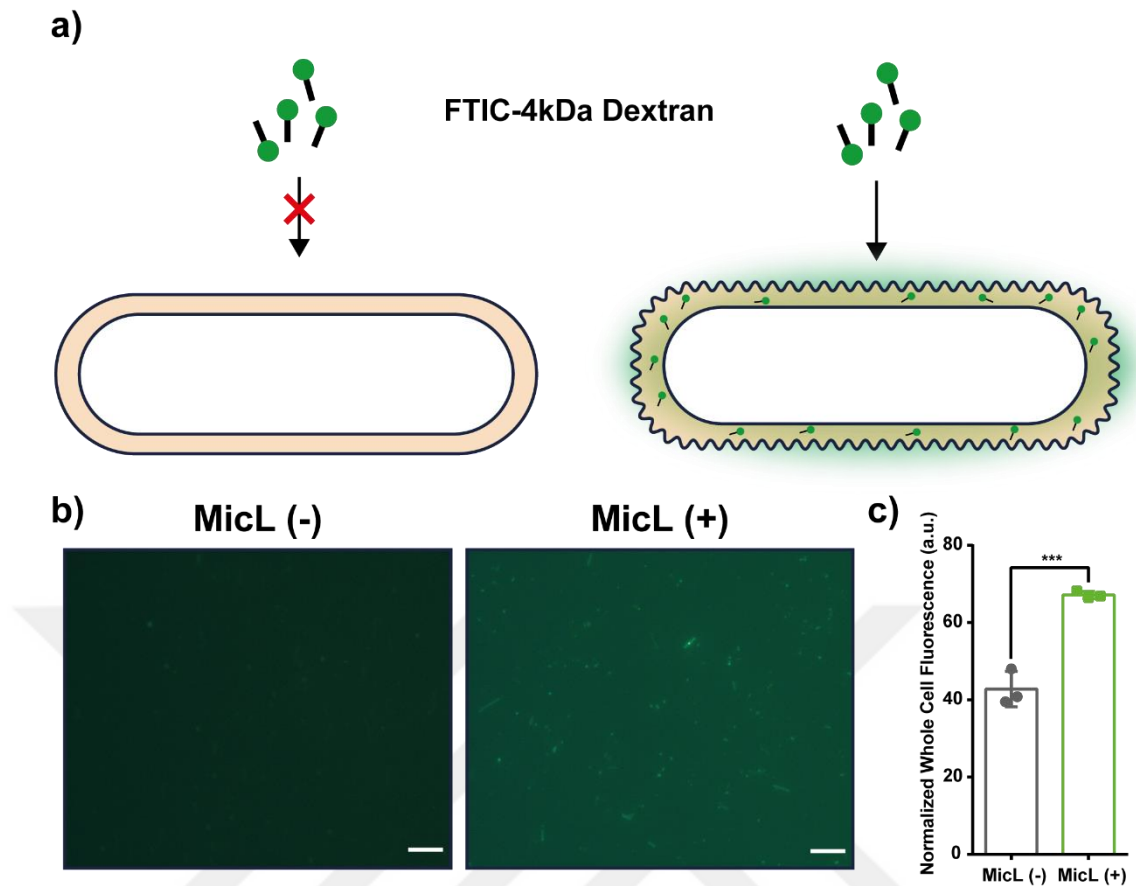


Figure 4-4. The iLOM system makes the outer membrane of bacteria leaky, allowing macromolecules to diffuse from the extracellular medium into the periplasm. a) The schematic of the assay performed to validate the diffusion of macromolecules from the extracellular space to the periplasm. b) Fluorescence microscope images of cells treated with 4kDa Dextran conjugated with FTIC. The image on the left shows EcN cells without expression of the micL sRNA, and the image on the right shows EcN cells expressing the micL sRNA. c) The whole-cell green fluorescence of cells with and without micL sRNA expression after the diffusion assay.

4.4.4. Secretion of DsbA, bALP, MalS and ChiA with iLOM-SS

Protein release from periplasmic space was tested with micL expressing cells. Four different proteins with different molecular weights were selected to determine the size limit of free macromolecule diffusion from periplasmic space to extracellular milieu with LOM phenotype. Firstly, pET22b pLacO micL / pNtetO bALP was constructed. pET22b micL vector, bALP gene and pNtetO gene were amplified with compatible overhangs via PCR; then, fragments were assembled. (Figure 4-5) bALP was tagged with six histidine.

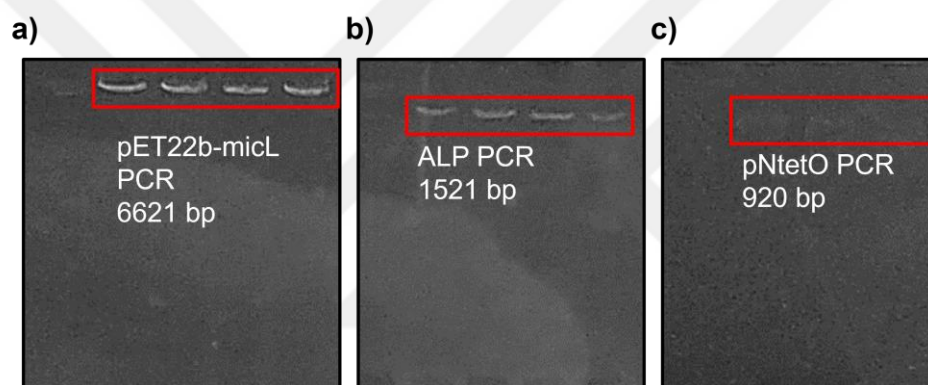


Figure 4-5. Construction of pLacO micL/ pNTetO bALP plasmid. a) PCR amplification of pET22b pLacO micL. b) PCR results of bALP encoding region from *E. coli* genome. c) Amplification of pNTetO. All DNA bands appeared at their expected sizes.

To clone the *dsbA*, *malS* and *chiA* genes, pET22b pLacO micL / pNtetO bALP was digested to remove the bALP gene whereas protein encoding genes were amplified with six histidine tag and Gibson overhangs via PCR. (Figure 5-6) Gibson assembly was utilized to insert genes into the linearized vector, individually.

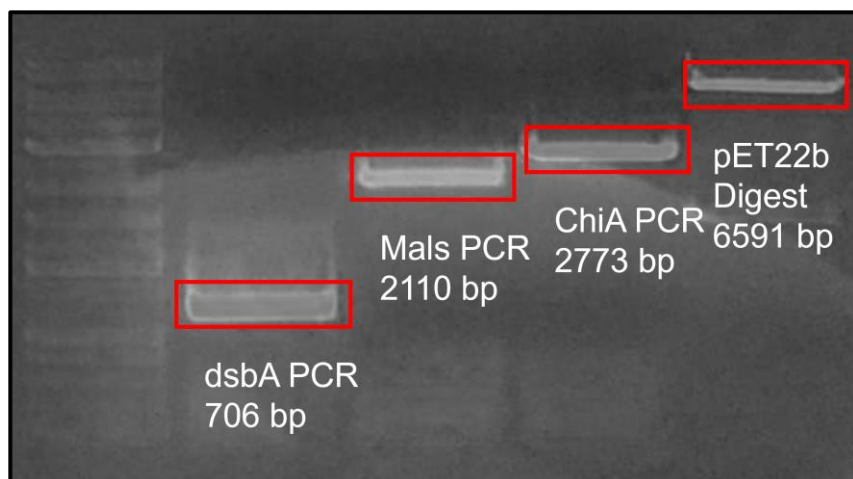


Figure 4-6. Cloning of his-tagged DsbA, MalS, and ChiA proteins at the downstream of pNtetO of pLacO micL/ pNTetO bALP plasmid. PCR amplification of dsbA (second lane), MalS (third lane), and ChiA (fourth lane) encoding regions with 6x his-tag. pET22b was digested for cloning. (fifth lane) All DNA bands appeared at their expected sizes.

After the cloning of vectors, sequence verified vectors were transformed into *E. coli* cells separately, and cells were induced to expression of micL and proteins for overnight. The cell free supernatants were collected and analyzed with western blot. The western blot analysis showed that all proteins expressed in the periplasm were released to extracellular milieu following the induction micL. (Figure 4-7)

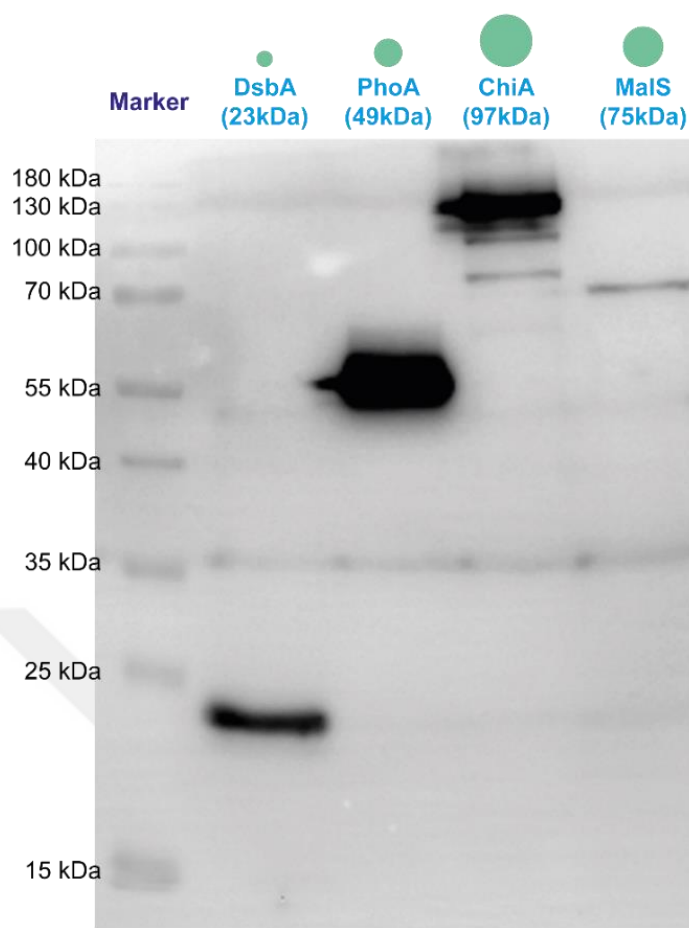


Figure 4-7. Western blot analysis of cell-free supernatant obtained from culture expressing the *micL* sRNA and the proteins.

4.4.5. Characterization of iLOM-SS Using *PhoA* as Reporter

The iLOM-SS was characterized by secreting bALP enzyme to extracellular milieu. A mock control and a secreting control were constructed. To do so, bALP was cloned in the vector that expresses a mock RNA under control of pLacO instead of *micL*. The pNTetO-bALP region was amplified with PCR and inserted in the mock RNA expressing vector with Gibson assembly. (Figure 4-8a) Secreting control was constructed by cloning the secretion tag of YebF to bALP in the mock RNA expressing vector. The YebF coding region was obtained with overhangs with PCR and cloned with Gibson assembly. (Figure 4-8b)

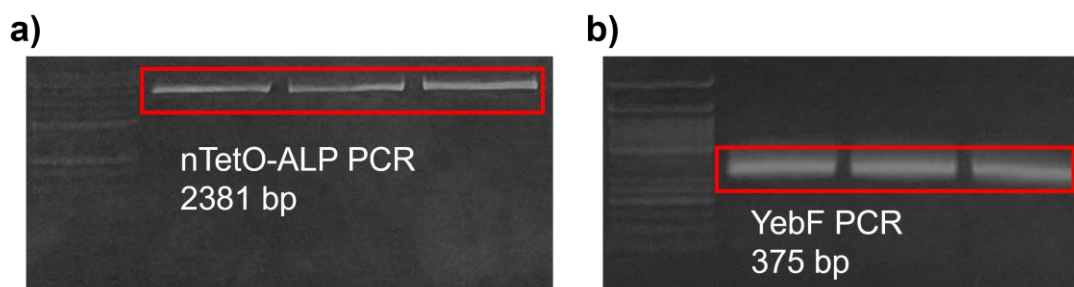


Figure 4-8. Cloning of control plasmids, pET22b pLacO mock RNA / pNTetO bALP and pET22b pLacO mock RNA / pNTetO yebF-bALP. a) PCR amplification of pNtetO bALP regions for cloning into pET22b pLacO mock RNA plasmid. b) Amplification of yebf region from *E. coli* genome to clone at the upstream of bALP. All DNA bands appeared at their expected sizes.

bALP was utilized as reporter for characterization of iLOM-SS because it can degrade the para-nitrophenol phosphate (PnPP) substrate to yellow colored para-nitrophenol (PnP) which can easily quantified by determining solution absorption at 405nm. (Figure 4-9a)

Firstly, the functionality of iLOM-SS was evaluated across various *E. coli* strains. To do so, the reporter plasmid harboring an aTc-inducible *phoA* gene and an IPTG-inducible *micL* sRNA was introduced into various *E. coli* strains, including the wild-type K-12 strain, the BL21(DE3) strain, and the probiotic EcN strain. Following overnight growth, cells were induced to express both ALP and *micL* sRNA, and the ALP activity in the cell supernatants was subsequently measured. Notably, *micL* expression resulted in significantly elevated ALP activity compared to uninduced cells across all tested strains. (Figure 4-9) These findings clearly demonstrate the functionality of the iLOM-SS in diverse *E. coli* strains. Due to its probiotic properties, the EcN strain was selected for subsequent characterization experiments.

Inducer titration experiment was carried out to determine whether iLOM-SS provide titratable protein secretion in EcN. Bacterial cells harboring the reporter plasmid were induced with varying concentrations of either aTc or IPTG, while the concentration of the other inducer was maintained at its maximum (250 ng/ml for aTc and 1 mM for IPTG). After a 4-hour induction period, cell supernatants were collected for ALP activity determination. The results revealed that titration of aTc, and consequently *phoA* expression, enabled control over the amount of secreted ALP. However, titration of *micL* sRNA expression using IPTG did not provide titratable protein secretion within the tested concentration range. (Figure 4-9c)

The secretory capacity of iLOM-SS was benchmarked against the widely used YebF[234] fusion tag in *E. coli*. The *phoA* gene was inserted into the C-terminus of YebF, and cells expressing only *phoA* were included as a mock control. Overnight cultures of plasmid-carrying bacterial cells were diluted into fresh media and induced upon reaching an OD600 of 0.1. Cell supernatants were collected at 30-minute intervals for 4 hours to determine ALP activity. The results demonstrated that induced iLOM-SS could release a detectable amount of ALP enzyme into the extracellular media within 2 hours, with continuous protein accumulation in the extracellular milieu throughout the experiment. In contrast, the YebF fusion tag failed to translocate the ALP enzyme, similar to mock control. (Figure 4-9d)

Although a detectable amount of ALP enzyme was released at least 2 hours post-induction, it was hypothesized that the bottleneck might be the ALP production rather than iLOM-SS activation. To test this, bacterial cells were induced overnight to express ALP. The following day, cells were diluted into fresh media at an OD600 of 0.8, and *micL* sRNA expression was induced to activate iLOM-SS. Cell

supernatants were collected every 20 minutes for 100 minutes, and extracellular ALP levels were assessed using an activity assay. A significant difference in accumulated ALP activity was observed between micL-expressing cells and the mock control 60 minutes post-induction, indicating that iLOM-SS activation occurs approximately one hour after micL sRNA expression. (Figure 4-9e)

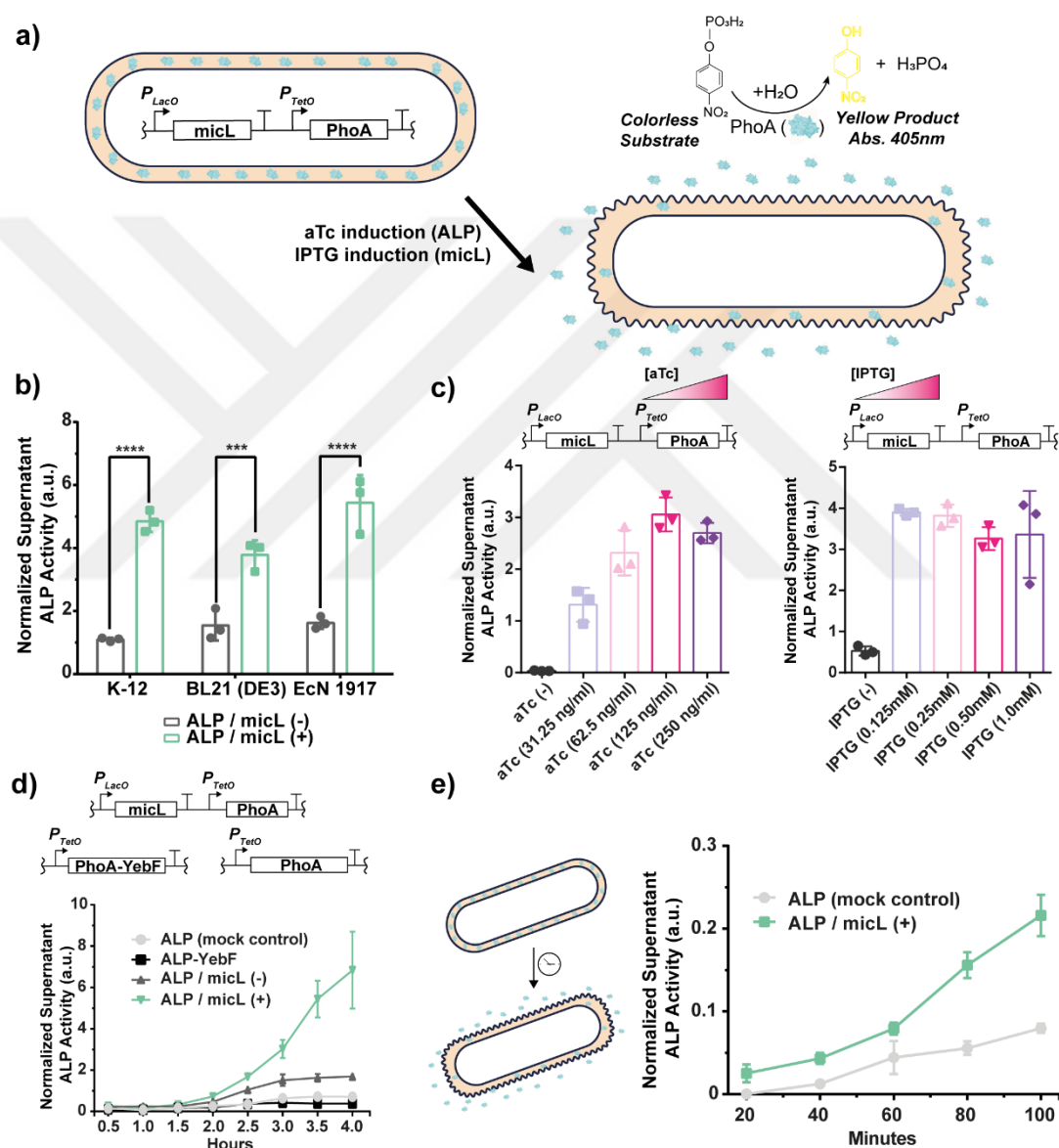


Figure 4-9. Characterization of iLOM-based secretion system in EcN. a) Schematic presentation of enzyme model used to detect secreted proteins amount in the extracellular milieu. PhoA encodes an alkaline phosphatase enzyme which

degrades the colorless PnPP to yellow-colored nPP. Enzymatic activity of ALP was used to assess and compare secreted protein amounts in supernatants upon induction of micL and PhoA. b) Validation of iLOM-based secretion system in different strains of *E. coli*. The plasmid that encodes micL and PhoA was transformed into *E. coli* K-12, BL21 (DE3) and EcN strains. Following induction with aTc and IPTG, ALP activity of supernatants obtained from cells was measured with PnPP degradation assay. The student's t-test was used to assess the significance. c) Expression titration of micL with IPTG and PhoA with aTc to evaluate the controllability of iLOM-based secretion in terms of secreted protein amount. During titration experiments, the non-titrated inducer was added at maximum concentration which was 1mM for IPTG and 250ng/ml for aTc. Four different concentrations were determined for each inducer. d) Comparison iLOM-based secretion of ALP with YebF carrier protein-based secretion and non-secreting control in time domain. Supernatants of cells were collected in each 30 minutes for 4 hours to determine secreted ALP activity in the extracellular space. e) Determination of time required for activation of the iLOM-SS after the induction of micL. PhoA was induced beforehand to translocate ALP enzymes to the cell periplasm. Then, cells were diluted 100-fold into fresh media supplemented with both aTc and IPTG. Subsequently, supernatants were collected for 100 minutes with 20 minutes intervals. ALP activity of obtained supernatants was determined with PnPP assay. All experiments were conducted with three independent biological replicates.

4.4.6. Conditional and Reversible Secretion by iLOM-SS

The reversibility of LOM phenotype induction by micL expression was tested. The main hypothesis was that the leaky outer membrane (LOM) phenotype could be

reversed, and secretion halted by rescuing lpp translation in the absence of micL sRNA expression. If the hypothesis is true, stop genetic switches can be constructed to halt protein secretion with the inducible Leaky Outer Membrane based Secretion System (iLOM-SS) in case of need.

To test this, a toggle switch for micL expression was constructed. The pECJ-3 plasmid was amplified to clone the micL region at the downstream of pTetO. (Figure 4-10a). Coding region for micL was extended with PCR to add compatible overhangs for Gibson assembly. (Figure 4-10b) The obtained fragments were assembled, and sequence verified.

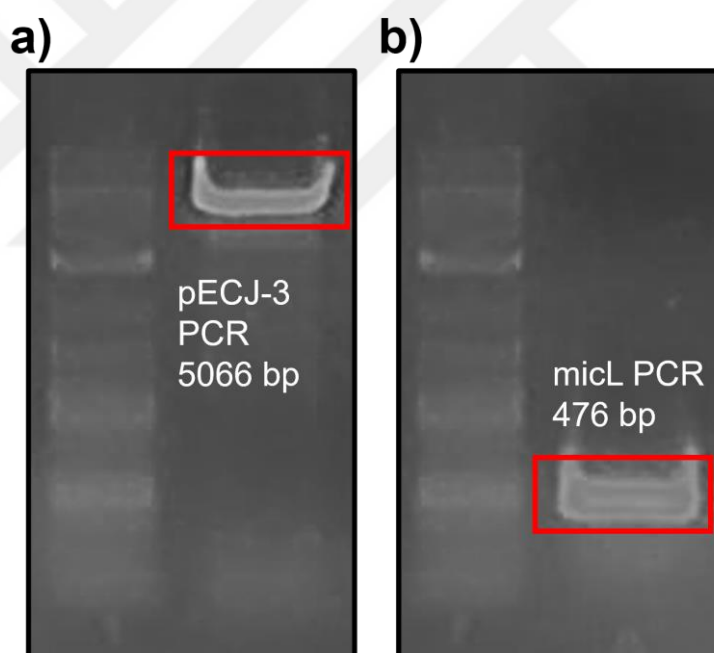


Figure 4-10. Construction of pECJ-micL plasmid that encodes TetR/LacI based toggle switch. a) PCR amplification of pECJ plasmid with overhangs for assembly. b) PCR results of micL amplification for Gibson assembly. All DNA bands appeared at their expected sizes.

Prior to secretion and switching experiments, LacI expression from genome was eliminated by inserting the SpecR coding region to ORF of lacI gene. Cas9 augmented recombineering was utilized for the insertion. In the augmented recombineering, the pCas plasmid encodes the Cas9, tracrRNA and recombinase enzymes whose functions are necessary to obtain viable *E. coli* cells following the Cas9 activity because *E. coli* cells lack of any non-homologous repair mechanisms. gRNA loaded Cas9 are targeted to LacI gene for inducing double strand break which are repaired with homologous recombination by action of recombinases and the donor DNA. The donor DNA contains the homologous regions for recombinant and antibiotic resistance cassette for selection the edited cells in the presence of antibiotic. (Figure 4-11a). Firstly, gRNA was designed and cloned into CRISPR array plasmid using the same protocol described in Chapter 3. (Figure 4-11b). The donor DNA was assembled by PCR in which SpecR cassette was flanked by homolog regions of LacI gene. The pCas plasmid, the CRISPR array plasmid and the donor DNA were electroporated into cells, and recovered cells were spread on the agar supplemented with spectinomycin antibiotic. The grown cells were screened with colony PCR by amplifying SpecR and right arm homolog region. Colony PCR revealed that SpecR was inserted into LacI region. (Figure 4-11d). A positive colony was picked for toggle switch experiments.

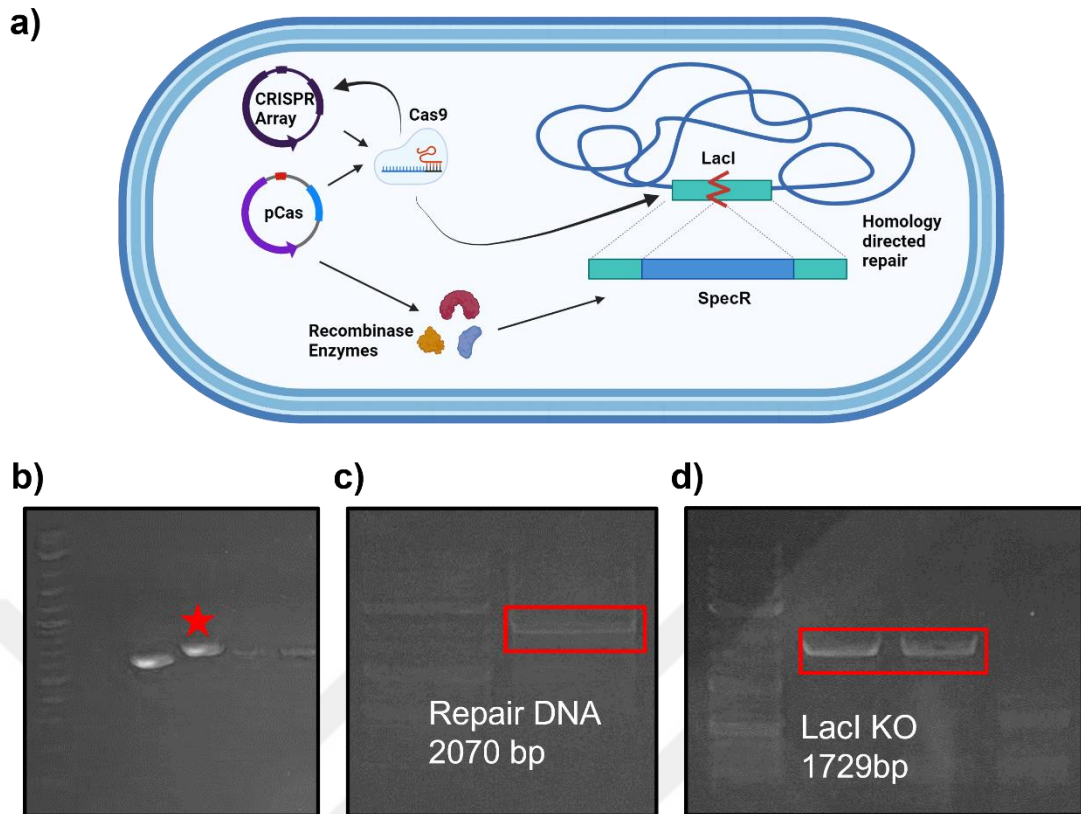


Figure 4-11. Knocking out of LacI gene from *E. coli* genome. a) The illustration of Cas9 based gene knock-out in *E. coli* cells. a) Colony PCR results to validate the cloning of LacI targeting gRNA into E1 CRISPR array plasmids. b) Assembly of donor DNA for repairing the double stranded break caused by Cas9 at the LacI locus. The donor DNA encodes a spectinomycin resistance cassette for selection. c) Colony PCR results showing the insertion of SpecR cassette into LacI region. All DNA bands were observed at their expected sizes.

The constructed toggle genetic switch expressed micL sRNA with pTetO and mCherry with pLacO. This created two states: a micL sRNA-expressing secretory state and an mCherry-expressing non-secretory state. A plasmid that encodes arabinose-inducible bALP as the secretion reporter was co-transformed with the toggle switch plasmid into cells. (Figure 4-12a)

The cells were induced sequentially with aTc and IPTG for six days to change their states between secretory and non-secretory. After each induction, the cell state changes were confirmed by measuring the mCherry signal from the cells. As expected, cells expressed mCherry following IPTG induction, which changes the cell state to the non-secretory state. The mCherry signal decreased when cells were induced with aTc, which changes the cell state to the secretory state. (Figure 4-12b, and 4-12c) To further validate the system, the cells in each state were induced with arabinose to check whether ALP could be released from the cells. Cells in the secretory state could release ALP to the extracellular medium, while no ALP activity was observed in the supernatant of cells in the non-secretory state. (Figure 4-12b) These results demonstrate that the hypothesis was correct and that an emergency genetic circuit could be constructed to stop protein release from iLOM-SS by blocking the expression of micL sRNA.

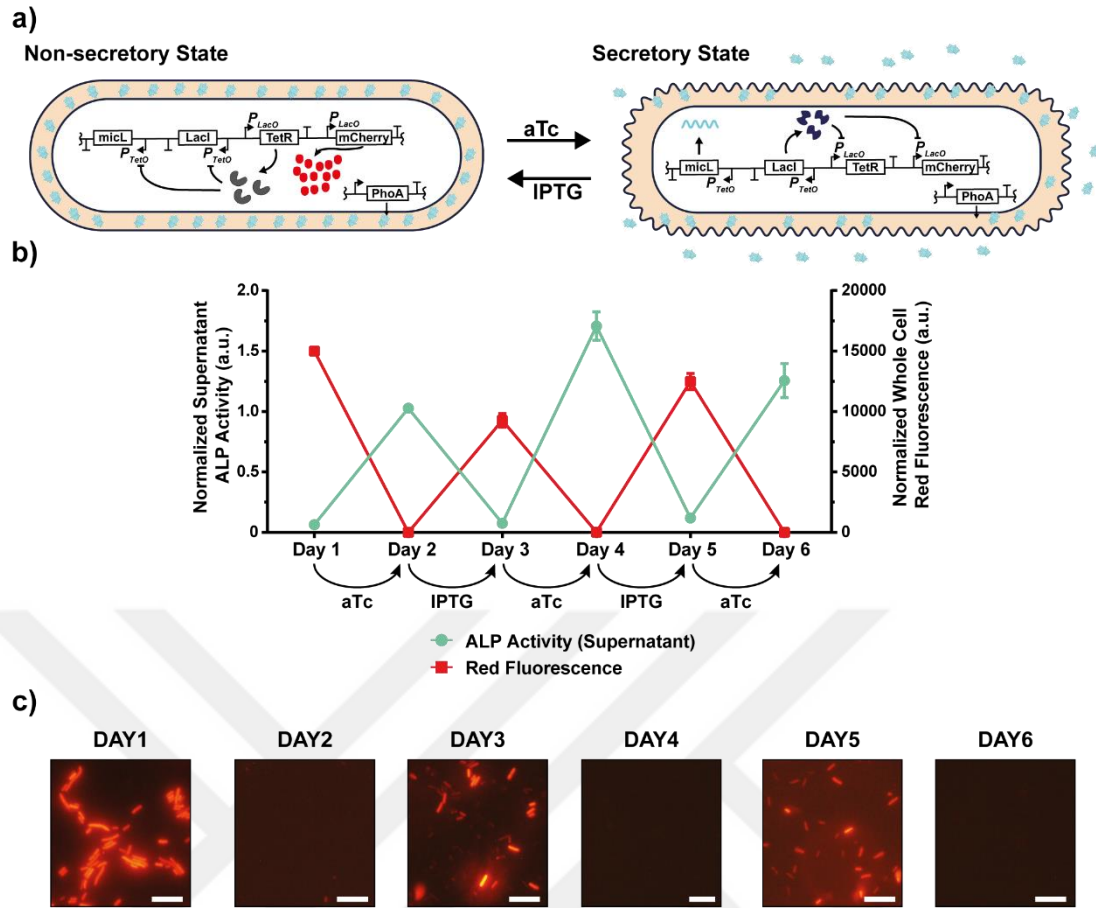


Figure 4-12. Leaky membrane can be restored in the absence of micL expression to halt the secretion of periplasmically translocated proteins with the iLOM-SS. a) Schematic representation of genetic circuit that was used to test the hypothesis of leaky membrane restoration. A toggle switch gene circuit with LacI/TetR repressor pair was constructed. The coding region of micL sRNA was placed downstream of pTetO promoter whereas mCherry expression was controlled by pLacO. In the non-secretory state, cells were expected to produce mCherry and TetR proteins; on the other hand, micL sRNA and LacI were expressed by the cells in the secretory state. b) Supernatant ALP activity and whole cell red fluorescence of the cells sequentially induced by IPTG and aTc for six days. c) Red fluorescence microscopy images of the cells each day. White bars are 5 μ m. All experiments were conducted with three biological replicates.

4.4.7. Construction of Compact-micL

A compact version of iLOM-SS was engineered to reduce the metabolic burden on cells, wherein the expression of the Point of Interest (POI) and micL sRNA occurred on the same transcriptional unit, eliminating the need for a second expression cassette. The expression cassette was constructed in the pET22b mock RNA /pNtetO bALP vector. The vector was digested with HindIII for linearization. (Figure 4-13a) The micL RNA was extended and amplified to add the ribozymes sites with two PCR. (Figure 4-13b, Figure 4-13c). The plasmid and the micL DNA fragment were assembled with Gibson Assembly.

During the cloning, the micL sRNA is placed in between hammerhead ribozyme (ribozyme I) end and hepatitis delta virus (HDV) ribozyme (ribozyme II). The constructed DNA region is named compact micL (c-micL) and directly fused to 3' end of PoI coding mRNA. After the transcription, ribozyme I and ribozyme I catalyze the self-processing of the c-micL transcript into three RNA molecules: the PoI coding mRNA fused with ribozyme I, the micL sRNA, and ribozyme II. The micL sRNA then induces the outer membrane leakiness, allowing the PoI to diffuse into the extracellular milieu. (Figure 4-14a)

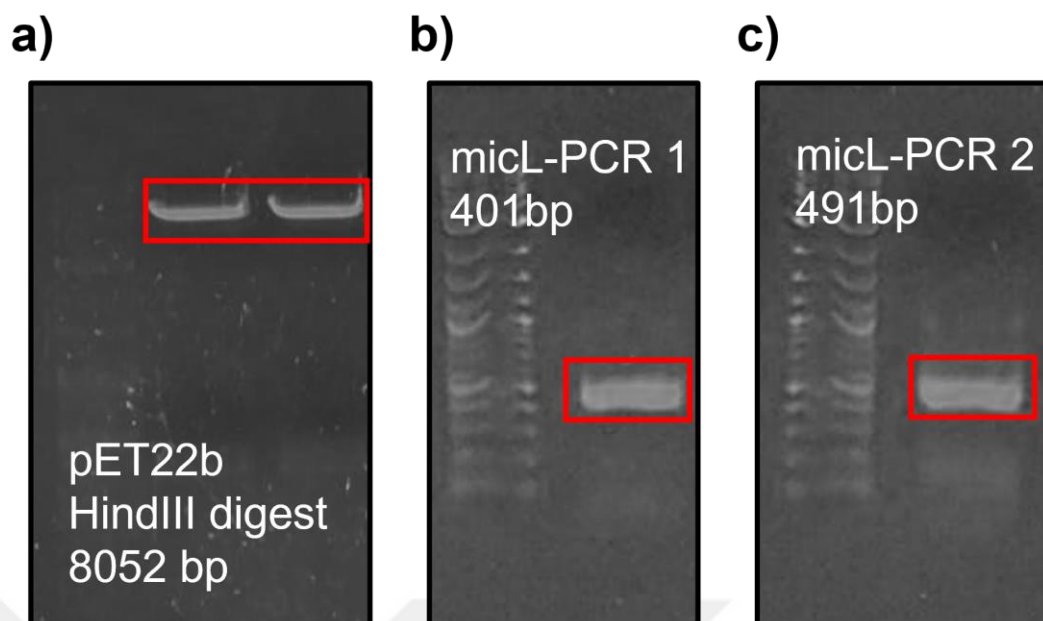


Figure 4-13. Construction of pET22b c-micL bALP plasmid. a) Digestion of pET22b plasmid with HindIII b) Extension of micL sRNA for addition of ribozymes via PCR c) The amplification of extended micL sRNA with PCR for cloning. All DNA bands appeared at their expected sizes.

BALP was employed as the reporter enzyme to test whether compact micL provide induction of LOM based protein secretion. Cells bearing either bALP or bALP fused with c-micL were induced overnight then, the phosphatase activity of cell free supernatant from cultures were determined. Analysis of supernatant bALP activities revealed that cells expressing the fusion of bALP with c-micL were able to release the bALP enzyme into the extracellular medium at levels significantly higher than those observed in the mock control. (Figure 4-14b)

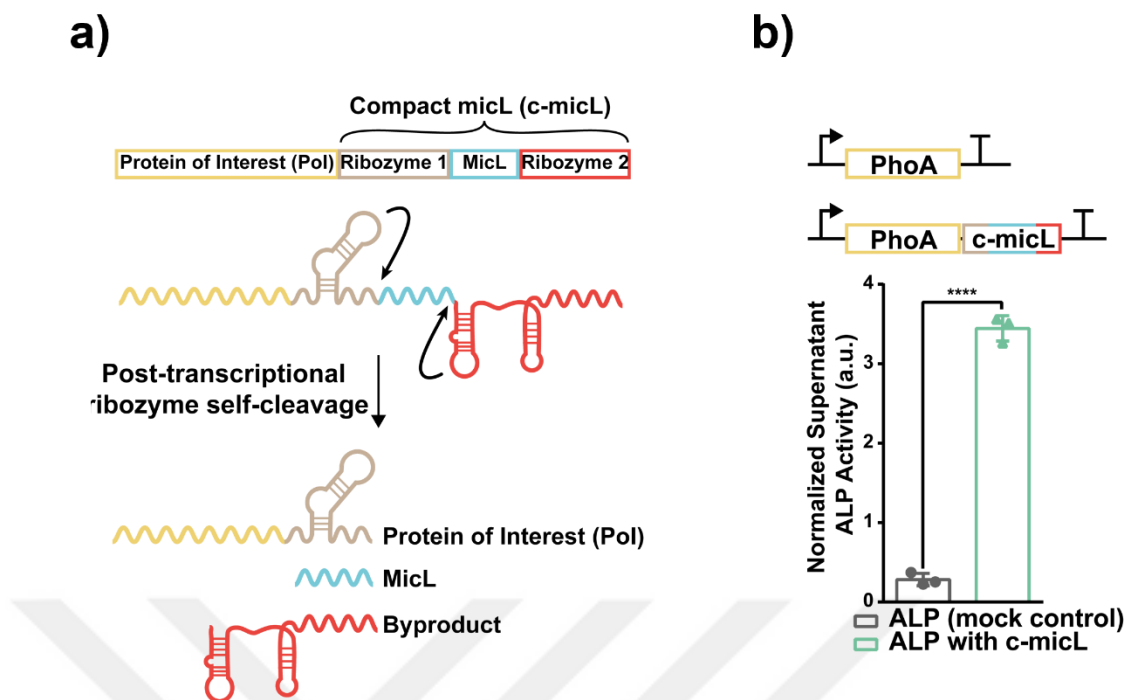


Figure 4-14. Development of single transcription unit for iLOM-based secretion.

a) Visual depiction of single transcription unit topology for the iLOM-SS. b) Functional validation of iLOM-based secretion of ALP with c-micL. Supernatant activities of cells that were carrying either only PhoA gene (mock control) or PhoA fused with c-micL were measured. All experiments were carried out with three biological replicates.

4.4.8. Functionality Assessment of Compact-micL in Thiosulfate Sensing Circuit

Operability of c-micL in two-component system-based biosensors were determined. Thiosulfate sensing ThsS/ThsR circuit was selected as the model which previously characterized in EcN. [231] For testing, bALP and bALP fused with c-micL was cloned in the pKD237 vector that encodes the TCS promoter. The pKD237 vector was amplified with PCR to linearize the DNA whereas bALP-cmicL and bALP was obtained with Gibson overhangs via PCR amplification. (Figure 4-15a, and 4-15b)

The bALP and bALP-cmicL were cloned into pKD237 vector with Gibson Assembly.

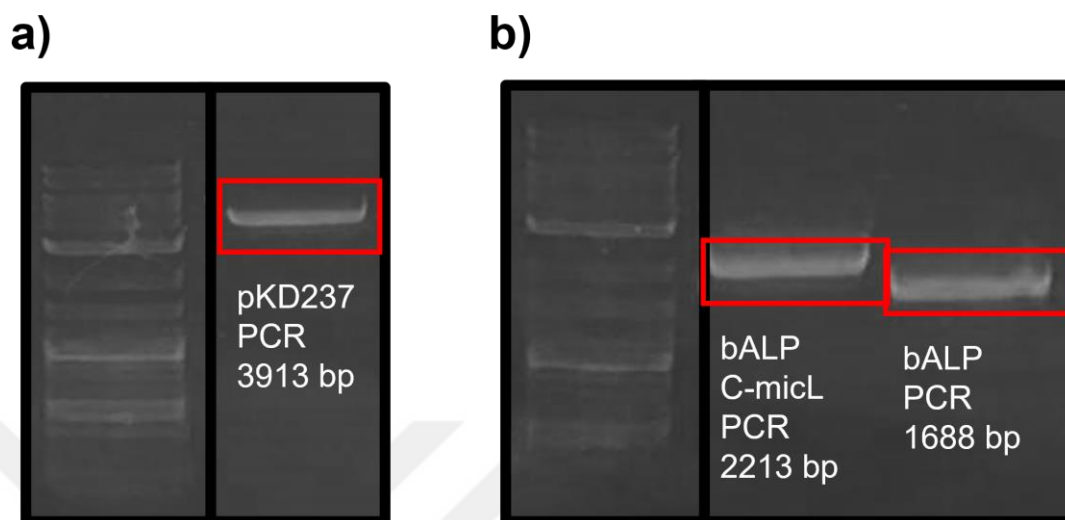


Figure 4-15. Construction of pKD327 bALP and bALP-cmicL plasmids. a) Amplification of pKD237 with compatible overhangs for assembly. b) Extension of bALP and bALP-micL to add gibsion overhangs via PCR. All DNA bands appeared at their expected sizes.

In the thiosulfate sensing TCS, ThsS is the transmembrane protein that binds the thiosulfate molecules outside of cells. Upon binding of thiosulfate, ThsS is activated and phosphorylates the intracellular ThsR transcription. Then, phosphorylated ThsR binds its cognate promoter p(phsa) and starts to transcription of downstream genes. (Figure 4-16a) The cells bearing full TCS circuit with either bALP or bALP-cmicL were induced with thiosulfate for overnight. The next day, supernatants were collected from cultures and phosphatase activity of supernatant were determined. Based on phosphatase assay, bALP activity of cells bearing the bALP-cmicL constructs was significantly higher than other groups thereby indicating that cmicL was functional in TCS circuit. (Figure 4-16a)

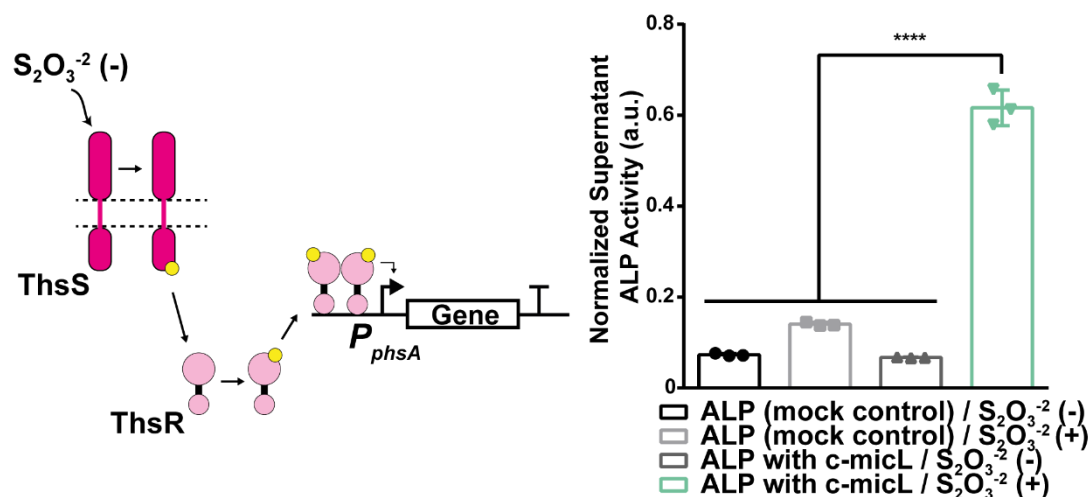


Figure 4-16. Functional validation of c-micL in thiosulfate sensing TCS based biosensors. To sense thiosulfate, ThsS/ThsR two component system was used. bALP secretion was utilized as output in the genetic circuit. As negative control, bALP was used. The phosphatase activity of cell-free supernatant was measured with PnPP assay. The experiment was performed with three replicates.

4.4.9. Functionality Assessment of Compact-micL in Recombinase Based Memory Circuits

Performance of cmicL was determined in the recombinase-based recording circuit. To test the performance of cmicL, registry plasmids were constructed that harbor inverted proD flanked by BxbI recognition sites and either bALP or bALP-micL at the downstream of inverted proD. For cloning, the pZa plasmid was digested with restriction enzymes to linearize the DNA. (Figure 4-17a) The inverted proD region, bALP coding region and bALP-cmicL region were amplified via PCR with compatible overhangs. (Figure 4-17b, 4-17c and 4-17d) The bALP and bALP-cmicL regions were cloned separately into the pZa vector along with inverted proD via Gibson assembly.

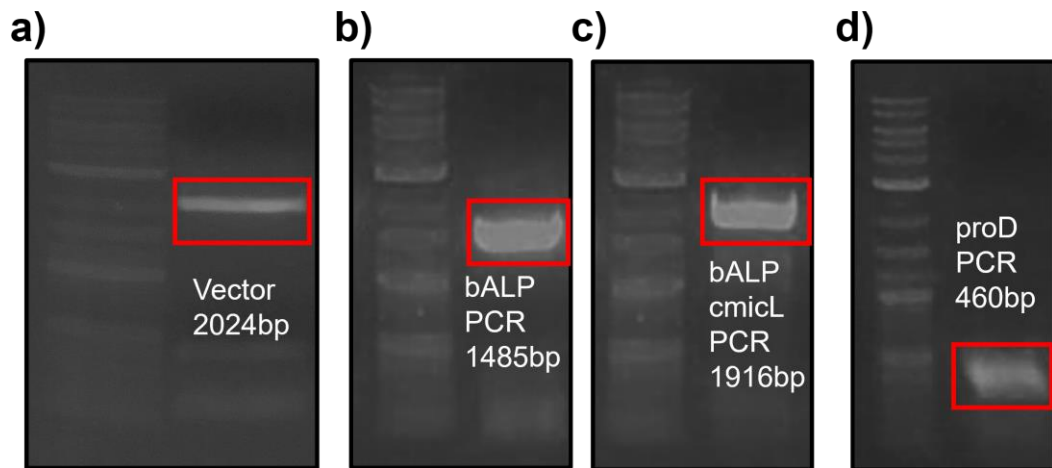


Figure 4-17. Construction of registry plasmids with bALP and bALP-micL. a, b, c, d) Amplification of pZA vector (a), bALP (b), bALP-cmicL (c), and inverted proD flanked by attP and attB sites of bxbI (d). bALP and bALP-cmicL was cloned into pZA vector with inverted proD, separately. All DNA bands appeared at their expected sizes.

For permanent DNA recording, the BxbI recombinase is expressed under aTc regulated promoter. In presence of aTc, the expressed BxbI binds its attP and attB sites on the registry plasmid and catalyzes the inversion of sequence inbetween. Thus, the aTc presence is recorded in the registry plasmid irreversibly. (Figure 4-18a) To test the system, cells bearing the circuit plasmids were induced with aTc for overnight. Next day, cells were rediluted by 10.000-fold into fresh media without aTc and grown again for overnight. Following the growth of culture, supernatants were collected for measuring the amount of secreted bALP by phosphatase assay. The activity results highlighted that cells bearing recombinase circuit with bALP-cmicL were able to record presence of aTc and secreted bALP as the output even there was no aTc remaining. (Figure 4-18b) Thus, the experiments showed that cmicL was operatable with recording genetic circuits based on recombinases.

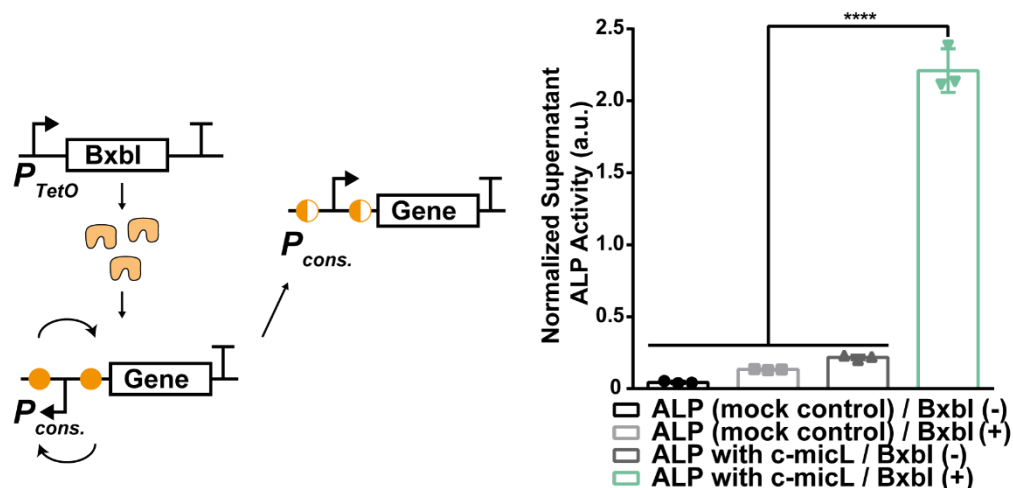


Figure 4-18. Functional validation of c-micL in recombinase-based memory recoding circuit. Inversion of proD was used to record presence of the inducer. In the presence of the inducer, BxbI was expressed subsequently direction of proD would be reverted. As the output, bALP secretion was measured from cell-free supernatant in which bALP without cmicL was used as a control.

4.4.10. Functionality Assessment of Compact-micL in Quorum Sensing Circuit

Functioning of cmicL in quorum sensing circuit was validated. The quorum sensing of *Vibrio fischeri* was constructed in the pET22b vector. The vector was amplified via PCR with Gibson overhangs. (Figure 4-19a) The region containing the sfGFP expression cassette with pLux was amplified. (Figure 4-19b) C-micL region was also amplified via PCR. (Figure 4-19c) The fragments were assembled in the pET22b vector resulting in synthesis of pET22b pLux sfGFP c-micL vector. The cloned vector was further linearized with PCR for insertion of LuxI gene with constitutively active promoter. (Figure 4-19d)

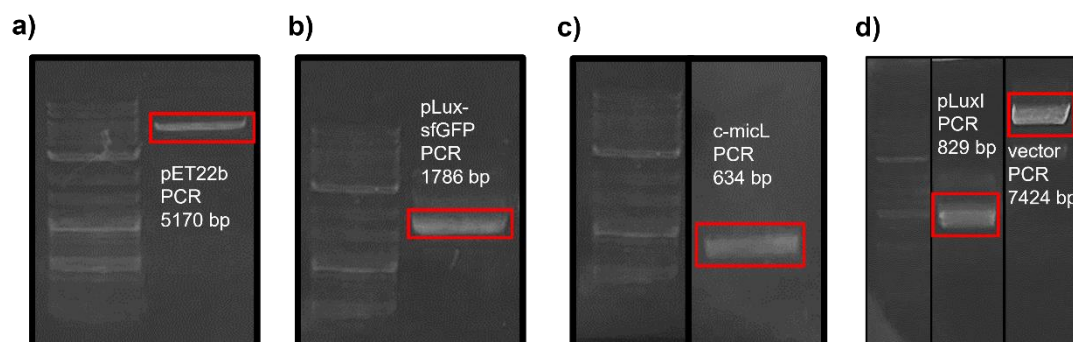


Figure 4-19. Construction of quorum sensing circuit to test functionality of c-micL. a) PCR amplification of pET22b vector. b) Extension of pLux-sfGFP with PCR to add overhangs for assembly. c) Amplification of c-micL for assembly. The obtained fragments were assembled with Gibson assembly to obtain pET22b pLux-sfGFP c-micL plasmid. d) PCR amplification of pET22b pLux-sfGFP c-micL and pLuxI. During LuxI PCR, constitutively active promoter was added at the upstream of LuxI. All DNA bands appeared at their expected sizes.

The quorum sensing gene network of *Vibrio fischeri* (*V. fischeri*) consists of LuxR transcriptional factor and its cognate promoter pLux, and LuxI enzyme which synthesizes the AHL molecule. Constant production of LuxI enzyme inside cells causes the accumulation of AHL in the culture media. At the threshold concentration of AHL, the LuxR is activated by AHL thereby initiating the transcription of genes at the downstream of pLux. The accumulation of AHL is directly proportional to the number of cells present in the media. (Figure 4-20a) To test the system, a full quorum sensing circuit that contains the c-micL at the downstream of pLux was constructed. In addition, two controls were used in the experimental setup. In the first control, c-micL quorum sensing circuit without LuxI was used whereas full quorum sensing circuit without c-micL was also included as a control. Cells were co-transformed with arabinose inducible bALP to determine whether periplasmic

proteins were released to the media. Cells bearing the plasmids were grown for overnight and rediluted by 100-fold into media consists of fresh and spent media mixed with 1:1 ratio. Diluted cells were grown again overnight, and their supernatants were collected to determine amount of secreted bALP via phosphatase assay. Based on the supernatant bALP activity, cells bearing the full quorum sensing circuit with c-micL had significantly more bALP activity than the control groups thereby indicating that c-micL is functioning in the quorum sensing circuit. (Figure 4-20b)

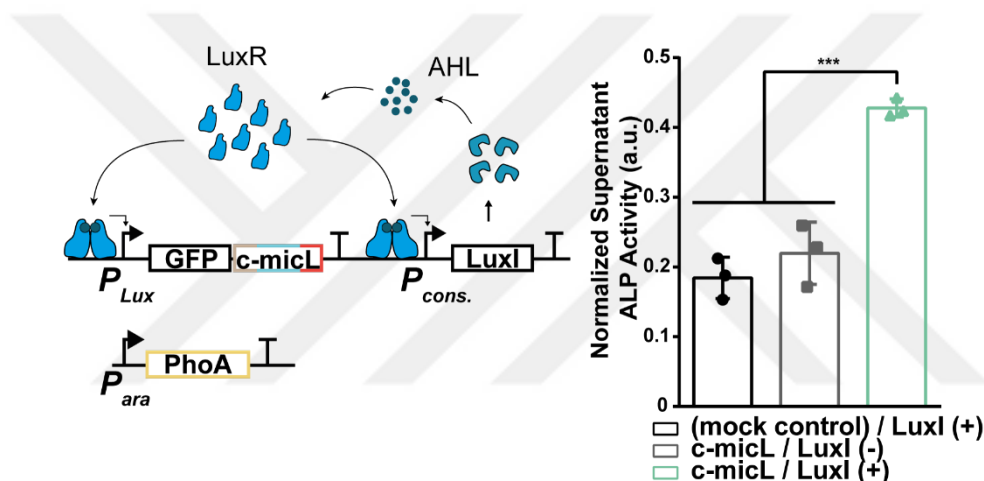


Figure 4-20. Functional validation of c-micL activity in quorum sensing circuit.

GFP gene and c-micL were expressed under control of pLux promoter whose induction was controlled by LuxI produced AHL molecule. The LuxI gene was expressed with constitutively active thereby leading to constant accumulation of AHL in the media. Thus, the amount of cell present in media is directly proportional to amount of AHL produced. As a output, bALP was expressed whose secretion from cells were determined from cell free supernatant.

4.4.11. Effects of micL Expression on EcN Physiology

Effects of micL expression on EcN cells was analysed by determining the three physiological parameters after the micL induction, growth rate of the EcN cells, motility of bacteria cells and invasion of Caco-2 cells. Following to the micL induction, cell growth of WT EcN cells and MicL expressing EcN cells were measured for 4 hours. Growth curve indicated that micL expression have no major effect on bacterial growth. (Figure 4-21a) The invasion of Caco-2 cells by micL expressing EcN was determined. Caco-2 colorectal carcinoma cells were chosen as the model cell line for the invasion assay because EcN cells will counter cells in the gut when administered as living therapeutics. To do so, EcN cells were induced for micL expression, and collected. Subsequently, EcN cells were washed with 1x PBS and resuspended in DMEM at concentration of approximately 5×10^8 cells/ml. Meanwhile, Caco-2 cells that were grown till their maximum confluency were washed 1x PBS. 100 μ l of EcN cells suspended in cells were added on Caco-2 cells and incubated. Following the incubation, Caco-2 cells were washed and invaded EcN cells were collected with Trinton X-100 treatment. CFU were counted by serial dilution and spreading the collected EcN cells on agar plate. Results indicated that invasion ability of micL expressing EcN for Caco-2 is decreased (Figure 2-21b). Following the invasion assay, motility of micL expressing EcN cells were determined with motility assay. In this assay, 10 μ l of overnight grown cells were seeded in the middle of the agar plate with %3 agar. Next day, cells spread was measured from the agar plate. Motility experiments revealed that motility of micL expressing EcN was decreased compared to control group. (Figure 2-21c)

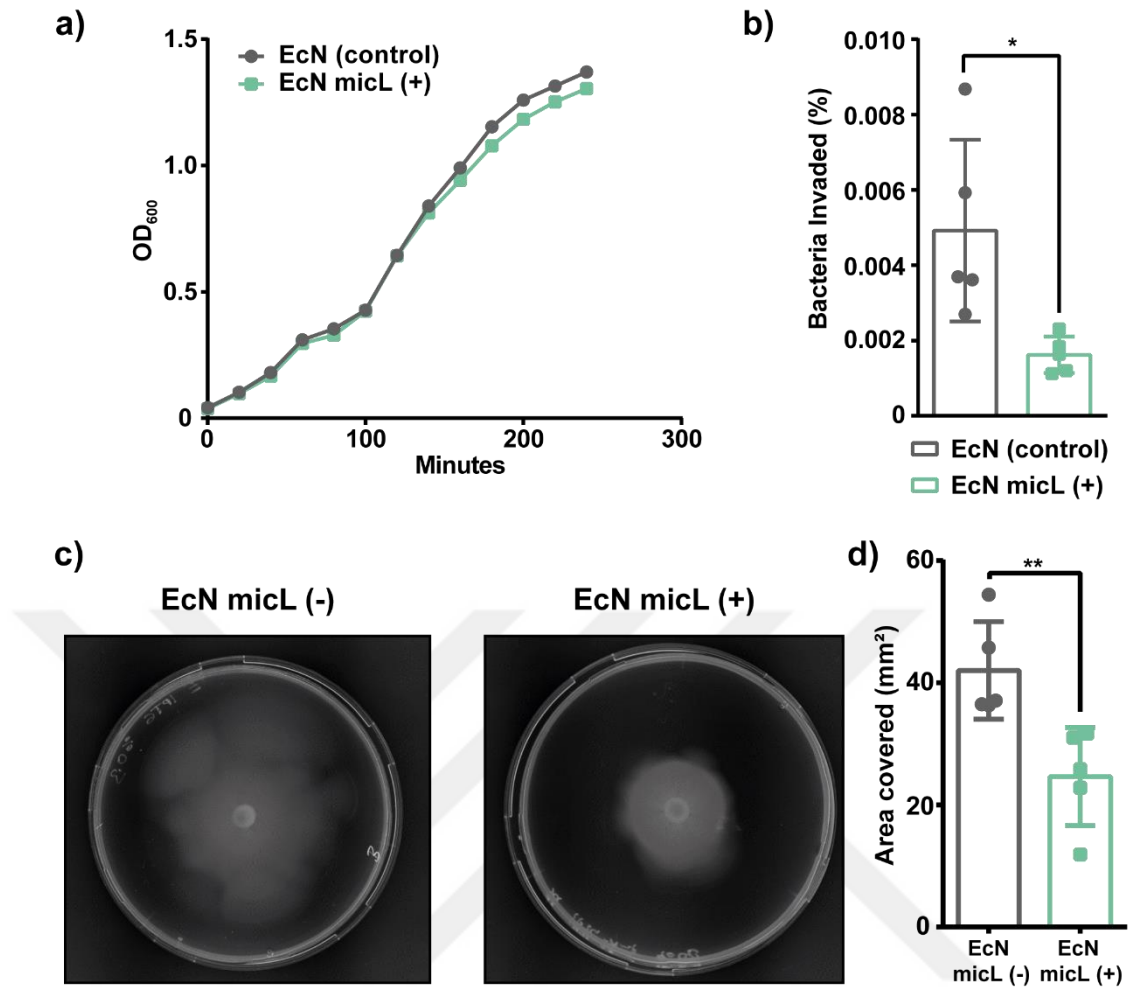


Figure 4-21. Effects of micL expression on EcN physiology. a) Growth curve of WT and micL sRNA expressing EcN. Optical density at 600nm was measured to monitor growth for 4 hours. Three biological replicates were used for the growth assay. b) Caco-2 invasion assay with WT and micL sRNA expressing EcN strains. The experiment was conducted with five biological replicates. c) Motility of WT and micL sRNA expressing EcN strains. The experiment was performed with five biological replicates.

4.4.12. Safety Assays on Caco-2 cells for iLOM-Induced Leaked Periplasmic Content of EcN cells

To assess the safety of iLOM-SS used in EcN, effects of leaked periplasmic content were measured on Caco-2 cells since periplasmic content would be leaked along with the recombinant proteins. (Figure 4-22a) Firstly, the effect of leaked periplasmic content from EcN on Caco-2 proliferation was measured using MTT. To do so, EcN were grown in DMEM and micL expression was induced for 3 hours. The DMEM was collected and filtered following the induction. Subsequently, DMEM media was supplemented with FBS, L-glutamine and antibiotics to get full growth media. DMEM that contains the leaked periplasmic content of EcN was added onto Caco-2 cells and cells were incubated for 3 days. After incubation, MTT assay was performed to assess the proliferation of Caco-2 cells. MTT assay results revealed that leaked periplasmic content didn't have toxic effect on Caco-2 cells. There was a statistically significant difference between DMEM only and DMEM obtained from micL expressing EcN groups, but the increase was small which was approx. $\times 1.05$ fold. (Figure 4-22b) In addition to proliferation, the effect of leaked periplasmic content on Caco-2 migration was assessed via migration assay. In this experiment, a hole was created in the middle of well that contained Caco-2 cells at maximum confluency. Following the hole creation, cells were incubated for a day, and supplemented DMEM obtained from EcN cells were onto cells. Closure of holes was monitored for 3 days using microscope. The dimensions of the holes was measured from microscopy pictures that were obtained for each day. The percentage closure of the holes was calculated. The results revealed that leaked periplasmic content has no effect on migration of Caco-2 cells. Holes were healed at the same rate for each group. (Figure 4-22d)

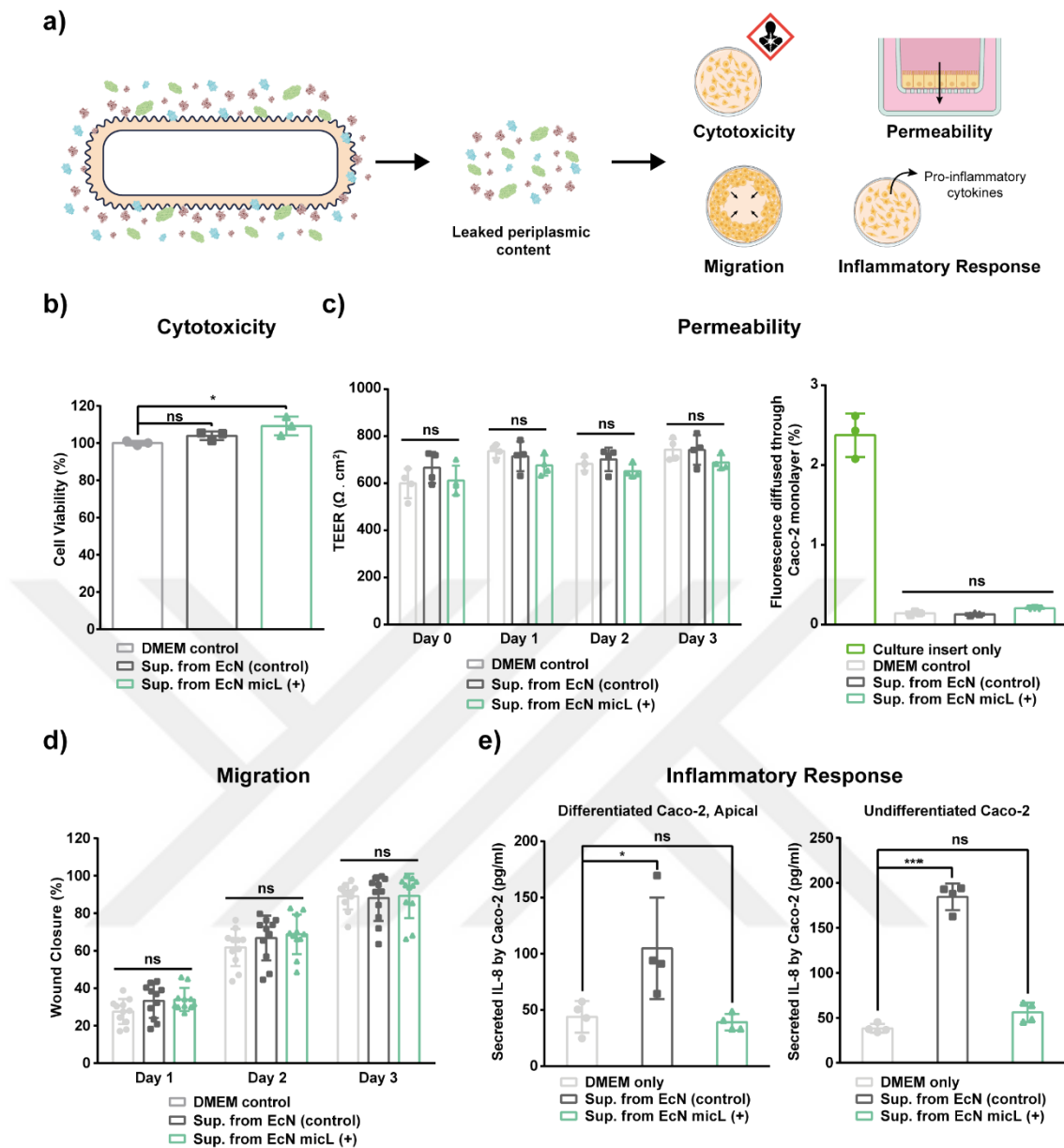


Figure 4-22. In vitro safety assessments of iLOM-based secretion from EcN. a) Schematic presentation of parameters that were tested with Caco-2 cell line. During the iLOM-based secretion, periplasmic content including proteins and other metabolites leaks through to extracellular space along with the PoI. b) Cytotoxicity of leaked periplasmic content from EcN. The experiment was conducted with three independent replicates c) Effects of leaked periplasmic content on permeability of differentiated Caco-2 monolayer. TEER (right) and FTIC conjugated dextran (4 kDa) diffusion (left) were determined to assess the permeability. TEER measurement was

performed with four replicates while diffusion assay was done with three replicates. d) Wound closure assay was conducted to analyze the cell migration upon treatment with leaked periplasmic content. Percentage closure was calculated to normalize the data. Eleven replicates were used in the migration experiment. e) IL-8 secretion by Caco-2 cells treated with leaked periplasmic content. IL-8 amount was determined for undifferentiated (left) and apical side of differentiated (right) Caco-2 cells with ELISA. Four biological replicates were used.

The effect of leaked periplasmic content on Caco-2 monolayer permeability was also assessed. To do so, Caco-2 cells were differentiated to monolayer polarized enterocytes-like cells for 21 days on cell culture inserts. Differentiation was validated by measuring the transepithelial electrical resistance (TEER). After obtaining the monolayer polarized Caco-2 cells, supplement DMEM obtained from micL expressing was added onto apical side of the polarized cells. TEER was measured in each day for 3 days. TEER measurement showed that the periplasmic leaked content has no effect on permeability of polarized Caco-2 cells. (Figure Figure 2-22c) As further validation, diffusion of FTIC conjugated 4kDa dextran through polarized Caco-2 cells was measured at the third day. FTIC-4kDa diffusion assay also indicated that periplasmic leaked content didn't disturb barrier activity of the polarized Caco-2 cells. (Figure 2-22c)

IL-8 secretion by both differentiated and undifferentiated Caco-2 cells were measured upon treatment with leaked periplasmic content obtained from EcN. Cells were treated with leaked periplasmic content described above for 3 days. Following the incubation, cell media was collected, and IL-8 concentration was determined with ELISA. Although, supplemented DMEM obtained from WT EcN causes a

significant amount of IL-8 expression compared to DMEM control, there is no expression difference between control and supplemented DMEM obtained from micL expressing EcN. It is revealed that leaked periplasmic content didn't result in induction of IL-8 secretion. (Figure 2-22e)

4.4.13. IL-1Ra Secretion with iLOM-SS Equipped EcN

As model anti-inflammatory cytokine, IL-1 receptor agonist (IL-1Ra) was secreted by EcN using iLOM-SS. IL-1Ra cloned in pET22b pLacO micL / pNtetO bALP vector at the location of bALP. To do so, the vector was digested to remove bALP from plasmid and linearize the plasmid DNA. (Figure 4-23a) IL-1Ra was amplified with compatible overhangs for Gibson assembly via PCR. (Figure 4-24b) The vector and IL1-Ra encoding fragment were assembled.

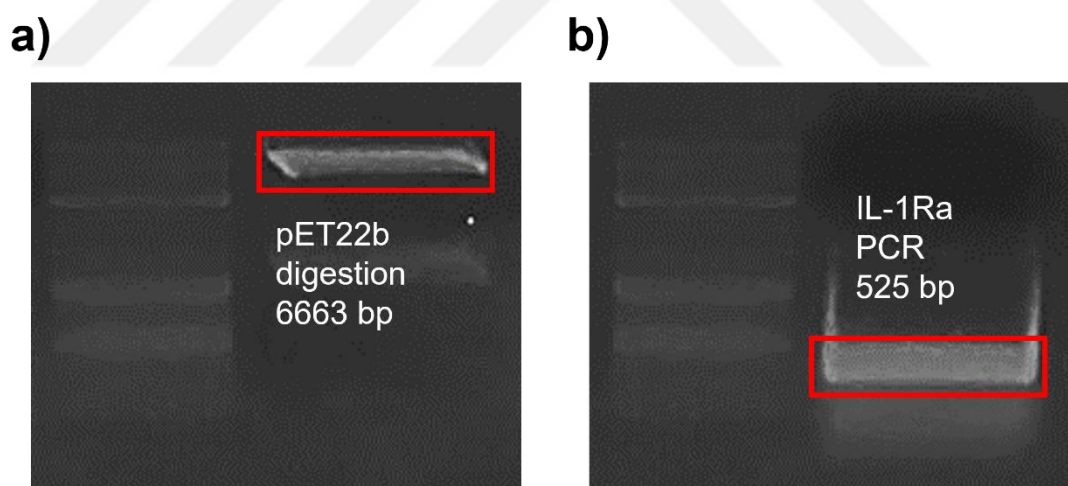


Figure 4-23. Construction of pET22b pLacO micL / pNTetO IL-1Ra plasmid. a) Digestion of pET22b pLacO micL / pNtetO plasmid to remove bALP. b) PCR amplification of IL-1Ra with compatible overhangs for Gibson assembly. All DNA bands appeared at their expected sizes.

Cells were induced for expression of anakinra as well micL for overnight. Subsequently, cell supernatant was collected and presence of anakinra was validated with both ELISA and Western Blot analysis. ELISA results revealed that micL based secretion system was improved the extracellular secretion of anakinra 7-fold compared to control group that has only expression cassette for periplasmic translocation of anakinra. (Figure 2-24b) Western blot analysis with anti-anakinra showed that anakinra was secreted to extracellular space. (Figure 2-24c) The activity of secreted anakinra was validated with A375.S2 *in vitro* model. A375.S2 is sensitive IL-1 α which causes activation of apoptosis cascade via IL-1R. Anakinra can prevent the activation of this cascade by preventing the binding of IL-1 α to IL-1R. (Figure 2-24a) Activity of anakinra can be measured by MTT assay from IL-1 α treated A375.S2 cells. To test activity of anakinra, EcN cells were induced and supernatant that contains the secreted anakinra was collected. The collected supernatant was concentrated for 50x with ultrafiltration. Different amount of the concentrated supernatant was added onto IL-1 α treated A375.S2. After 3 days incubation, MTT assay performed to check viable cells. MTT results showed that secreted anakinra was active and prevent apoptosis induced by IL-1 α . (Figure 2-24d).

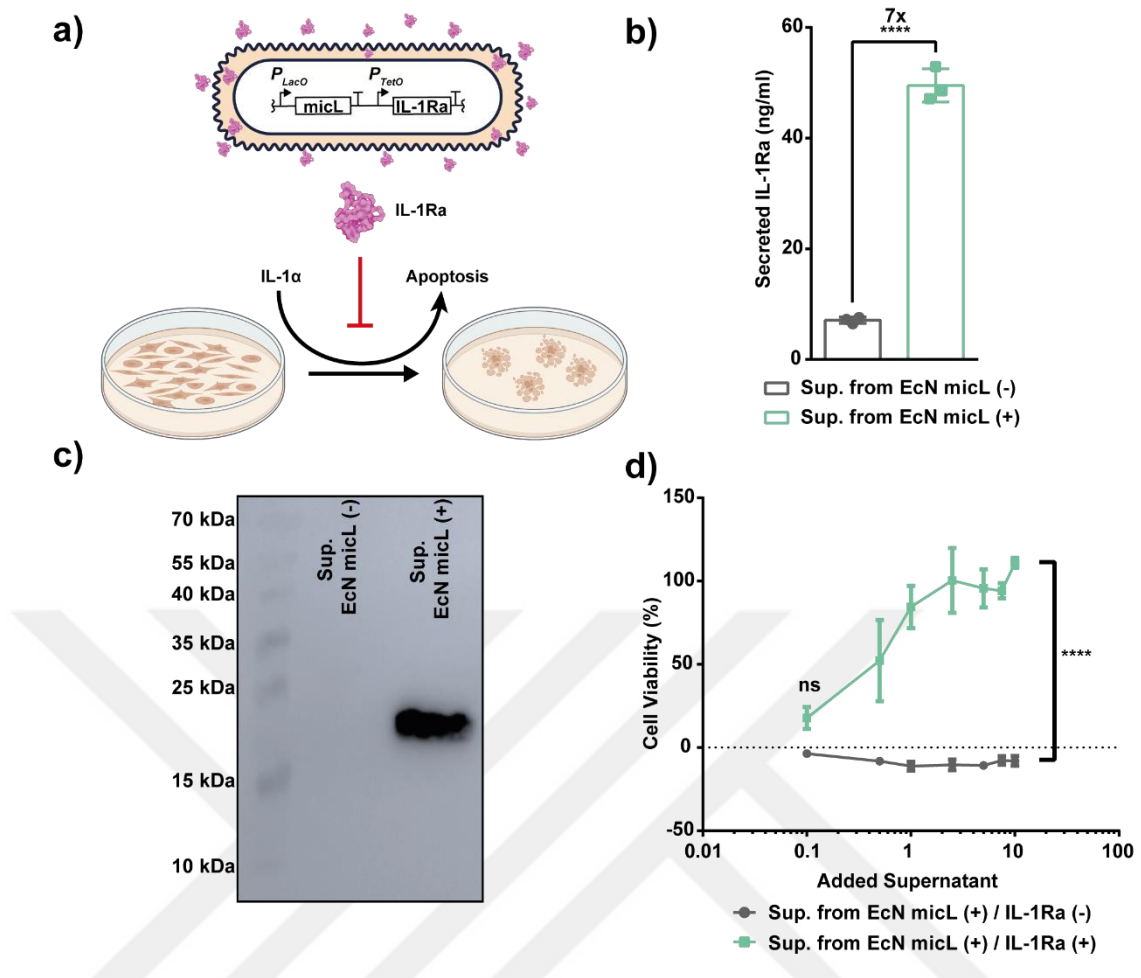


Figure 4-24. Active *IL-1Ra* secretion using iLOM-SS by *EcN*. a) Schematic presentation of the genetic circuit to secrete *IL-1Ra* with iLOM-based secretion system and in vitro assay to determine the activity of *IL-1Ra*. The activity of secreted *IL-1Ra* was determined using A375.S1 cells. In the presence of active *IL-1Ra*, *IL-1α* induced apoptosis of A375.S2 cells are inhibited which can be detected by measuring cell viability after the *IL-1α* treatment. b) Secreted *IL-1Ra* amount to extracellular space with iLOM-based secretion system. The concentration was determined via *IL-1Ra* specific ELISA without any enrichment method e.g., ultrafiltration. c) Verification of *IL-1Ra* at the supernatant following the induction. Anti-*IL-1Ra* antibody was used to capture the Western blot image. d) Cell viability of A375.S2 cells determined by utilizing MTT assay after the treatment with *IL-1α* and the

supernatant obtained from either IL-1Ra secreting EcN cells or control cells. The experiments were conducted with at least three independent replicates.

4.4.14. Nanobody Secretion with iLOM-SS Equipped EcN

As a model antibody, 2Rs15d nanobody was secreted using iLOM-SS from EcN. The nanobody was cloned in the pET22b pLacO micL / pNtetO bALP vector at the location of bALP. The vector was digested to take out bALP coding region from the plasmid. (Figure 4-25a) Nanobody encoding region was extended via PCR to add Gibson overhangs. (Figure 4-25b) The digested vector and fragment encoding nanobody was assembled.

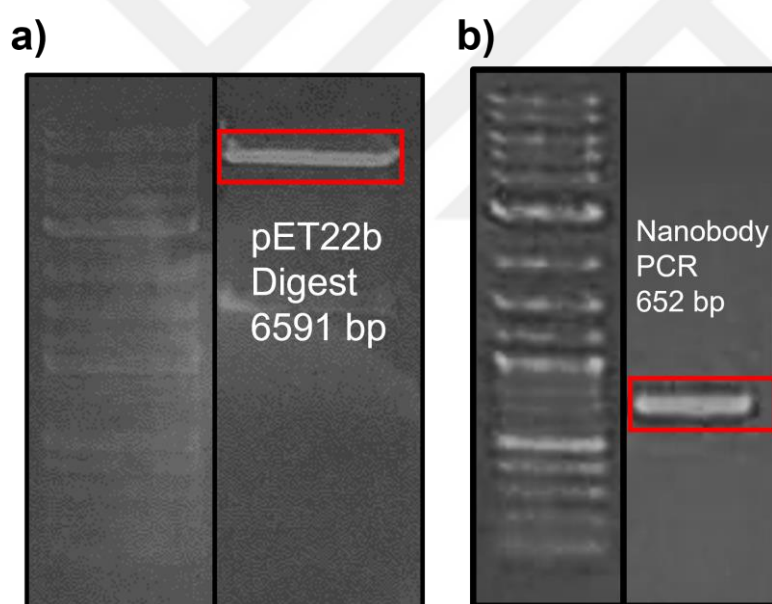


Figure 4-25. Construction of pET22b pLacO micL / pNTetO nanobody plasmid.

a) Digestion of pET22b pLacO micL / pNtetO plasmid to remove bALP. b) PCR amplification of 2Rs15d nanobody with compatible overhangs for Gibson assembly.

All DNA bands appeared at their expected sizes.

Following the cloning of nanobody in the pET22b pLacO micL / pNtetO vector, the signal sequence of nanobody was changed to determine the effect of signal peptide on periplasmic translocation thereby extracellular release of the protein. (Figure 4-26a) To change pelB-ss signal peptide, the vector was digested and linearized. The DNA fragments encoding the signal peptides which are dsbA-ss, ompA-ss, lpp-ss, malE-ss, LTB-ss, gIII-ss were constituted with same protocol described for the cloning of gRNAs. The DNA fragments were ligated to the vector, separately. For the TorA-ss, the DNA encoding the signal peptides was amplified from genome with compatible overhangs, then assembled into the vector via Gibson Assembly. (Figure 4-26b)

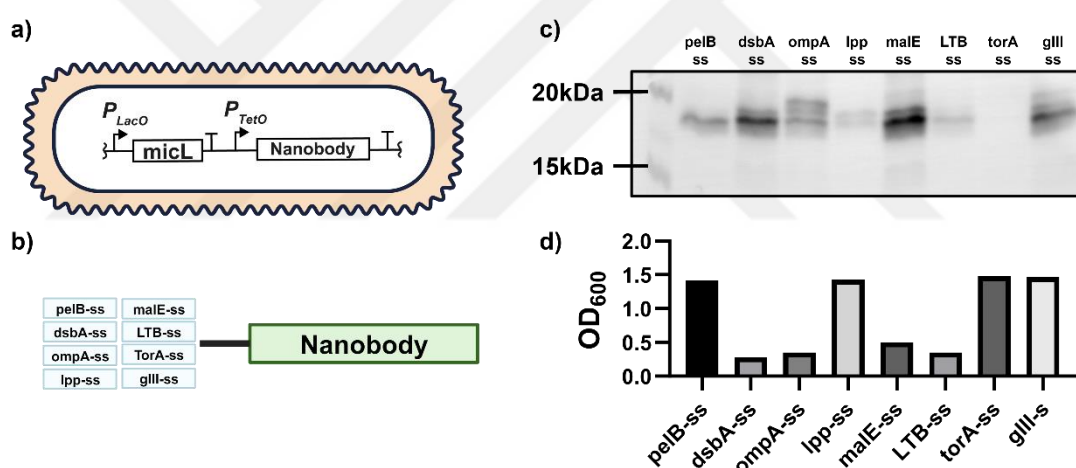


Figure 4-26. Nanobody secretion using iLOM-SS by EcN. a) The presentation of expression cassette used in the secretion of the nanobody. Expression of *micL* was controlled by pLacO promoter whereas nanobody expression was induced by using pNTetO. b) The depiction of signal peptides used in the secretion of the nanobody. Secretion signals of pelB, dsbA, ompA, lpp, malE, LTB, TorA and gIII was used. c) Western blot analysis of secreted nanobody. Cell-free supernatants from overnight

induced culture were used in the analysis. d) OD₆₀₀ of culture at the time of supernatant collection.

Cells bearing the nanobody expression cassette with iLOM-SS vectors were induced overnight. Next day, the supernatant of induced culture was collected and analyzed with western blot. Western blot analysis revealed that all signal peptides expect for the TorA-ss was able to drive translocation of nanobody to the periplasm. (Figure 4-26c). Nonetheless, OD₆₀₀ of induced culture clearly indicated that the use of dsbA-ss, ompA-ss, malE-ss and LTB-ss put heavy burden on cell metabolism. (Figure 4-26d)

4.5. Discussions

In this study, a protein secretion system called iLOM-ss was engineered, which did not exist in nature previously, to allow *E. coli* strains to release a wide range of proteins into the extracellular space. Within the iLOM-ss, proteins translocated to the periplasm were released into the extracellular milieu by inducing the LOM phenotype through the translational suppression of the lpp protein by micL sRNA. This secretion system was conceived to overcome the limitations associated with natural secretion systems, including high metabolic demands and narrow substrate profiles. The iLOM-ss was shown to be functional in various *E. coli* strains, including the industrially recognized BL21 (DE3) protein expression strain [235] and the clinically proven probiotic EcN strain. [236] While the primary focus of characterization was on iLOM-SS in the EcN strain for therapeutic cargo delivery, the versatility of iLOM-SS among *E. coli* strains suggests potential applications in various contexts where protein secretion is required.

Titratable action and stop switches are essential features for biological therapy agents, especially for living therapeutics because life-threatening adverse reactions can occur after administration, some of which cannot be predicted prior to human clinical trials. [237] Therefore, there is a need to design genetic switches for controlling the action of advanced biological agents such as engineered bacteria, mRNA, and CAR-T cells. [238]–[241] In this study, the secreted protein amount with iLOM-SS was shown to be titratable with a small molecule inducer. Although the titration experiments were conducted with aTc, which is not applicable for human administration, there are promoter/transcriptional factor pairs whose inducers are safe to use in humans. [242] Also, a stop switch is designed for halting the protein secretion in iLOM-SS with small molecules in case of need. Altogether with these tools, controlling the secretory action of iLOM-SS augmented microbes is possible with external clues. However, future studies are required to confirm their performance *in vivo*.

The current genetic tools for controlling protein secretion are not fast enough for many physiological processes in the human body, such as blood sugar regulation because many of them rely on transcription and translation of a gene which require relatively long times compared to the time needed. [243] Here, it is demonstrated that iLOM-SS can release proteins within minutes of being activated if the protein is already translocated to the periplasm. This makes iLOM-SS a promising system for designing bacteria as protein delivery systems for applications that require a fast response.

The fundamental concept of modularity in genetic tools is a cornerstone in the field of biological engineering, as it expedites the iterative process of designing, building,

and testing multiple genetic circuits with simultaneously. [244] To implement this modularity, a compact version of iLOM-SS was developed, enabling the co-expression of micL sRNA and the protein of interest within a shared transcriptional unit. To highlight the modularity of this compact iLOM-SS, its operability was validated across diverse genetic circuit architectures designed to detect small molecules or register the presence of a small molecule inducer. The compact iLOM-SS can serve as a versatile platform, empowering researchers, and enthusiasts in the realm of synthetic biology to craft genetic networks with protein secretion as an output signal. Thanks to the user-friendly and plug-and-play design of the compact iLOM-SS, the democratization of protein secretion is poised to revitalize the landscape of synthetic biology research, where protein secretion stands as a requirement.

As proof of concept, the secretion of IL-1Ra and a nanobody were demonstrated in this study. iLOM-SS augmented EcN has the potential to be employed in the development of sentinel bacteria for the treatment of a variety of gut pathologies, including inflammatory bowel disease and celiac disease because many studies showed that *E. coli* strains can express various proteins recombinantly in their active forms. [245]–[249] Additionally, it was observed that the leaked periplasmic content of EcN during secretion with iLOM-SS did not induce pro-inflammatory or cytotoxic effects *in vitro* on the Caco-2 human epithelial cell model, thus indicating that LPC (leaked periplasmic content) does not exhibit any apparent negative impacts on human cells and can be tolerated. Nevertheless, these findings should be validated with *in vivo* models to ensure the safety of protein delivery facilitated by iLOM-SS augmented EcN.

CHAPTER 5

Engineering Human Gut Commensal

Lachnospiraceae Species for Protein Delivery

5.1. Objective of the Study

The human body possess a diverse microbiota which has been linked with overall well-being and health of the body because of the constant and intricate interactions between the body and the microbes to maintain which provide means to maintain the homeostasis. Lachnospiraceae species which consists of strictly anaerobic microbes often positively associated with healthy microbiota by providing beneficial metabolites for the gut metabolism. Therefore, Lachnospiraceae species can provide a unique platform for engineering microbial therapeutics, however many of these organisms are understudied. Thus, many molecular tools are lacking to engineer these microbes for therapeutical purposes. Here we developed signal peptides as well as a protein display shuttle for driving the extracellular protein translocation in Lachnospiraceae species

5.2. Introduction

The human gut is an ecological niche for a vast number of microbes whose interactions with the human body can be beneficial, neutral, or harmful for the host. [250] Strict anaerobic bacterial species are the most prevalent organisms in gut compared to other microorganisms. [251] Generally, healthy microbes are defined by relative abundance of bacteria species belonging to phyla of Firmicutes (recently

changed to Bacillota), Bacteroidetes (recently changed to Bacteroidota), Proteobacteria, Actinobacteria. [252] Although compositions and diversities of the microbiota for a given body site e.g., ileum is relatively constant and is not affected by minor perturbations, the microbiota differs greatly among unrelated individuals. However, studies have suggested that many functional molecular pathways are ubiquitously shared between different people. [253] In case of strong assault on the microbiota such as antibiotic treatment, the constellations of microbial species drastically changed to a dysbiotic state [254], [255] in which human health is impacted directly or indirectly due to the complicated relationships microbes with other microbes and with the host organ. [256]–[258]

Bacteria species found in Lachnospiraceae family constitute one of major families found in health adult microbiota, and their abundance is significantly decreased in dysbiosis. [259], [260] Many studies have revealed that Lachnospiraceae species are positively associated with the human health owing their inherent characteristics and metabolism. [261]–[264] For instance, some members of Lachnospiraceae family can convert the primary bile acids to secondary bile acid compounds thereby preventing the colonization of *C. difficile* pathogen. [265] Most notably, many Lachnospiraceae species can ferment cellobiose and other sugars to produce immunologically relevant such as short chain fatty acids such as butyrate [266] which contributes positively both immune regulatory mechanisms and metabolic functions. [267], [268]

Owing to their unique abilities, Lachnospiraceae species offer an unexplored avenue for developing living therapeutics however there are few molecular tools available to tinker their functions. Extracellular protein translocation including secretion and surface display is one of the key functions required for intervening disease

mechanism. Many disease-related molecular targets e.g., proteins, and metabolites cannot free diffuse through the cellular membrane thus the therapeutical effector proteins should be in the extracellular space to exert their functions such as depletion of a cytokine, degradation of a key metabolite and activating a receptor on other cells. [269] Here, we developed molecular technologies for extracellular protein translocation in Lachnospiraceae species.

5.3. Materials and Methods

5.3.1. Construction of Plasmids Used in the Study

All plasmids used in this study were constructed with Gibson Assembly. Plasmid sequences were verified with either Sanger sequencing or NGS. All enzymes used in this for plasmid construction were purchased from New England Biolabs (NEB). Primers were ordered from IDT Inc. If needed, the gene synthesis service of University of Chicago was utilized. All plasmid maps, primers, and sequencing results were given in Appendixes. LB media and LB agar with appropriate antibiotics were used for *E. coli* cultivation unless otherwise noted. *E. coli* S17 λ pir strain was used for cloning experiments.

Signal peptides were cloned upstream of nanoluciferase in *E. coli* - Clostridia shuttle plasmid. The plasmid was linearized with PCR by using readD3 and JHA.J1 primers. The signal peptides were amplified from the genome of *R. gnavus*. For ssREA-2, readD3 and readD4 were used in the PCR. For ssREA-3, readD5 and readD6 were used in the PCR. For ssREA-10, readD17 and readD18 were used in the PCR. For ssREA-11, readD19 and readD20 were used in the PCR. For ssREA-12, readD21 and readD22 were used in the PCR. For ssREA-13, readD23 and readD24 were used in the PCR. For

ssREA-14, the signal peptide coding region was synthesized via assembly PCR using reaD25, reaD26, reaD27, and reaD28 primers. ssJack-1 containing was obtained from Mimee Lab. The PCR products and linearized vector were assembled separately.

To the display shuttles, the plasmid encoding the ssREA-3 signal peptide was amplified with reaD7 and JHA.C08. The variants of IpbA proteins were amplified from *R. gnavus* genome. The 379-607 variant was amplified employing reaD9 and reaD8, while for the 510-607 variant, amplification was carried out with reaD10 and reaD8. Similarly, the 560-607 variant was amplified with the primer pair reaD11 and reaD8, and for the 585-607 variant, the amplification was achieved using reaD12 and reaD8. Each of these fragments was independently assembled with the linearized vector.

5.3.2. Cultivation of Lachnospiraceae species

R. gnavus and *C. comes* were grown in brain heart infusion media (premix, Thermo Bacto) supplemented with yeast extract (5g/L), resazurin (1mg/1L), hemin (5mg/ml), menadione (1mg/L), and 1x vitamin solution (Table 5.1.), and 0.05% cysteine. For the solid media, the brain heart infusion agar mix (Thermo Bacto) was used. Prior to the addition of vitamins and other supplements, the premix, yeast extract and resazurin were dissolved then the mixture was autoclaved. Both liquid and agar media were prereduced in the anaerobic chamber at least 24 hours before the inoculation. Cells were streaked on the agar plate; then, single colonies were picked to grow in the liquid media. For proteomics experiments, *R. gnavus* cells were grown in the YPG media without hemin. (Table 5.2. and Table 5.3.) Cells were supplement

with 100µg/ml kanamycin and 10µg/ml thiamphenicol following the conjugation.

For induction experiments, 250ng/µl aTc was used.

Table 5.1. 5000X vitamin solution

Component	Amount for 100 mL
Biotin	10 mg
Cobalamin (Vitamin B12)	0.5 mg
Para-aminobenzoic acid	25 mg
Folic acid	10 mg
Pyridoxamine	50 mg
Thiamine	25 mg
Riboflavin	25 mg

Table 5.2. YPG Media

Compound	Amount per Liter
Peptone	20 g
Yeast Extract	10 g
Phosphate Buffer	1 mL
Resazurin	4 mL

TYG Salt solution	40 mL
0.8% Calcium chloride	1 mL
dH ₂ O	950 mL
<i>After Autoclaving</i>	
Cysteine-HCl	10 mL
Hemin	5 mL
Vitamin K1 (menadione)	1 mL
Glucose (0.5%)	5g

Table 5.3. TYG salts solution

Ingredient	per liter
MgSO ₄ . 7H ₂ O	0.5 g
NaHCO ₃ (bicarbonate)	10 g
NaCL	2.g
Water	to 1000 ml

5.3.3. Sample Preparation for Proteomic Analysis for Secreted and Intracellular Proteomes of *R. gnavus*

R. gnavus cells were grown as described. Cell pellet was collected via centrifugation at maximum speed for 5 minutes. The collect cell pellet was washed three times with 1x PBS. Then, cells were lysed in SDS lysis buffer (%5 SDS, 50mM, TEAB, pH 7.55). At the same time, protease inhibitor cocktail was added into solution. The dissolved pellet was incubated on ice for 30 minutes for complete lysis. The soluble fraction was collected via centrifugation at maximum speed for 30 minutes at +4 °C. The cell free supernatant after the growth was concentrated 500x fold via 10 kDa ultrafiltration tube. Following the concentration, protease cocktail was added into concentrated supernatant to minimize the protein degradation.

5.3.4. Identification of Signal Peptides in *R. gnavus* Genome using SignalP6

Open reading frames in *R. gnavus* genome (ATCC-29149) were detected using the prokka tool. [270] Then, ORFs were analyzed using the Signal6P software to identify potential signal peptides. [271]

5.3.5. De novo Design of the Signal Peptide Based on Conserved Regions in Signal Peptides of *R. gnavus*

All predicted sec signals peptides were aligned using Glam2 to find conserved motif among the signal peptides.[272] Most conserved regions determined by the Glam2 tool were collected and rearranged based on the general signal peptide structure of sec translocon.

5.3.6. Manual Literature Search to Find Experimentally Validated Secreted Proteins in Lachnospiraceae Species

Pubmed was used to search papers related to protein secretion in Lachnospiraceae species. “protein secretion Lachnospiraceae”, “extracellular proteins Lachnospiraceae”, “*R. gnavus* protein”, and “*R. gnavus* extracellular enzyme” keywords were used. The papers were scanned manually to find papers that includes experimental validation for secreted proteins.

5.3.7. Conjugation of Recombinant Plasmids into Lachnospiraceae Species

The shuttle plasmids were conjugated into *R. gnavus* and *C. comes* by using agar spotting. Briefly, plasmid bearing *E. coli* s17 λ pir cells were grown overnight aerobically. Next day, *E. coli* cells were collected and washed once with 1x PBS supplemented with cysteine. Subsequently, cells were transferred into anaerobic chamber. Cells were further washed with two times with prereduced 1x PBS supplemented with cysteine. *E. coli* cells were mixed with overnight grown recipient Lachnospiraceae cells at 5:1 ratio. The cells mixture was spot on the prereduced BHIS agar and incubated at 37 °C for 24 hours in the anaerobic chamber. After the incubation, the bacterial lawn was scrapped from top the agar and dissolved in the prereduced 1X PBS. Half of the cells were spread on the prereduced BHIS agar supplemented with kanamycin and thiamphenicol. Following the 2 days incubation, single colonies were picked from the plate and streaked on agar plate supplemented with antibiotics. The streak plates were incubated two more days to get recombinant recipient cells that were carrying the plasmid.

5.3.8. Nanoluciferase Assay

Conjugated Lachnospiraceae cells were grown in the prereduced BHIS broth media supplemented with antibiotics for overnight. Next day, cells were rediluted 1:00 into fresh prereduced BHIS media supplemented with aTc and antibiotics. Subsequent to 24-hour induction, cells and cell-free supernatants were separated by centrifugation. Cell pellets were resuspended in the 1X PBS. Both cell-free supernatants and suspended cell solutions were diluted by 10-fold with 1X PBS. Diluted samples were used to determine nanoluciferase activity via NanoGlo Kit (Promega)

5.3.9. Trypsin Surface Trimming of Lachnospiraceae Species

Cells carrying the plasmids for nanoluciferase surface display shuttles were induced same as described above. Following the induction, cells were collected with centrifugation and their supernatants were discarded. Cell pellets were suspended in the PBS, then trypsin enzyme was added into cell suspension at the final concentration of 0.025% (w/v). After this, cells were incubated at room temperature for one hour. Subsequently, cells were centrifuged, and supernatants were discarded to remove the trypsin enzyme. Retained nanoluciferase activity after the trypsin trimming was measured same as described above.

5.4. Results

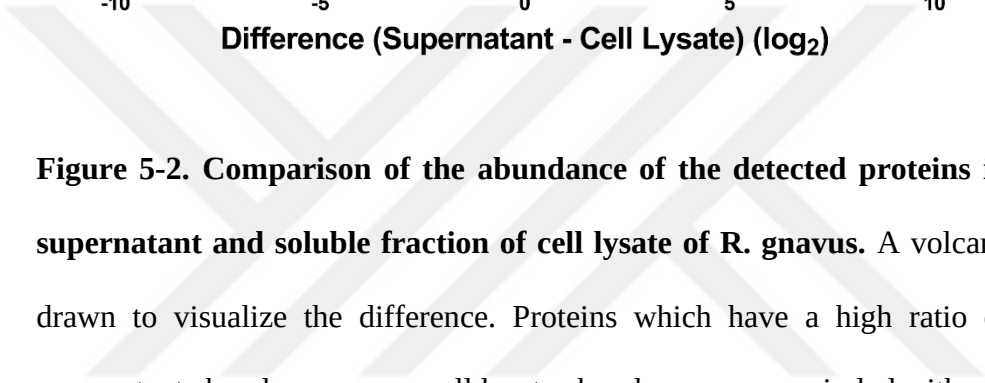
5.4.1. Signal Peptide Selection Approaches for Protein Secretion in Lachnospiraceae

Four different strategies were employed to determine candidate signal peptides whose protein secretion efficiency would be screened in Lachnospiraceae species. In

the first strategy, a proteomics study was carried out to identify secreted proteins during in vitro growth of *Ruminococcus gnavus*. For the second strategy, literature was manually scanned with keywords given in the methods section for *R. gnavus* to find proteins whose secretion was validated experimentally in other studies. In the third strategy, the genome of *R. gnavus* was analyzed with bioinformatic tools to find all signal peptides, and the signal peptide of the protein with highest probability of having a signal peptide was selected. In the fourth strategy, a synthetic signal peptide was constructed based on the alignment of all potential sec signal peptides identified from the third strategy.

5.4.2. Shotgun Proteomic Analysis of *R. gnavus* Intracellular Soluble Fraction and Secretome

Secreted proteins from *R. gnavus* organism during its in vitro growth were determined using mass spectroscopy. *R. gnavus* cells were grown in YPG media (Anaerobic Systems) that lacked heme. After the growth, cell free supernatant was obtained via centrifugation and concentrated using an ultrafiltration tube for 100x fold. Cell pellets were washed with PBS; subsequently, was lysed with lysis buffer (SDS lysis buffer (%5 SDS, 50mM TEAB, pH 7.55). SDS-PAGE gel was run to check protein bands before MS analysis. Clear proteins were observed for both concentrated supernatant and cell lysate. (Figure 5-1a) Proteins present in cell lysate and concentrated supernatant were determined using mass-spectroscopy. (Figure 5-1b)



supernatant abundance versus cell lysate abundance were encircled with red circles.

0 1 1 (0)

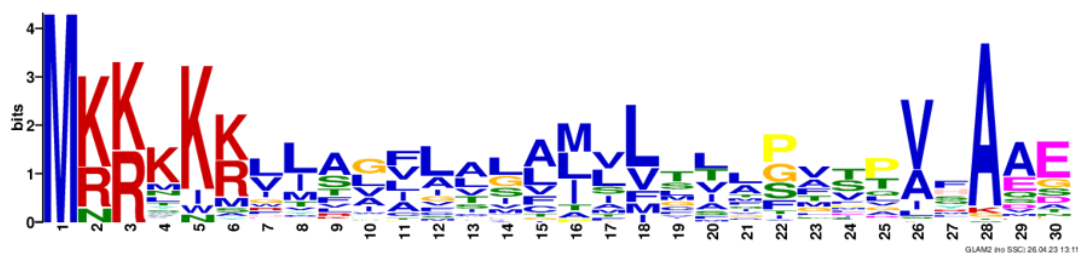


Figure 5-3. The depiction of most conserved amino acids among sec signal peptides of *R. gnavus* with logo plot.

5.4.4. Characterization of Secretion Efficiency of Signal Peptides in *R. gnavus* and *C. comes*

The signal peptides selected for characterization were given in Table 5-1. For characterization, Sequences of selected signal peptides were retrieved with additional five amino acids from downstream of the calculated cleavage site by signalP6.

Table 5-1. List of signal peptide candidates selected for secretion efficiency test in *Lachnospiraceae* species

Mimce Lab code	Prokka code	Uniprot code	Annotation	Source	Signal Peptide Sequence
ssJack-1	GMBBHPNO_02554	A0A829NNJ0	Cell surface protein Shp haem-binding domain-containing protein	SignalP6 (highest score for cleavage prediction)	MRKKQKNGMLQKMLMA AVFLLVLTVPQTARAA
ssREA2	KLDLGBH_01678 Sialidase B_signalp	A7B557	BNR/Asp-box repeat protein	Literature	MNKYKKIVSIAATAVMC VAPMSVYAQEAQT
ssREA3	KLDLGBH_03088 hypothetical protein_signalp	A0A829NZS7	Gram-positive cocci surface proteins LPX1G domain-containing protein	Literature	MYRKKLMRACTGLMV VTATANVAPVFAEAP A
ssREA9	KLDLGBH_01584	A0A829NUE0	Bacteriocin	Proteomics / Sup only	MRAKKVIVGVASAMLI SSIPVMAFQVTNG
ssREA10	KLDLGBH_01933	A7B118	AB hydrolase-1 domain-containing protein	Proteomics / Sup only	MRRYKKRISRVAVAMLF VFLIGNSNMVHAMEKTQ
ssREA11	KLDLGBH_02199	A7B6C5	Peptidase_C39_2 domain-containing protein	Proteomics / Valconplot (compared to its CL, most abundant in sup)	MRKIFKRFASLLTIFILTIM SIVPVHASENTS
ssREA12	KLDLGBH_03036	A0A2N5NIT6	LacI family transcriptional regulator	Proteomics / Sup only	MKKKVLAVLLGGCMVA GLLAG
ssREA13	KLDLGBH_03201 hypothetical protein_signalp	A0A412NDK9	NlpC/P60 domain-containing protein	Proteomics / Valconplot – (compared to its CL, second most abundant in sup)	MKMKKVRSVLLAAALTC MIVTPVMAAPDDT
pREA14	-	-	-	Synthetic	MKKKKLLALFLALAMV LTLVPVPVFAAE

The selected signal peptide candidates were cloned into the *E. coli* – Lachnospiraceae shuttle plasmid. The shuttle plasmid was amplified with PCR. (Figure 5-4) All the signal peptides except ssREA14 were obtained from genome with Gibson overhangs for cloning experiments. The ssREA14 was synthesized using assembly PCR with compatible overhangs for Gibson assembly. (Figure 5-4b and 5-4c) The signal peptides were cloned into plasmids separately at the upstream of the nanoluciferase by using Gibson assembly.

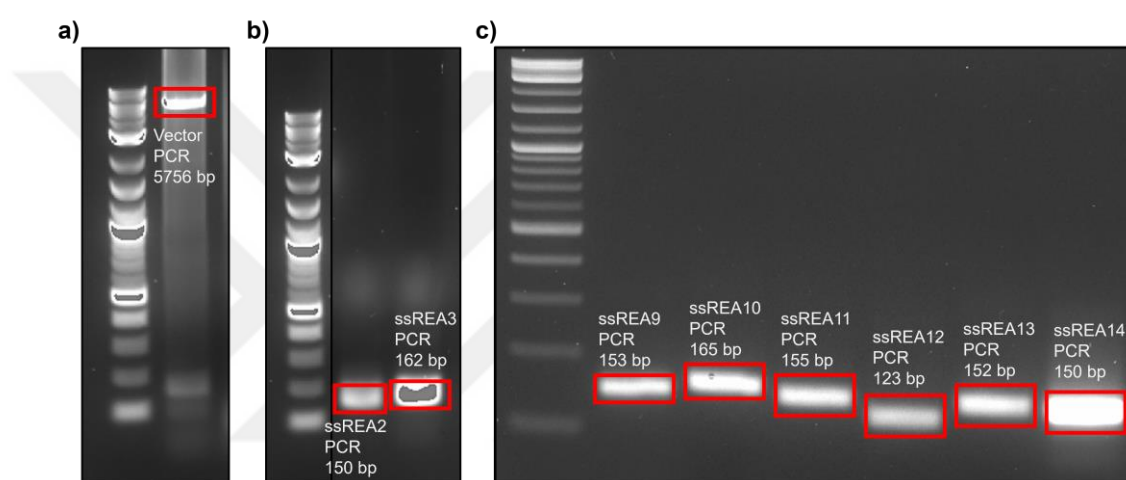


Figure 5-4. Cloning of signal peptides at the upstream of nanoluciferase in the conjugative plasmid. a) PCR amplification of the vector plasmid with gibbon overhangs. b, c) Amplification of signal peptides ssREA2 (b), ssREA3(b), ssREA9 (c), ssREA10 (c), ssREA11 (c), ssREA12 (c), ssREA13 (c), and ssREA14 (c) All DNA bands appeared at their expected sizes.

The conjugative plasmids used in the experiments contain an *E. coli* origin of replication, gram positive replicon (pCB102), catP cassette for chloramphenicol resistance, oriT for conjugation and full expression cassette inducible with aTc in which production of nanoluciferase is controlled. Nanoluciferase enzyme was utilized as the reporter to assess the secretion efficiency. (Figure 5-5)

For transferring recombinant vectors into *R. gnavus* and *C. comes*, plasmids were transformed into *E. coli* S17 λ pir strain with chemical transformation for the conjugation. Both recipient cells and *E. coli* cells carrying the conjugation were grown overnight. On the day of the conjugation experiment, *E. coli* cells were collected with centrifugation at 4500xg for 5 minutes and washed with 1x PBS supplemented with 0.05% cysteine. Subsequently, the cell pellet was transferred into the anaerobic growth chamber. Two additional PBS washes were performed with a pre-reduced solution in the chamber. Following the washing steps, recipient Lachnospiraceae cells grown anaerobically in BHIS media was mixed with *E. coli* pellet. The cell suspension was dropped onto a prereduced BHIS agar plate without any antibiotics. The bacterial lawn was grown for one day. Next day, all bacterial lawn was collected by scraping and dissolved in the prereduced cysteine-supplemented PBS. The cell suspension was spread onto BHIS agar plate containing kanamycin and thiamphenicol. Plates were incubated at 37 °C for 2 days. After the incubation, single colonies were picked and streaked onto new antibiotic-supplemented BHIS plate with X-gal and incubated for 2 days to visualize conjugant colonies. (Figure 5-5b)

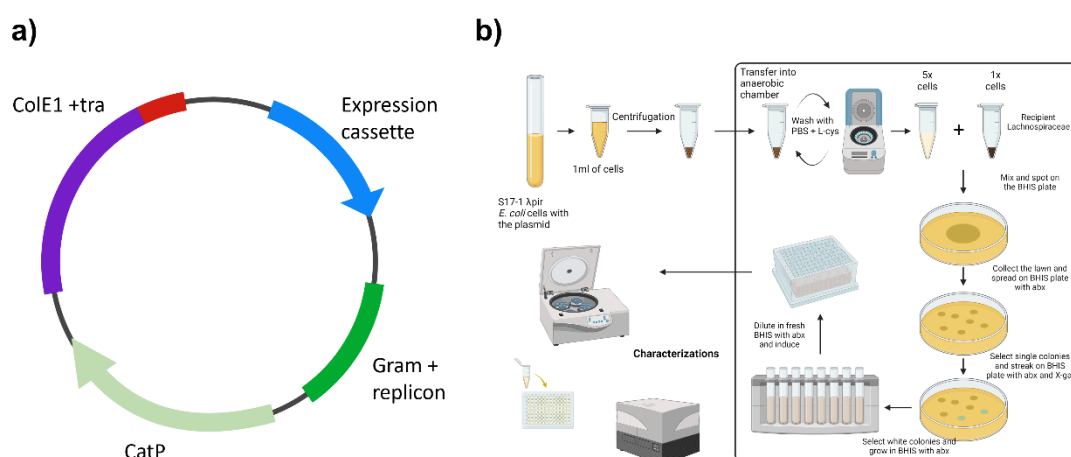
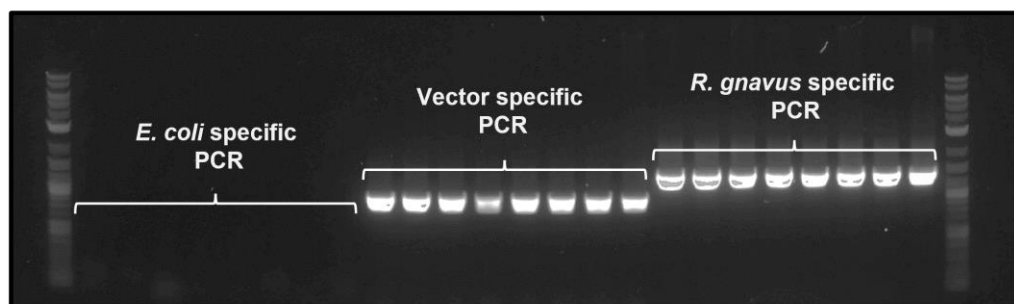


Figure 5-5. Genetic engineering of Lachnospiraceae species. a) Depiction of conjugative plasmid used to transfer DNA into recipient Lachnospiraceae. b) Schematic description of the conjugation process.

Following the conjugation, colony PCR was performed to ensure that single colonies selected from plates were recipient cells carrying the plasmids without any *E. coli* contamination. Three reactions were run for each selected colony to amplify 16S RNA region of recipient Lachnospiraceae cells, either *R. gnavus* or *C. Comes* (1.3kb), 16S RNA region of *E. coli* (1.3kb) and a region from the recombinant plasmids (800bp). Based on the PCR results, it is concluded that conjugation was successful for both *R. gnavus* and *C. comes*, and the recombinant Lachnospiraceae cells were free from *E. coli* contamination. (Figure 5-6a, 5-6b)

a)



b)

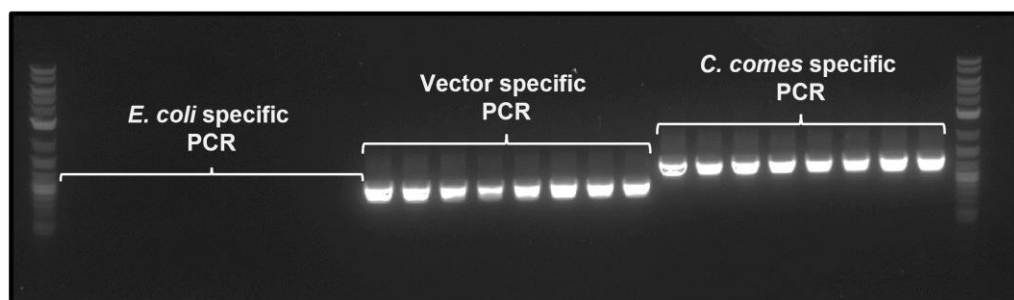


Figure 5-6. Colony PCR to verify conjugation into *R. gnavus* (a) and *C. comes* (b). The first eight lanes are products of the PCR that amplify *E. coli* specific gene. The second eight lanes are PCR products that amplify specific region from vector. The last eight lanes are products of the PCR that amplify 16S RNA region of the recipient cell. The order of each eight lanes belonging to the same PCR is ssREA2, ssREA3, ssREA9, ssREA10, ssREA11, ssREA12, ssREA14, ssREA14. All DNA bands appeared at their expected sizes.

The recombinant Lachnospiraceae cells carrying the plasmids were grown and induced to express nanoluciferase with signal peptides. To induce the protein expression, signal colonies were picked and grown overnight in the prereduced BHIS liquid media supplemented with kanamycin and thiamphenicol. The next day, cells were diluted 100-fold into fresh media with aTc and antibiotics and grown overnight again. Following the induction, cells were transferred into 96-well plate and centrifuged to obtain cell free supernatant and cell pellet. Cell pellet was suspended in 1X PBS. To measure the nanoluciferase activity, NanoGlo kit (Promega) was used. Before measurement, both cell suspension in PBS and cell free supernatant was diluted 10-fold to prevent saturation during the measurement.

Luminescence measurement revealed that extracellular nanoluciferase activity was higher than intracellular control with all peptides except ssREA12 for both *R. gnavus* and *C. comes*. The highest extracellular activity was obtained with ssJack-1, ssREA11 and ssREA14 peptides species whereas REA14 peptide had the highest ratio of supernatant-to-intracellular luciferase activity thereby which indicates the highest extracellular protein translocation efficiency for both of Lachnospiraceae. (Figure 5-7)

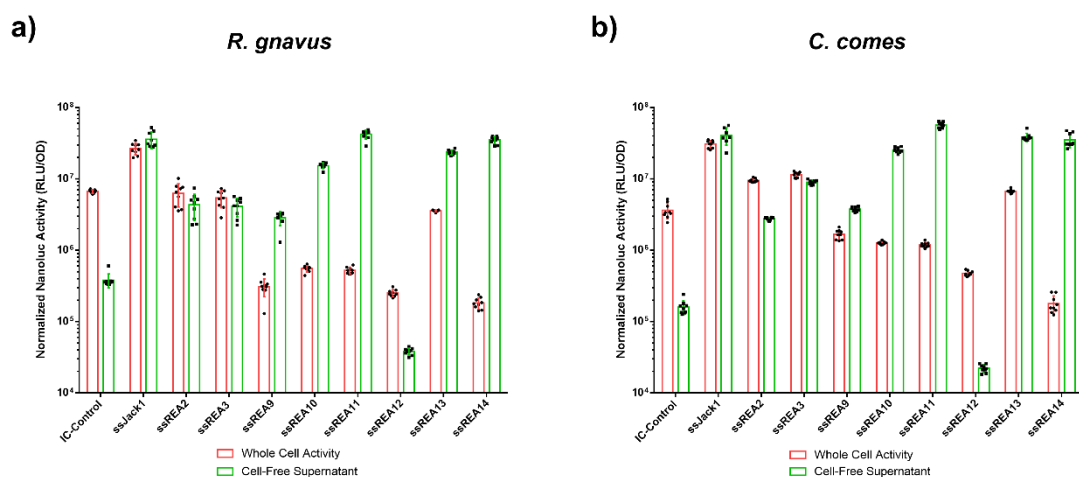


Figure 5-7. Intracellular and extracellular luciferase activity of *R. gnavus* (a) or *C. comes* (b) cells. The cells were the carrying plasmids that encode nanoluciferase fused with different signal peptides.

5.4.5. Protein Display Shuttle for *R. gnavus* and *C. comes*

A protein display shuttle was engineered and optimized for Lachnospiraceae species to express membrane bound proteins faced extracellularly. To do so, IbpA protein was redesigned to translocate proteins at the cell surface. IbpA is an extracellular membrane bound protein identified in *R. gnavus* and possesses four repeats of a superantigen domain and a sortase motif at its C-terminal. To repurpose the IbpA protein as a surface display shuttle, the protein was truncated gradually from N-terminus while keeping its native signal peptide to find minimal required domains for protein presentation at the cell surface. Four different variants were constructed progressively removing sequence from N-terminus of the IbpA protein. During the all variants, the N-terminal signal peptide was retained; meanwhile, all superantigen domains and sequence upstream of the superantigen domain were removed from protein sequence, and as a cargo protein, nanoluciferase was fused downstream of signal peptide with a flexible linker. (Figure 5-8)

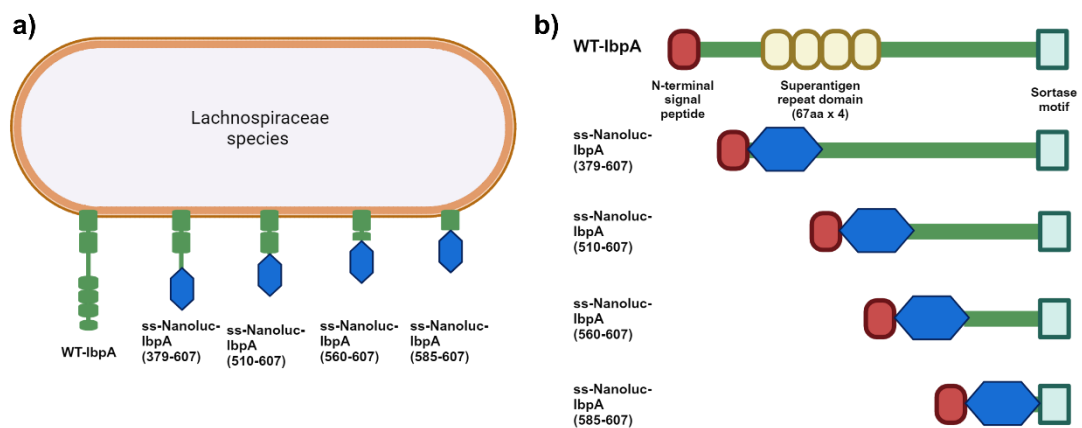


Figure 5-8. The protein display shuttle for Lachnospiraceae cells. a) The illustration of IbpA proteins utilized for surface protein display. b) Gene topology of IbpA variants designed to be tested for determining the surface display efficiency. In the design step, the N- terminus region and super antigen repeat domains were removed and nanoluciferase was used as the cargo protein for the surface display.

To clone the IbpA variants, conjugation vector with ssREA-3 signal peptide was amplified with PCR. (Figure 5-9a) All four different variants of IbpA were obtained from *R. gnavus* with compatible overhangs for Gibson assembly via PCR. (Figure 5-9b) The regions encoding the IbpA variants were cloned in the vector with Gibson assembly.

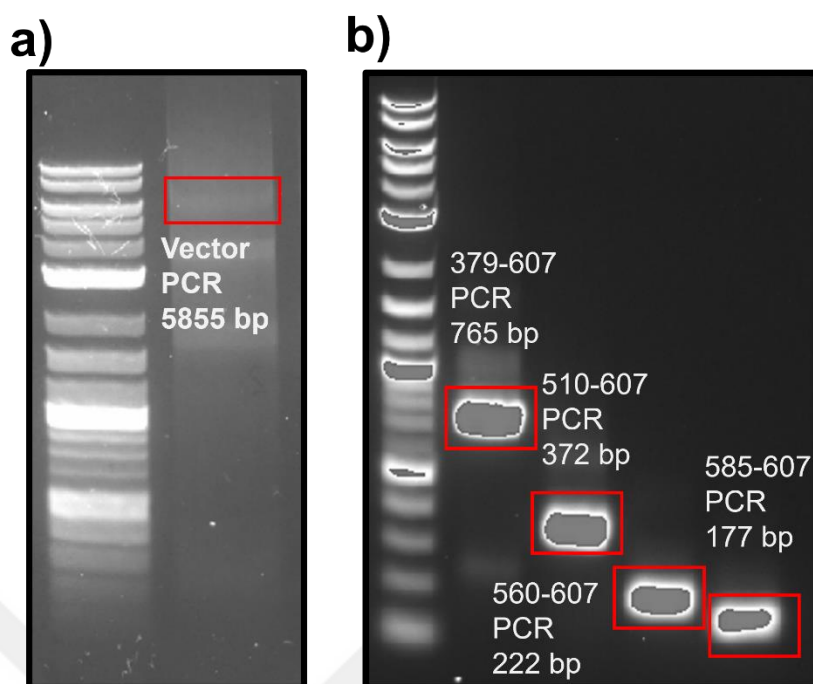


Figure 5-9. Construction of the display shuttle plasmids. a) PCR amplification of the vector plasmid with ssREA-3 nanoluciferase gene. b) Amplification of IbpA variants from *R. gnavus*. IbpA variants were cloned at the downstream of nanoluciferase. All DNA bands appeared at their expected sizes.

To test protein display, trypsin surface trimming experiment was utilized. In the trypsin surface trimming, cells are treated with trypsin to digest the accessible proteins at the extracellular space meanwhile trypsin can't access the intracellular protein due to cellular membrane. Thus, the accessible proteins can be determined by checking retained protein after the trypsinization. (Figure 5-10a)

All IbpA variants were conjugated to both *R. gnavus* and *C. comes* to test the extracellular protein display. Cells were induced and centrifuged to obtain cell pellet. Subsequently, cells were treated with trypsin for one hour at 37°C. Following the incubation, cells were washed with PBS two times and their whole cell luciferase activity was assessed same as described above and compared with untreated cells.

Whole cell luciferase activities revealed that both *R. gnavus* and *C. comes* cells carrying 379-607 and 510-607 variants could display proteins on the cell surface because cells carrying those plasmids lost their luciferase activity drastically after the trypsin assay. (Figure 5-10b, and 5-10c)

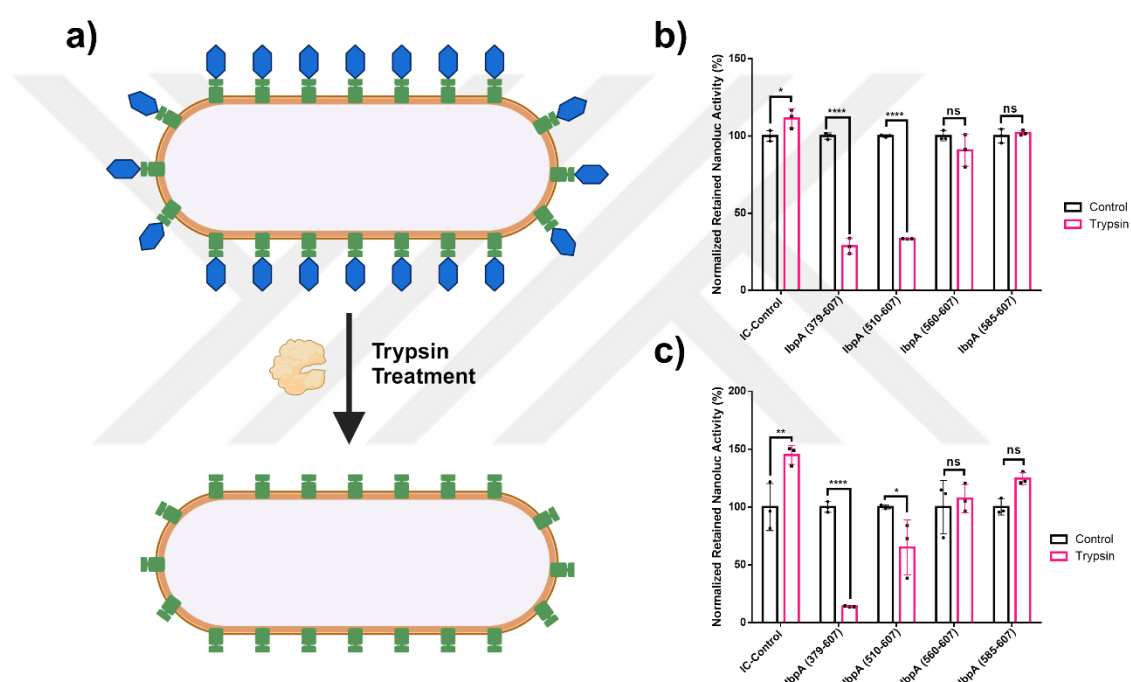


Figure 5-10. Validation of nanoluciferase surface display with trypsin surface trimming assay. a) The depiction of trypsin surface trimming assay. Cells were treated with trypsin to digest proteins at the cell surface. If nanoluciferase are displayed on the extracellular milieu, nanoluciferase will be digested by trypsin thereby decreasing the whole cell nanoluciferase activity. a, b) Retained nanoluciferase activity of *R. gnavus* (b) and *C. comes* (c) cells expressing the Ibpa variant with nanoluciferase after trypsin treatment.

5.5. Discussions

Species in Lachnospiraceae family are the native resident microbes of healthy microbiome which have beneficial intrinsic abilities thereby offering intriguing opportunities for engineering living microbial biotherapeutics. [273] Yet, there is few molecular technologies available for Lachnospiraceae species. Here, we focused on characterizing sec signal peptides for protein secretion in Lachnospiraceae species as well as designing a protein surface anchoring system. Our motivation was to develop efficient secretory protein production tools in Lachnospiraceae. Secretory protein production is the indispensable task for the *in situ* delivery of protein payloads via engineered microbes. Both efficacy and off-target toxicity can be drastically improved by secretion therapeutical proteins at the disease site. [274]–[276]

A four-pronged approach was taken to identify potential candidates signal peptides which could drive extracellular protein translocation effectively. Based on the characterization data, the signal peptides found by using the proteomic screens of secretome yielded four of the five most efficient peptides functioning both *R. gnavus* and *C. comes* whereas most the efficient peptide for protein secretion was predicted by finding the conserved regions among all signal peptides in *R. gnavus* via Glam2 tool. All other peptides which were found by either searching literature or using signal6p were less effective compared to other peptides, yet they also were functional thereby indicating that all strategies employed in this study could be utilized to predict signal peptides in other non-model organisms.

Engineering native microbes offer an advantageous platform to develop microbiome-based therapeutics and diagnostics for several reasons. Firstly, native microbes are well adapted to occupy certain ecological niche within the gut thereby facilitating

their engraftment into the preexisting microbiota. [277] The stable engraftment of engineered microbes in the gut can provide a means to continuously monitor physiological parameters and autonomously intervene the pathological conditions upon sensing. Such theranostic microbes can change the treatment options of chronic GI pathologies e.g., inflammatory bowel diseases. In addition, native gut microbes process interesting molecular pathways to transform and synthesize biomolecules such as antimicrobial peptides, modified bile and organic acids, and short chain fatty chain acids. [278]–[280] These molecular capabilities can be augmented with synthetic biology approaches for developing smart therapeutical cells, completely circumventing reconstruction these beneficial pathways in the model microbial chassis such as *E. coli*.

CHAPTER 6

Conclusion and Future Perspectives

6.1. Overview

Biopharmaceutical drugs have been changing the treatment paradigms for many indications by offering the access for novel mechanism of actions required to intervene the biology of diseases. Meanwhile, emerging technologies such as *in vitro* DNA synthesis, CRISPR-Cas gene editing system, combinatorial protein libraries and many others integrate the biological engineering principles i.e., synthetic biology in the biotechnological drug research for improving the capabilities of existing biopharmaceuticals as well as discovering the novel biopharmaceutical modalities offered by mother nature. Preliminary outcomes of this integration already have hit the clinics by offering cures for diseases previously considered intractable. Nonetheless, there is still a substantial technological gap to fulfill clinical needs for current and future diseases. In this thesis, I described a prophylactic protein-based drug candidate against SARS-CoV-2 and its variants. In addition, genetic technologies to engineer probiotic and commensal based microbial therapeutics were demonstrated.

6.2. Major Contributions to the Present Literature

The works elaborated in this thesis have made following conceptual and methodological contributions to the existing literature on living therapeutics as well as lectin-based protein drugs.

SARS-CoV-2 caused a global pandemic from its emergence in 2020 to 2023. The widespread transmission of the virus without any available medicinal interventions put a significant burden on existing healthcare thereby claiming millions of lives. Here in Chapter 2, I described a potent antiviral lectin protein, griffithsin (rGRFT), against SARS-CoV-2 and its variants that evolved during the course of pandemic. The rGRFT protein was proven to block the SARS-CoV-2 and the variants of SARS-CoV-2 which can escape the vaccine induced adaptive immunity. Besides, safety studies *in vitro* and *in vivo* models have revealed that rGRFT is well-tolerated. In addition, the recombinant expression of rGRFT was carried out in *E. coli* BL21 (DE3) strain which provided the cost-effective production process. Altogether, the work demonstrates that rGRFT can offer cheap and effective antivirals against future variants of SARS-CoV-2.

EcN is the one of most studied microorganisms for engineering therapeutical strains owing to the vast literature available on biology and genetics of *E. coli* strains as well as the century long safety record in humans. Beyond the academics works, there are commercial efforts to push engineered *EcN*-therapeutics into the clinics for several indications. In Chapter 3, I expanded the genetic toolbox for engineering *EcN* by providing a plasmid-based stable gene transfer system. For this purpose, cryptic plasmids of *EcN* were cured from cells with CRISPR-Cas system subsequently, cryptic plasmids were constructed as expression vectors to transfer recombinant DNA into the *EcN* and their stability in the cells were validated. This work provides means to fasten research on engineering *EcN* strains for therapeutical purposes by eliminating the cumbersome genome insertions for stable DNA transfer.

In Chapter 4, a synthetic secretion system was developed for EcN strains for delivery of protein-based payloads to the gut. Due to the outer membrane of the cells, protein secretion in EcN necessitates the use of secretory machineries which are hard to express in recombinant host due to their gigantic sizes. To circumvent the use of secretory machines, the outer membrane of EcN cells was transiently destabilized by knocking down the structural proteins at the periplasm thereby driving the free diffusion of recombinant proteins which had been translocated to periplasm prior. The secretion system is named as inducible leaky outer membrane-based secretion system (iLOM-SS). The iLOM-SS was proven to be modular and efficient in EcN for secretion of different proteins. This work demonstrated engineering of a prominent secretion system which could be implemented in the probiotic EcN strain and could serve as a resourceful tool for shuttling diverse proteins to the extracellular space.

Microbes residing in the gut have constant interactions with the body and affect many physiological parameters as a result of these interactions. Some of the members of this microbial community such as Lachnospiraceae species are positively correlated with overall well-being and health of the human body. Although there is vast potential for developing novel microbial therapeutics by using these microorganisms, many molecular tools to implement desired functions are lacking, which hinders engineering efforts. Here in Chapter 5, I characterized sec based protein secretion peptides for driving efficient protein translocation to extracellular space in Lachnospiraceae species. In addition, a protein display shuttle was engineered by repurposing a surface protein of *R. gnavus* for extracellular protein representation. This work provides necessary tools for protein secretion and surface display in Lachnospiraceae which can be used to augment their inherent functions for

therapeutical purposes. In addition, methodologies used in this study could serve as a blueprint to develop protein secretion strategies for other non-model microorganisms in the gut microbiota.

6.3. Future Perspectives

6.3.1. Broad Antivirals for Future Pandemics

Viral pandemics pose a serious threat to global health, with researchers estimating that a new pandemic could emerge more frequently than ever due to rapid globalization and ecological disruption. [281]–[283] These factors increase the likelihood of viral spillovers to humans which include many types of rapidly evolving zoonotic viruses with the potential to cause global pandemics. [284] Vaccines and neutralizing antibody therapies offer protection for the body against viral assaults, yet viruses accumulate evasive mutations quickly for these preventative measures thereby rendering them ineffective. [285] Additionally, neither vaccines nor antibody therapies can offer a broad-spectrum protection against a class of viruses as they are designed to target a specific epitope of a particular virus. The development of new vaccines and antibody therapies for each and every virus, including emerging and re-emerging strains, is a long and costly process. [286], [287] Therefore, there is an urgent need for pan-viral antiviral agents that target conserved regions on viruses.

Glycan structures on viral capsids have been shown to play critical roles in virus penetration into cells. [288] Therefore, glycan binding proteins/peptides can offer a unique mechanism of action that is shared among the different virus types to inhibit viral replication. Lectin proteins are ubiquitously found in all kingdoms of life which

can bind multivalently to carbohydrates with high affinities. [289] Hence, the earth biome can be mined to discover pan-virus antivirals with minimal toxicity on humans which harbors a vast number of lectins and lectin-like proteins. Furthermore, production of these lectin proteins in bacteria will drastically decrease the cost thereby which enables people to access these medications. In conclusion, lectin proteins can be a plausible addition to medical armaments against viral pandemics.

6.3.2. Engineering Probiotic and Commensal Bacteria Species for Protein Delivery

Protein-based drugs are a prominent class of therapeutics with the potential to treat a wide range of diseases. However, their full potential is often limited by the adverse events caused by their systemic administration. [290] Synthetic biology, in combination with the inherent properties of probiotic strains, provides a platform for engineering probiotic bacteria to sense the disease site for discrimination from the healthy tissues. Therefore, probiotic strains can offer a novel approach to deliver protein-based drugs to the target disease site, minimizing their interactions on other organs and systems. For instance, EcN strains can prevent colonization of Enterobacteriaceae in the inflamed gut and it can colonize the necrotic cores of tumors. [291], [292] This inherent preference of EcN can be further enhanced by adding biosensor layers to detect environmental cues associated with disease sites. Coupled with the intensive literature on recombinant protein expression in *E. coli*, EcN cells can be engineered to secrete a diverse array of proteins by the described system in this thesis iLOM-SS. [293] High capacity and modularity of iLOM-SS can serve to build genetic circuits with complex actuators modules for delivering multiple proteins at once or multiple proteins with a sequence. In addition, fast

response time of iLOM-SS offer to build protein delivery system for various protein hormones such as insulin which is required to be released as soon as possible when signal is received. Lastly, compact version of iLOM-SS makes this system wide available to other laboratories that don't have expertise in secretory protein production thereby accelerating research on EcN and other *E. coli* strains which are in the need of protein secretion.

In addition, some commensal strains found in healthy microbiota harbors unique molecular abilities in the context of therapeutics which can be further augmented by secretion of therapeutical proteins. [273] However, there are no molecular tools available to engineering recombinant proteins from these organisms. The methodologies presented in this thesis on engineering protein secretion for nonmodel Lachnospiraceae species can pay the way for development of similar tools in other nonmodel gut microbes such as Bifidobacterium species. In addition, the capabilities of Lachnospiraceae species can be enhanced by secretion of various proteins. For example, the oxygen sensitivity of these species might be mitigated by secretion of oxide mutase thereby increasing their evolutionary fitness in the inflamed gut. Increasing the fitness of butyrate producing strains of Lachnospiraceae might help to ameliorate the inflammation by restoring immune balance at the gut.

6.3.3. Engineering Native Human Commensal as Theranostic Cells

Microbes inhabiting the gut have been adapted to colonize stably at certain parts of the GI tract. In the normal situations, the stable colonization of gut by members of healthy microbiota also exerts positive effects on the human body by modulating immune system and constructing a strong resistance against pathogens and pathobionts. [294] Therefore, established communities at the gut microbiota are

robust and unaffected by minor perturbations. The robustness of microbiota can be exploited to develop long-term therapies by engineering the members of endogenous microbiota. Stably engraftment engineered microbes can reprogrammed to sense environmental clues for pathological conditions such as IBD or bleeding. [295] Such biosensor microbes can monitor the gut conditions continuously for diagnosis of disease without requiring invasive techniques, including colonoscopy. In addition, the microbes can be further augmented to compute to several molecular signatures and release active payloads for the treatments. These fully autonomous microbes with the ability to conditionally deliver active therapeutics at the disease will transform paradigm for many chronic diseases.

BIBLIOGRAPHY

- [1] M. Kesik-Brodacka, “Progress in biopharmaceutical development,” *Biotechnology and Applied Biochemistry*, vol. 65, no. 3, pp. 306–322, 2018, doi: 10.1002/bab.1617.
- [2] S. Dasani, R. Palanki, P. Menon, and S. K. Bose, “Chapter 85 - Biopharmaceuticals,” in *Translational Surgery*, A. E. M. Eltorai, J. A. Bakal, P. C. Newell, and A. J. Osband, Eds., in Handbook for Designing and Conducting Clinical and Translational Research. , Academic Press, 2023, pp. 535–538. doi: 10.1016/B978-0-323-90300-4.00056-2.
- [3] G. Walsh and E. Walsh, “Biopharmaceutical benchmarks 2022,” *Nat Biotechnol*, vol. 40, no. 12, Art. no. 12, Dec. 2022, doi: 10.1038/s41587-022-01582-x.
- [4] E. R. Mardis, “Next-Generation Sequencing Platforms,” *Annual Review of Analytical Chemistry*, vol. 6, no. 1, pp. 287–303, 2013, doi: 10.1146/annurev-anchem-062012-092628.
- [5] J. Tian, K. Ma, and I. Saaem, “Advancing high-throughput gene synthesis technology,” *Mol. BioSyst.*, vol. 5, no. 7, pp. 714–722, Jun. 2009, doi: 10.1039/B822268C.
- [6] A. H. Khan, G. J. Tye, and R. Noordin, “CRISPR-Cas9 Genome Editing Tool for the Production of Industrial Biopharmaceuticals,” *Mol Biotechnol*, vol. 62, no. 9, pp. 401–411, Sep. 2020, doi: 10.1007/s12033-020-00265-9.
- [7] X. Yan, X. Liu, C. Zhao, and G.-Q. Chen, “Applications of synthetic biology in medical and pharmaceutical fields,” *Sig Transduct Target Ther*, vol. 8, no. 1, Art. no. 1, May 2023, doi: 10.1038/s41392-023-01440-5.
- [8] X. Tan, J. H. Letendre, J. J. Collins, and W. W. Wong, “Synthetic biology in the clinic: engineering vaccines, diagnostics, and therapeutics,” *Cell*, vol. 184, no. 4, pp. 881–898, Feb. 2021, doi: 10.1016/j.cell.2021.01.017.

- [9] D. S. Dimitrov, "Therapeutic Proteins," *Methods Mol Biol*, vol. 899, pp. 1–26, 2012, doi: 10.1007/978-1-61779-921-1_1.
- [10] B. Leader, Q. J. Baca, and D. E. Golan, "Protein therapeutics: a summary and pharmacological classification," *Nat Rev Drug Discov*, vol. 7, no. 1, Art. no. 1, Jan. 2008, doi: 10.1038/nrd2399.
- [11] H. A. D. Lagassé *et al.*, "Recent advances in (therapeutic protein) drug development," *F1000Res*, vol. 6, p. 113, Feb. 2017, doi: 10.12688/f1000research.9970.1.
- [12] M. L. Chiu, D. R. Goulet, A. Teplyakov, and G. L. Gilliland, "Antibody Structure and Function: The Basis for Engineering Therapeutics," *Antibodies*, vol. 8, no. 4, Art. no. 4, Dec. 2019, doi: 10.3390/antib8040055.
- [13] R. Jefferis, K. Kato, and W. R. (Bill) Strohl, Eds., *Structure and Function of Antibodies*. MDPI - Multidisciplinary Digital Publishing Institute, 2021. doi: 10.3390/books978-3-03943-898-3.
- [14] D. N. Forthal, "Functions of Antibodies," *Microbiology spectrum*, vol. 2, no. 4, p. 1.
- [15] B. A. Keyt, R. Baliga, A. M. Sinclair, S. F. Carroll, and M. S. Peterson, "Structure, Function, and Therapeutic Use of IgM Antibodies," *Antibodies*, vol. 9, no. 4, Art. no. 4, Dec. 2020, doi: 10.3390/antib9040053.
- [16] D. Sterlin and G. Gorochov, "When Therapeutic IgA Antibodies Might Come of Age," *Pharmacology*, vol. 106, no. 1–2, pp. 9–19, Sep. 2020, doi: 10.1159/000510251.
- [17] T. Z. Yuan *et al.*, "Rapid discovery of diverse neutralizing SARS-CoV-2 antibodies from large-scale synthetic phage libraries," *mAbs*, vol. 14, no. 1, p. 2002236, Dec. 2022, doi: 10.1080/19420862.2021.2002236.
- [18] Q. Liu *et al.*, "Functional GLP-1R antibodies identified from a synthetic GPCR-focused library demonstrate potent blood glucose control," *mAbs*, vol. 13, no. 1, p. 1893425, Jan. 2021, doi: 10.1080/19420862.2021.1893425.
- [19] D. M. Mason *et al.*, "Optimization of therapeutic antibodies by predicting antigen specificity from antibody sequence via deep learning," *Nat Biomed Eng*, vol. 5, no. 6, Art. no. 6, Jun. 2021, doi: 10.1038/s41551-021-00699-9.
- [20] A. C. Hunt *et al.*, "A rapid cell-free expression and screening platform for antibody discovery," *Nat Commun*, vol. 14, no. 1, Art. no. 1, Jul. 2023, doi: 10.1038/s41467-023-38965-w.
- [21] W. Schaefer *et al.*, "Immunoglobulin domain crossover as a generic approach for the production of bispecific IgG antibodies," *Proceedings of the National Academy of Sciences*, vol. 108, no. 27, pp. 11187–11192, Jul. 2011, doi: 10.1073/pnas.1019002108.
- [22] R. Bargou *et al.*, "Tumor Regression in Cancer Patients by Very Low Doses of a T Cell–Engaging Antibody," *Science*, vol. 321, no. 5891, pp. 974–977, Aug. 2008, doi: 10.1126/science.1158545.
- [23] R. Lucchi, J. Bentanachs, and B. Oller-Salvia, "The Masking Game: Design of Activatable Antibodies and Mimetics for Selective Therapeutics and Cell Control," *ACS Cent. Sci.*, vol. 7, no. 5, pp. 724–738, May 2021, doi: 10.1021/acscentsci.0c01448.
- [24] O. Erster *et al.*, "Site-specific targeting of antibody activity in vivo mediated by disease-associated proteases," *Journal of Controlled Release*, vol. 161, no. 3, pp. 804–812, Aug. 2012, doi: 10.1016/j.jconrel.2012.05.035.

- [25] P. Holliger and P. J. Hudson, "Engineered antibody fragments and the rise of single domains," *Nat Biotechnol*, vol. 23, no. 9, Art. no. 9, Sep. 2005, doi: 10.1038/nbt1142.
- [26] D. A. Richards, "Exploring alternative antibody scaffolds: Antibody fragments and antibody mimics for targeted drug delivery," *Drug Discovery Today: Technologies*, vol. 30, pp. 35–46, Dec. 2018, doi: 10.1016/j.ddtec.2018.10.005.
- [27] R. Simeon and Z. Chen, "In vitro-engineered non-antibody protein therapeutics," *Protein & Cell*, vol. 9, no. 1, pp. 3–14, Jan. 2018, doi: 10.1007/s13238-017-0386-6.
- [28] B. Owens, "Faster, deeper, smaller—the rise of antibody-like scaffolds," *Nature Biotechnology*, vol. 35, no. 7, Art. no. 7, Jul. 2017, doi: 10.1038/nbt0717-602.
- [29] S. Rothenberger *et al.*, "The trispecific DARPin ensovibep inhibits diverse SARS-CoV-2 variants," *Nat Biotechnol*, vol. 40, no. 12, Art. no. 12, Dec. 2022, doi: 10.1038/s41587-022-01382-3.
- [30] M. Gebauer and A. Skerra, "Engineered Protein Scaffolds as Next-Generation Therapeutics," *Annual Review of Pharmacology and Toxicology*, vol. 60, no. 1, pp. 391–415, 2020, doi: 10.1146/annurev-pharmtox-010818-021118.
- [31] "Cytokines in the balance," *Nat Immunol*, vol. 20, no. 12, Art. no. 12, Dec. 2019, doi: 10.1038/s41590-019-0557-0.
- [32] P. A. Kołodziejski *et al.*, "The Role of Peptide Hormones Discovered in the 21st Century in the Regulation of Adipose Tissue Functions," *Genes (Basel)*, vol. 12, no. 5, p. 756, May 2021, doi: 10.3390/genes12050756.
- [33] T. B. Kjeldsen *et al.*, "Molecular Engineering of Insulin Icodec, the First Acylated Insulin Analog for Once-Weekly Administration in Humans," *J. Med. Chem.*, vol. 64, no. 13, pp. 8942–8950, Jul. 2021, doi: 10.1021/acs.jmedchem.1c00257.
- [34] R. Hernandez, J. Pöder, K. M. LaPorte, and T. R. Malek, "Engineering IL-2 for immunotherapy of autoimmunity and cancer," *Nat Rev Immunol*, vol. 22, no. 10, Art. no. 10, Oct. 2022, doi: 10.1038/s41577-022-00680-w.
- [35] M. Evans, P. M. Schumm-Draeger, J. Vora, and A. B. King, "A review of modern insulin analogue pharmacokinetic and pharmacodynamic profiles in type 2 diabetes: improvements and limitations," *Diabetes Obes Metab*, vol. 13, no. 8, pp. 677–684, Aug. 2011, doi: 10.1111/j.1463-1326.2011.01395.x.
- [36] I. B. Hirsch, "Insulin Analogues," *New England Journal of Medicine*, vol. 352, no. 2, pp. 174–183, Jan. 2005, doi: 10.1056/NEJMra040832.
- [37] S. Franzè, F. Cilurzo, and P. Minghetti, "Insulin Biosimilars: The Impact on Rapid-Acting Analogue-Based Therapy," *BioDrugs*, vol. 29, no. 2, pp. 113–121, Apr. 2015, doi: 10.1007/s40259-015-0121-x.
- [38] J. Scheller, A. Chalaris, D. Schmidt-Arras, and S. Rose-John, "The pro- and anti-inflammatory properties of the cytokine interleukin-6," *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research*, vol. 1813, no. 5, pp. 878–888, May 2011, doi: 10.1016/j.bbamcr.2011.01.034.
- [39] J. Deckers *et al.*, "Engineering cytokine therapeutics," *Nat Rev Bioeng*, vol. 1, no. 4, Art. no. 4, Apr. 2023, doi: 10.1038/s44222-023-00030-y.
- [40] J. Emmerich *et al.*, "Abstract 1744: STK-012, an alpha/beta selective IL-2 mutein for the activation of the antigen-activated T cells in solid tumor," *Cancer Research*, vol. 81, no. 13_Supplement, p. 1744, Jul. 2021, doi: 10.1158/1538-7445.AM2021-1744.

- [41] R. A. Saxton *et al.*, “Structure-based decoupling of the pro- and anti-inflammatory functions of interleukin-10,” *Science*, vol. 371, no. 6535, p. eabc8433, Mar. 2021, doi: 10.1126/science.abc8433.
- [42] E. Shemesh *et al.*, “Enzyme replacement and substrate reduction therapy for Gaucher disease,” *Cochrane Database of Systematic Reviews*, no. 3, 2015, doi: 10.1002/14651858.CD010324.pub2.
- [43] P. Harmatz *et al.*, “The effect of galsulfase enzyme replacement therapy on the growth of patients with mucopolysaccharidosis VI (Maroteaux-Lamy syndrome),” *Molecular Genetics and Metabolism*, vol. 122, no. 1, pp. 107–112, Sep. 2017, doi: 10.1016/j.ymgme.2017.03.008.
- [44] U. R. R. Somaraju and A. Solis-Moya, “Pancreatic enzyme replacement therapy for people with cystic fibrosis,” *Cochrane Database of Systematic Reviews*, no. 8, 2020, doi: 10.1002/14651858.CD008227.pub4.
- [45] W. Kim *et al.*, “Trends in Enzyme Therapy for Phenylketonuria,” *Molecular Therapy*, vol. 10, no. 2, pp. 220–224, Aug. 2004, doi: 10.1016/j.ymthe.2004.05.001.
- [46] T. A. Triplett *et al.*, “Reversal of indoleamine 2,3-dioxygenase-mediated cancer immune suppression by systemic kynurenine depletion with a therapeutic enzyme,” *Nat Biotechnol*, vol. 36, no. 8, Art. no. 8, Sep. 2018, doi: 10.1038/nbt.4180.
- [47] M. C. Hochberg, “New directions in symptomatic therapy for patients with osteoarthritis and rheumatoid arthritis,” *Seminars in Arthritis and Rheumatism*, vol. 32, no. 3, Supplement, pp. 4–14, Dec. 2002, doi: 10.1053/sarh.2002.37215.
- [48] E. Peters, S. Heemskerk, R. Masereeuw, and P. Pickkers, “Alkaline Phosphatase: A Possible Treatment for Sepsis-Associated Acute Kidney Injury in Critically Ill Patients,” *American Journal of Kidney Diseases*, vol. 63, no. 6, pp. 1038–1048, Jun. 2014, doi: 10.1053/j.ajkd.2013.11.027.
- [49] T. Zhang *et al.*, “Clinical Potential of an Enzyme-based Novel Therapy for Cocaine Overdose,” *Sci Rep*, vol. 7, no. 1, Art. no. 1, Nov. 2017, doi: 10.1038/s41598-017-14105-5.
- [50] S. M. Dostal, Y. Fang, J. C. Guerrette, T. C. Scanlon, and K. E. Griswold, “Genetically Enhanced Lysozyme Evades a Pathogen Derived Inhibitory Protein,” *ACS Chem. Biol.*, vol. 10, no. 4, pp. 1110–1117, Apr. 2015, doi: 10.1021/cb500976y.
- [51] J. C. Jackson, S. P. Duffy, K. R. Hess, and R. A. Mehl, “Improving Nature’s Enzyme Active Site with Genetically Encoded Unnatural Amino Acids,” *J. Am. Chem. Soc.*, vol. 128, no. 34, pp. 11124–11127, Aug. 2006, doi: 10.1021/ja061099y.
- [52] B. N. J. Watson, J. A. Steens, R. H. J. Staals, E. R. Westra, and S. van Houte, “Coevolution between bacterial CRISPR-Cas systems and their bacteriophages,” *Cell Host & Microbe*, vol. 29, no. 5, pp. 715–725, May 2021, doi: 10.1016/j.chom.2021.03.018.
- [53] Z. Zhao, P. Shang, P. Mohanraju, and N. Geijsen, “Prime editing: advances and therapeutic applications,” *Trends in Biotechnology*, vol. 41, no. 8, pp. 1000–1012, Aug. 2023, doi: 10.1016/j.tibtech.2023.03.004.
- [54] T. Li *et al.*, “CRISPR/Cas9 therapeutics: progress and prospects,” *Sig Transduct Target Ther*, vol. 8, no. 1, Art. no. 1, Jan. 2023, doi: 10.1038/s41392-023-01309-7.

- [55] H.-K. Liao *et al.*, “In Vivo Target Gene Activation via CRISPR/Cas9-Mediated Trans-epigenetic Modulation,” *Cell*, vol. 171, no. 7, pp. 1495–1507.e15, Dec. 2017, doi: 10.1016/j.cell.2017.10.025.
- [56] C. A. Mitchell, K. Ramessar, and B. R. O’Keefe, “Antiviral lectins: Selective inhibitors of viral entry,” *Antiviral Res*, vol. 142, pp. 37–54, Jun. 2017, doi: 10.1016/j.antiviral.2017.03.007.
- [57] S. C. Gordts *et al.*, “NICTABA and UDA, two GlcNAc-binding lectins with unique antiviral activity profiles,” *Journal of Antimicrobial Chemotherapy*, vol. 70, no. 6, pp. 1674–1685, Jun. 2015, doi: 10.1093/jac/dkv034.
- [58] D. S. Dimitrov, “Virus entry: molecular mechanisms and biomedical applications,” *Nat Rev Microbiol*, vol. 2, no. 2, Art. no. 2, Feb. 2004, doi: 10.1038/nrmicro817.
- [59] M. Mazalovska and J. C. Kouokam, “Plant-Derived Lectins as Potential Cancer Therapeutics and Diagnostic Tools,” *Biomed Res Int*, vol. 2020, p. 1631394, May 2020, doi: 10.1155/2020/1631394.
- [60] A. Gupta, J. L. Andresen, R. S. Manan, and R. Langer, “Nucleic acid delivery for therapeutic applications,” *Advanced Drug Delivery Reviews*, vol. 178, p. 113834, Nov. 2021, doi: 10.1016/j.addr.2021.113834.
- [61] R. S. Finkel *et al.*, “Nusinersen versus Sham Control in Infantile-Onset Spinal Muscular Atrophy,” *New England Journal of Medicine*, vol. 377, no. 18, pp. 1723–1732, Nov. 2017, doi: 10.1056/NEJMoa1702752.
- [62] I. Urits *et al.*, “A Review of Patisiran (ONPATTRO®) for the Treatment of Polyneuropathy in People with Hereditary Transthyretin Amyloidosis,” *Neurol Ther*, vol. 9, no. 2, pp. 301–315, Dec. 2020, doi: 10.1007/s40120-020-00208-1.
- [63] Y. Wang, G. Sheng, S. Juranek, T. Tuschl, and D. J. Patel, “Structure of the guide-strand-containing argonaute silencing complex,” *Nature*, vol. 456, no. 7219, Art. no. 7219, Nov. 2008, doi: 10.1038/nature07315.
- [64] P. S. Kowalski, A. Rudra, L. Miao, and D. G. Anderson, “Delivering the Messenger: Advances in Technologies for Therapeutic mRNA Delivery,” *Molecular Therapy*, vol. 27, no. 4, pp. 710–728, Apr. 2019, doi: 10.1016/j.ymthe.2019.02.012.
- [65] M. A. Islam, E. K. G. Reesor, Y. Xu, H. R. Zope, B. R. Zetter, and J. Shi, “Biomaterials for mRNA delivery,” *Biomater. Sci.*, vol. 3, no. 12, pp. 1519–1533, Nov. 2015, doi: 10.1039/C5BM00198F.
- [66] U. Sahin, K. Karikó, and Ö. Türeci, “mRNA-based therapeutics — developing a new class of drugs,” *Nat Rev Drug Discov*, vol. 13, no. 10, Art. no. 10, Oct. 2014, doi: 10.1038/nrd4278.
- [67] A. N. Kuhn *et al.*, “Phosphorothioate cap analogs increase stability and translational efficiency of RNA vaccines in immature dendritic cells and induce superior immune responses in vivo,” *Gene Ther*, vol. 17, no. 8, Art. no. 8, Aug. 2010, doi: 10.1038/gt.2010.52.
- [68] J. M. Henderson *et al.*, “Cap 1 Messenger RNA Synthesis with Co-transcriptional CleanCap® Analog by In Vitro Transcription,” *Current Protocols*, vol. 1, no. 2, p. e39, 2021, doi: 10.1002/cpz1.39.
- [69] K. Karikó *et al.*, “Incorporation of Pseudouridine Into mRNA Yields Superior Nonimmunogenic Vector With Increased Translational Capacity and Biological Stability,” *Molecular Therapy*, vol. 16, no. 11, pp. 1833–1840, Nov. 2008, doi: 10.1038/mt.2008.200.

- [70] K. Karikó, H. Muramatsu, J. Ludwig, and D. Weissman, "Generating the optimal mRNA for therapy: HPLC purification eliminates immune activation and improves translation of nucleoside-modified, protein-encoding mRNA," *Nucleic Acids Research*, vol. 39, no. 21, p. e142, Nov. 2011, doi: 10.1093/nar/gkr695.
- [71] X. Hou, T. Zaks, R. Langer, and Y. Dong, "Lipid nanoparticles for mRNA delivery," *Nat Rev Mater*, vol. 6, no. 12, Art. no. 12, Dec. 2021, doi: 10.1038/s41578-021-00358-0.
- [72] A. Hınçer, R. E. Ahan, E. Aras, and U. Ö. Ş. Şeker, "Making the Next Generation of Therapeutics: mRNA Meets Synthetic Biology," *ACS Synth. Biol.*, vol. 12, no. 9, pp. 2505–2515, Sep. 2023, doi: 10.1021/acssynbio.3c00253.
- [73] R. Jain *et al.*, "MicroRNAs Enable mRNA Therapeutics to Selectively Program Cancer Cells to Self-Destruct," *Nucleic Acid Ther*, vol. 28, no. 5, pp. 285–296, Oct. 2018, doi: 10.1089/nat.2018.0734.
- [74] R. Kedmi *et al.*, "A modular platform for targeted RNAi therapeutics," *Nature Nanotech*, vol. 13, no. 3, Art. no. 3, Mar. 2018, doi: 10.1038/s41565-017-0043-5.
- [75] S. Mc Cafferty *et al.*, "In Vivo Validation of a Reversible Small Molecule-Based Switch for Synthetic Self-Amplifying mRNA Regulation," *Molecular Therapy*, vol. 29, no. 3, pp. 1164–1173, Mar. 2021, doi: 10.1016/j.ymthe.2020.11.010.
- [76] R. Chen *et al.*, "Engineering circular RNA for enhanced protein production," *Nat Biotechnol*, vol. 41, no. 2, Art. no. 2, Feb. 2023, doi: 10.1038/s41587-022-01393-0.
- [77] A. R. Lucas, Ed., *Viruses as Therapeutics: Methods and Protocols*, vol. 2225. in *Methods in Molecular Biology*, vol. 2225. New York, NY: Springer US, 2021. doi: 10.1007/978-1-0716-1012-1.
- [78] H. R. Gelderblom, "Structure and Classification of Viruses," in *Medical Microbiology*, 4th ed., S. Baron, Ed., Galveston (TX): University of Texas Medical Branch at Galveston, 1996. Accessed: Oct. 19, 2023. [Online]. Available: <http://www.ncbi.nlm.nih.gov/books/NBK8174/>
- [79] T. Travieso, J. Li, S. Mahesh, J. D. F. R. E. Mello, and M. Blasi, "The use of viral vectors in vaccine development," *npj Vaccines*, vol. 7, no. 1, Art. no. 1, Jul. 2022, doi: 10.1038/s41541-022-00503-y.
- [80] N. Tatsis *et al.*, "Chimpanzee-origin adenovirus vectors as vaccine carriers," *Gene Ther*, vol. 13, no. 5, Art. no. 5, Mar. 2006, doi: 10.1038/sj.gt.3302675.
- [81] M. Robert-Guroff, "Replicating and non-replicating viral vectors for vaccine development," *Current Opinion in Biotechnology*, vol. 18, no. 6, pp. 546–556, Dec. 2007, doi: 10.1016/j.copbio.2007.10.010.
- [82] D. Zhou, A. Cun, Y. Li, Z. Xiang, and H. C. J. Ertl, "A Chimpanzee-Origin Adenovirus Vector Expressing the Rabies Virus Glycoprotein as an Oral Vaccine against Inhalation Infection with Rabies Virus," *Molecular Therapy*, vol. 14, no. 5, pp. 662–672, Nov. 2006, doi: 10.1016/j.ymthe.2006.03.027.
- [83] L. Si *et al.*, "Generation of a live attenuated influenza A vaccine by proteolysis targeting," *Nat Biotechnol*, vol. 40, no. 9, Art. no. 9, Sep. 2022, doi: 10.1038/s41587-022-01381-4.
- [84] S. M. Hoy, "Onasemnogene Apeparvovec: First Global Approval," *Drugs*, vol. 79, no. 11, pp. 1255–1262, Jul. 2019, doi: 10.1007/s40265-019-01162-5.

- [85] C. Kang and L. J. Scott, “Voretigene Neparvovec: A Review in RPE65 Mutation-Associated Inherited Retinal Dystrophy,” *Mol Diagn Ther*, vol. 24, no. 4, pp. 487–495, Aug. 2020, doi: 10.1007/s40291-020-00475-6.
- [86] M. Poorebrahim *et al.*, “Production of CAR T-cells by GMP-grade lentiviral vectors: latest advances and future prospects,” *Critical Reviews in Clinical Laboratory Sciences*, vol. 56, no. 6, pp. 393–419, Aug. 2019, doi: 10.1080/10408363.2019.1633512.
- [87] H. J. Wagner, W. Weber, and M. Fussenegger, “Synthetic Biology: Emerging Concepts to Design and Advance Adeno-Associated Viral Vectors for Gene Therapy,” *Advanced Science*, vol. 8, no. 9, p. 2004018, 2021, doi: 10.1002/advs.202004018.
- [88] S. J. Russell, K.-W. Peng, and J. C. Bell, “Oncolytic virotherapy,” *Nat Biotechnol*, vol. 30, no. 7, Art. no. 7, Jul. 2012, doi: 10.1038/nbt.2287.
- [89] L. Nissim *et al.*, “Synthetic RNA-Based Immunomodulatory Gene Circuits for Cancer Immunotherapy,” *Cell*, vol. 171, no. 5, pp. 1138–1150.e15, Nov. 2017, doi: 10.1016/j.cell.2017.09.049.
- [90] H. K. Chung *et al.*, “A compact synthetic pathway rewires cancer signaling to therapeutic effector release,” *Science*, vol. 364, no. 6439, p. eaat6982, May 2019, doi: 10.1126/science.aat6982.
- [91] R. A. Edwards and F. Rohwer, “Viral metagenomics,” *Nat Rev Microbiol*, vol. 3, no. 6, Art. no. 6, Jun. 2005, doi: 10.1038/nrmicro1163.
- [92] N. Chanishvili, “Chapter 1 - Phage Therapy—History from Twort and d’Herelle Through Soviet Experience to Current Approaches,” in *Advances in Virus Research*, vol. 83, M. Łobocka and W. Szybalski, Eds., in Bacteriophages, Part B, vol. 83, Academic Press, 2012, pp. 3–40. doi: 10.1016/B978-0-12-394438-2.00001-3.
- [93] D. S. Schmitt, S. D. Siegel, and K. Selle, “Applications of designer phage encoding recombinant gene payloads,” *Trends in Biotechnology*, Oct. 2023, doi: 10.1016/j.tibtech.2023.09.008.
- [94] B. R. Lenneman, J. Fernbach, M. J. Loessner, T. K. Lu, and S. Kilcher, “Enhancing phage therapy through synthetic biology and genome engineering,” *Current Opinion in Biotechnology*, vol. 68, pp. 151–159, Apr. 2021, doi: 10.1016/j.copbio.2020.11.003.
- [95] K. Yehl *et al.*, “Engineering Phage Host-Range and Suppressing Bacterial Resistance through Phage Tail Fiber Mutagenesis,” *Cell*, vol. 179, no. 2, pp. 459–469.e9, Oct. 2019, doi: 10.1016/j.cell.2019.09.015.
- [96] J. Du *et al.*, “Enhancing bacteriophage therapeutics through in situ production and release of heterologous antimicrobial effectors,” *Nat Commun*, vol. 14, no. 1, Art. no. 1, Jul. 2023, doi: 10.1038/s41467-023-39612-0.
- [97] S. Meile *et al.*, “Engineered reporter phages for detection of Escherichia coli, Enterococcus, and Klebsiella in urine,” *Nat Commun*, vol. 14, no. 1, Art. no. 1, Jul. 2023, doi: 10.1038/s41467-023-39863-x.
- [98] M. A. Fischbach, J. A. Bluestone, and W. A. Lim, “Cell-Based Therapeutics: The Next Pillar of Medicine,” *Science Translational Medicine*, vol. 5, no. 179, pp. 179ps7–179ps7, Apr. 2013, doi: 10.1126/scitranslmed.3005568.
- [99] M. Mansouri and M. Fussenegger, “Therapeutic cell engineering: designing programmable synthetic genetic circuits in mammalian cells,” *Protein & Cell*, vol. 13, no. 7, pp. 476–489, Jul. 2022, doi: 10.1007/s13238-021-00876-1.

- [100] J. Zhang *et al.*, “Non-viral, specifically targeted CAR-T cells achieve high safety and efficacy in B-NHL,” *Nature*, vol. 609, no. 7926, Art. no. 7926, Sep. 2022, doi: 10.1038/s41586-022-05140-y.
- [101] S. Kaczanowska *et al.*, “Genetically engineered myeloid cells rebalance the core immune suppression program in metastasis,” *Cell*, vol. 184, no. 8, pp. 2033–2052.e21, Apr. 2021, doi: 10.1016/j.cell.2021.02.048.
- [102] J. H. Cho, A. Okuma, K. Sofjan, S. Lee, J. J. Collins, and W. W. Wong, “Engineering advanced logic and distributed computing in human CAR immune cells,” *Nat Commun*, vol. 12, no. 1, Art. no. 1, Feb. 2021, doi: 10.1038/s41467-021-21078-7.
- [103] C. T. Ellebrecht *et al.*, “Reengineering chimeric antigen receptor T cells for targeted therapy of autoimmune disease,” *Science*, vol. 353, no. 6295, pp. 179–184, Jul. 2016, doi: 10.1126/science.aaf6756.
- [104] K. Schreeb *et al.*, “Study Design: Human Leukocyte Antigen Class I Molecule A*02-Chimeric Antigen Receptor Regulatory T Cells in Renal Transplantation,” *Kidney International Reports*, vol. 7, no. 6, pp. 1258–1267, Jun. 2022, doi: 10.1016/j.ekir.2022.03.030.
- [105] C. Kemmer, M. Gitzinger, M. Daoud-El Baba, V. Djonov, J. Stelling, and M. Fussenegger, “Self-sufficient control of urate homeostasis in mice by a synthetic circuit,” *Nat Biotechnol*, vol. 28, no. 4, Art. no. 4, Apr. 2010, doi: 10.1038/nbt.1617.
- [106] P. Saxena, G. Charpin-El Hamri, M. Folcher, H. Zulewski, and M. Fussenegger, “Synthetic gene network restoring endogenous pituitary–thyroid feedback control in experimental Graves’ disease,” *Proceedings of the National Academy of Sciences*, vol. 113, no. 5, pp. 1244–1249, Feb. 2016, doi: 10.1073/pnas.1514383113.
- [107] M. Xie *et al.*, “ β -cell-mimetic designer cells provide closed-loop glycemic control,” *Science*, vol. 354, no. 6317, pp. 1296–1301, Dec. 2016, doi: 10.1126/science.aaf4006.
- [108] J. L. Sonnenburg and E. D. Sonnenburg, “Vulnerability of the industrialized microbiota,” *Science*, vol. 366, no. 6464, p. eaaw9255, Oct. 2019, doi: 10.1126/science.aaw9255.
- [109] M. R. Charbonneau, V. M. Isabella, N. Li, and C. B. Kurtz, “Developing a new class of engineered live bacterial therapeutics to treat human diseases,” *Nat Commun*, vol. 11, no. 1, Art. no. 1, Apr. 2020, doi: 10.1038/s41467-020-15508-1.
- [110] J. P. Lynch, L. Goers, and C. F. Lesser, “Emerging strategies for engineering Escherichia coli Nissle 1917-based therapeutics,” *Trends in Pharmacological Sciences*, vol. 43, no. 9, pp. 772–786, Sep. 2022, doi: 10.1016/j.tips.2022.02.002.
- [111] T. V. Plavec and A. Berlec, “Safety Aspects of Genetically Modified Lactic Acid Bacteria,” *Microorganisms*, vol. 8, no. 2, Feb. 2020, doi: 10.3390/microorganisms8020297.
- [112] K. A. Hedin, V. Kruse, R. Vazquez-Urbe, and M. O. A. Sommer, “Biocontainment strategies for in vivo applications of *Saccharomyces boulardii*,” *Frontiers in Bioengineering and Biotechnology*, vol. 11, 2023, Accessed: Nov. 14, 2023. [Online]. Available: <https://www.frontiersin.org/articles/10.3389/fbioe.2023.1136095>

- [113] T. V. Plavec and A. Berlec, "Engineering of lactic acid bacteria for delivery of therapeutic proteins and peptides," *Appl Microbiol Biotechnol*, vol. 103, no. 5, pp. 2053–2066, Mar. 2019, doi: 10.1007/s00253-019-09628-y.
- [114] L. Steidler *et al.*, "Treatment of Murine Colitis by *Lactococcus lactis* Secreting Interleukin-10," *Science*, vol. 289, no. 5483, pp. 1352–1355, Aug. 2000, doi: 10.1126/science.289.5483.1352.
- [115] N. Mao, A. Cubillos-Ruiz, D. E. Cameron, and J. J. Collins, "Probiotic strains detect and suppress cholera in mice," *Science Translational Medicine*, vol. 10, no. 445, p. eaao2586, Jun. 2018, doi: 10.1126/scitranslmed.aao2586.
- [116] J. Vockley *et al.*, "Efficacy and safety of a synthetic biotic for treatment of phenylketonuria: a phase 2 clinical trial," *Nat Metab*, vol. 5, no. 10, Art. no. 10, Oct. 2023, doi: 10.1038/s42255-023-00897-6.
- [117] C. B. Kurtz *et al.*, "An engineered *E. coli* Nissle improves hyperammonemia and survival in mice and shows dose-dependent exposure in healthy humans," *Science Translational Medicine*, vol. 11, no. 475, p. eaau7975, Jan. 2019, doi: 10.1126/scitranslmed.aau7975.
- [118] D. Lubkowitz *et al.*, "An engineered bacterial therapeutic lowers urinary oxalate in preclinical models and in silico simulations of enteric hyperoxaluria," *Molecular Systems Biology*, vol. 18, no. 3, p. e10539, Mar. 2022, doi: 10.15252/msb.202110539.
- [119] C. R. Gurbatri *et al.*, "Engineered probiotics for local tumor delivery of checkpoint blockade nanobodies," *Science Translational Medicine*, vol. 12, no. 530, p. eaax0876, Feb. 2020, doi: 10.1126/scitranslmed.aax0876.
- [120] J. P. Lynch *et al.*, "Engineered *Escherichia coli* for the in situ secretion of therapeutic nanobodies in the gut," *Cell Host & Microbe*, vol. 31, no. 4, pp. 634–649.e8, Apr. 2023, doi: 10.1016/j.chom.2023.03.007.
- [121] P. Praveschotinunt, A. M. Duraj-Thatte, I. Gelfat, F. Bahl, D. B. Chou, and N. S. Joshi, "Engineered *E. coli* Nissle 1917 for the delivery of matrix-tethered therapeutic domains to the gut," *Nat Commun*, vol. 10, no. 1, Art. no. 1, Dec. 2019, doi: 10.1038/s41467-019-13336-6.
- [122] F. C. P. Tiago *et al.*, "Adhesion to the yeast cell surface as a mechanism for trapping pathogenic bacteria by *Saccharomyces* probiotics," *Journal of Medical Microbiology*, vol. 61, no. 9, pp. 1194–1207, 2012, doi: 10.1099/jmm.0.042283-0.
- [123] D. Durmusoglu *et al.*, "In Situ Biomanufacturing of Small Molecules in the Mammalian Gut by Probiotic *Saccharomyces boulardii*," *ACS Synth. Biol.*, vol. 10, no. 5, pp. 1039–1052, May 2021, doi: 10.1021/acssynbio.0c00562.
- [124] K. Chen *et al.*, "A probiotic yeast-based immunotherapy against *Clostridioides difficile* infection," *Science Translational Medicine*, vol. 12, no. 567, p. eaax4905, Oct. 2020, doi: 10.1126/scitranslmed.aax4905.
- [125] R. E. Ahan *et al.*, "A Highly Potent SARS-CoV-2 Blocking Lectin Protein," *ACS Infect. Dis.*, vol. 8, no. 7, pp. 1253–1264, Jul. 2022, doi: 10.1021/acsinfecdis.2c00006.
- [126] A. E. Gorbalenya *et al.*, "The species Severe acute respiratory syndrome-related coronavirus: classifying 2019-nCoV and naming it SARS-CoV-2," *Nat Microbiol*, vol. 5, no. 4, Art. no. 4, Apr. 2020, doi: 10.1038/s41564-020-0695-z.
- [127] D. Mercatelli and F. M. Giorgi, "Geographic and Genomic Distribution of SARS-CoV-2 Mutations," *Frontiers in Microbiology*, vol. 11, 2020, Accessed:

- Oct. 12, 2023. [Online]. Available: <https://www.frontiersin.org/articles/10.3389/fmicb.2020.01800>
- [128] A. L. Rasmussen and S. V. Popescu, "SARS-CoV-2 transmission without symptoms," *Science*, vol. 371, no. 6535, pp. 1206–1207, Mar. 2021, doi: 10.1126/science.abf9569.
 - [129] J. F.-W. Chan *et al.*, "Genomic characterization of the 2019 novel human-pathogenic coronavirus isolated from a patient with atypical pneumonia after visiting Wuhan," *Emerging Microbes & Infections*, vol. 9, no. 1, pp. 221–236, Jan. 2020, doi: 10.1080/22221751.2020.1719902.
 - [130] R. Yan, Y. Zhang, Y. Li, L. Xia, Y. Guo, and Q. Zhou, "Structural basis for the recognition of SARS-CoV-2 by full-length human ACE2," *Science*, vol. 367, no. 6485, pp. 1444–1448, Mar. 2020, doi: 10.1126/science.abb2762.
 - [131] J. Lan *et al.*, "Structure of the SARS-CoV-2 spike receptor-binding domain bound to the ACE2 receptor," *Nature*, vol. 581, no. 7807, Art. no. 7807, May 2020, doi: 10.1038/s41586-020-2180-5.
 - [132] D. J. Benton *et al.*, "Receptor binding and priming of the spike protein of SARS-CoV-2 for membrane fusion," *Nature*, vol. 588, no. 7837, Art. no. 7837, Dec. 2020, doi: 10.1038/s41586-020-2772-0.
 - [133] M. Hoffmann *et al.*, "SARS-CoV-2 Cell Entry Depends on ACE2 and TMPRSS2 and Is Blocked by a Clinically Proven Protease Inhibitor," *Cell*, vol. 181, no. 2, pp. 271–280.e8, Apr. 2020, doi: 10.1016/j.cell.2020.02.052.
 - [134] J. Shang *et al.*, "Cell entry mechanisms of SARS-CoV-2," *Proceedings of the National Academy of Sciences*, vol. 117, no. 21, pp. 11727–11734, May 2020, doi: 10.1073/pnas.2003138117.
 - [135] E. E. Walsh *et al.*, "Safety and Immunogenicity of Two RNA-Based Covid-19 Vaccine Candidates," *N Engl J Med*, vol. 383, no. 25, pp. 2439–2450, Dec. 2020, doi: 10.1056/NEJMoa2027906.
 - [136] K. S. Corbett *et al.*, "Evaluation of the mRNA-1273 Vaccine against SARS-CoV-2 in Nonhuman Primates," *N Engl J Med*, vol. 383, no. 16, pp. 1544–1555, Oct. 2020, doi: 10.1056/NEJMoa2024671.
 - [137] R. Bos *et al.*, "Ad26 vector-based COVID-19 vaccine encoding a prefusion-stabilized SARS-CoV-2 Spike immunogen induces potent humoral and cellular immune responses," *npj Vaccines*, vol. 5, no. 1, Art. no. 1, Sep. 2020, doi: 10.1038/s41541-020-00243-x.
 - [138] P. M. Folegatti *et al.*, "Safety and immunogenicity of the ChAdOx1 nCoV-19 vaccine against SARS-CoV-2: a preliminary report of a phase 1/2, single-blind, randomised controlled trial," *The Lancet*, vol. 396, no. 10249, pp. 467–478, Aug. 2020, doi: 10.1016/S0140-6736(20)31604-4.
 - [139] J.-H. Tian *et al.*, "SARS-CoV-2 spike glycoprotein vaccine candidate NVX-CoV2373 immunogenicity in baboons and protection in mice," *Nat Commun*, vol. 12, no. 1, Art. no. 1, Jan. 2021, doi: 10.1038/s41467-020-20653-8.
 - [140] "Repurposed Antiviral Drugs for Covid-19 — Interim WHO Solidarity Trial Results," *N Engl J Med*, vol. 384, no. 6, pp. 497–511, Feb. 2021, doi: 10.1056/NEJMoa2023184.
 - [141] A. Baum *et al.*, "Antibody cocktail to SARS-CoV-2 spike protein prevents rapid mutational escape seen with individual antibodies," *Science*, vol. 369, no. 6506, pp. 1014–1018, Aug. 2020, doi: 10.1126/science.abd0831.
 - [142] P. Chen *et al.*, "SARS-CoV-2 Neutralizing Antibody LY-CoV555 in Outpatients with Covid-19," *N Engl J Med*, vol. 384, no. 3, pp. 229–237, Jan. 2021, doi: 10.1056/NEJMoa2029849.

- [143] G. Cerutti *et al.*, “Potent SARS-CoV-2 neutralizing antibodies directed against spike N-terminal domain target a single supersite,” *Cell Host & Microbe*, vol. 29, no. 5, pp. 819–833.e7, May 2021, doi: 10.1016/j.chom.2021.03.005.
- [144] S. Jiang, X. Zhang, Y. Yang, P. J. Hotez, and L. Du, “Neutralizing antibodies for the treatment of COVID-19,” *Nat Biomed Eng*, vol. 4, no. 12, Art. no. 12, Dec. 2020, doi: 10.1038/s41551-020-00660-2.
- [145] S. S. Abdool Karim and T. de Oliveira, “New SARS-CoV-2 Variants — Clinical, Public Health, and Vaccine Implications,” *N Engl J Med*, vol. 384, no. 19, pp. 1866–1868, May 2021, doi: 10.1056/NEJMc2100362.
- [146] N. G. Davies *et al.*, “Estimated transmissibility and impact of SARS-CoV-2 lineage B.1.1.7 in England,” *Science*, vol. 372, no. 6538, p. eabg3055, Apr. 2021, doi: 10.1126/science.abg3055.
- [147] L. Liu *et al.*, “Striking antibody evasion manifested by the Omicron variant of SARS-CoV-2,” *Nature*, vol. 602, no. 7898, Art. no. 7898, Feb. 2022, doi: 10.1038/s41586-021-04388-0.
- [148] K. D. McCormick, J. L. Jacobs, and J. W. Mellors, “The emerging plasticity of SARS-CoV-2,” *Science*, vol. 371, no. 6536, pp. 1306–1308, Mar. 2021, doi: 10.1126/science.abg4493.
- [149] K. B. Alexandre *et al.*, “Binding of the Mannose-Specific Lectin, Griffithsin, to HIV-1 gp120 Exposes the CD4-Binding Site,” *Journal of Virology*, vol. 85, no. 17, pp. 9039–9050, Sep. 2011, doi: 10.1128/jvi.02675-10.
- [150] K. Levendosky *et al.*, “Griffithsin and Carrageenan Combination To Target Herpes Simplex Virus 2 and Human Papillomavirus,” *Antimicrobial Agents and Chemotherapy*, vol. 59, no. 12, pp. 7290–7298, Nov. 2015, doi: 10.1128/aac.01816-15.
- [151] B. R. O’Keefe *et al.*, “Broad-Spectrum In Vitro Activity and In Vivo Efficacy of the Antiviral Protein Griffithsin against Emerging Viruses of the Family Coronaviridae,” *Journal of Virology*, vol. 84, no. 5, pp. 2511–2521, Mar. 2010, doi: 10.1128/jvi.02322-09.
- [152] J. K. Millet *et al.*, “Middle East respiratory syndrome coronavirus infection is inhibited by griffithsin,” *Antiviral Research*, vol. 133, pp. 1–8, Sep. 2016, doi: 10.1016/j.antiviral.2016.07.011.
- [153] S. Lusvardi and C. A. Bewley, “Griffithsin: An Antiviral Lectin with Outstanding Therapeutic Potential,” *Viruses*, vol. 8, no. 10, p. 296, Oct. 2016, doi: 10.3390/v8100296.
- [154] A. Micsonai *et al.*, “BeStSel: a web server for accurate protein secondary structure prediction and fold recognition from the circular dichroism spectra,” *Nucleic Acids Research*, vol. 46, no. W1, pp. W315–W322, Jul. 2018, doi: 10.1093/nar/gky497.
- [155] A. Hanifnezhad *et al.*, “Characterization of local SARS-CoV-2 isolates and pathogenicity in IFNAR^{-/-} mice,” *Heliyon*, vol. 6, no. 9, p. e05116, Sep. 2020, doi: 10.1016/j.heliyon.2020.e05116.
- [156] S. I. Hindawi *et al.*, “Inactivation of Middle East respiratory syndrome-coronavirus in human plasma using amotosalen and ultraviolet A light,” *Transfusion*, vol. 58, no. 1, pp. 52–59, 2018, doi: 10.1111/trf.14422.
- [157] B. R. O’Keefe *et al.*, “Broad-Spectrum In Vitro Activity and In Vivo Efficacy of the Antiviral Protein Griffithsin against Emerging Viruses of the Family Coronaviridae,” *Journal of Virology*, vol. 84, no. 5, pp. 2511–2521, Mar. 2010, doi: 10.1128/jvi.02322-09.

- [158] N. C. Kyriakidis, A. López-Cortés, E. V. González, A. B. Grimaldos, and E. O. Prado, “SARS-CoV-2 vaccines strategies: a comprehensive review of phase 3 candidates,” *npj Vaccines*, vol. 6, no. 1, Art. no. 1, Feb. 2021, doi: 10.1038/s41541-021-00292-w.
- [159] E. J. Anderson *et al.*, “Safety and Immunogenicity of SARS-CoV-2 mRNA-1273 Vaccine in Older Adults,” *N Engl J Med*, vol. 383, no. 25, pp. 2427–2438, Dec. 2020, doi: 10.1056/NEJMoa2028436.
- [160] F. P. Polack *et al.*, “Safety and Efficacy of the BNT162b2 mRNA Covid-19 Vaccine,” *N Engl J Med*, vol. 383, no. 27, pp. 2603–2615, Dec. 2020, doi: 10.1056/NEJMoa2034577.
- [161] M. Voysey *et al.*, “Safety and efficacy of the ChAdOx1 nCoV-19 vaccine (AZD1222) against SARS-CoV-2: an interim analysis of four randomised controlled trials in Brazil, South Africa, and the UK,” *The Lancet*, vol. 397, no. 10269, pp. 99–111, Jan. 2021, doi: 10.1016/S0140-6736(20)32661-1.
- [162] J. Sadoff *et al.*, “Interim Results of a Phase 1–2a Trial of Ad26.COV2.S Covid-19 Vaccine,” *N Engl J Med*, vol. 384, no. 19, pp. 1824–1835, May 2021, doi: 10.1056/NEJMoa2034201.
- [163] Y. Zhang *et al.*, “Safety, tolerability, and immunogenicity of an inactivated SARS-CoV-2 vaccine in healthy adults aged 18–59 years: a randomised, double-blind, placebo-controlled, phase 1/2 clinical trial,” *The Lancet Infectious Diseases*, vol. 21, no. 2, pp. 181–192, Feb. 2021, doi: 10.1016/S1473-3099(20)30843-4.
- [164] M. Hoffmann *et al.*, “SARS-CoV-2 variants B.1.351 and P.1 escape from neutralizing antibodies,” *Cell*, vol. 184, no. 9, pp. 2384–2393.e12, Apr. 2021, doi: 10.1016/j.cell.2021.03.036.
- [165] F. Amanat *et al.*, “A serological assay to detect SARS-CoV-2 seroconversion in humans,” *Nat Med*, vol. 26, no. 7, Art. no. 7, Jul. 2020, doi: 10.1038/s41591-020-0913-5.
- [166] O. C. Grant, D. Montgomery, K. Ito, and R. J. Woods, “Analysis of the SARS-CoV-2 spike protein glycan shield reveals implications for immune recognition,” *Sci Rep*, vol. 10, no. 1, Art. no. 1, Sep. 2020, doi: 10.1038/s41598-020-71748-7.
- [167] A. C. Walls, Y.-J. Park, M. A. Tortorici, A. Wall, A. T. McGuire, and D. Veasler, “Structure, Function, and Antigenicity of the SARS-CoV-2 Spike Glycoprotein,” *Cell*, vol. 183, no. 6, p. 1735, Dec. 2020, doi: 10.1016/j.cell.2020.11.032.
- [168] N. E. Ziolkowska *et al.*, “Domain-Swapped Structure of the Potent Antiviral Protein Griffithsin and Its Mode of Carbohydrate Binding,” *Structure*, vol. 14, no. 7, pp. 1127–1135, Jul. 2006, doi: 10.1016/j.str.2006.05.017.
- [169] J. Verma and N. Subbarao, “A comparative study of human betacoronavirus spike proteins: structure, function and therapeutics,” *Arch Virol*, vol. 166, no. 3, pp. 697–714, Mar. 2021, doi: 10.1007/s00705-021-04961-y.
- [170] W. Xu, M. Wang, D. Yu, and X. Zhang, “Variations in SARS-CoV-2 Spike Protein Cell Epitopes and Glycosylation Profiles During Global Transmission Course of COVID-19,” *Frontiers in Immunology*, vol. 11, 2020, Accessed: Oct. 12, 2023. [Online]. Available: <https://www.frontiersin.org/articles/10.3389/fimmu.2020.565278>
- [171] L. Casalino *et al.*, “Beyond Shielding: The Roles of Glycans in the SARS-CoV-2 Spike Protein,” *ACS Cent. Sci.*, vol. 6, no. 10, pp. 1722–1734, Oct. 2020, doi: 10.1021/acscentsci.0c01056.

- [172] J. C. Kouokam *et al.*, “Investigation of Griffithsin’s Interactions with Human Cells Confirms Its Outstanding Safety and Efficacy Profile as a Microbicide Candidate,” *PLOS ONE*, vol. 6, no. 8, p. e22635, Aug. 2011, doi: 10.1371/journal.pone.0022635.
- [173] L. Girard *et al.*, “Impact of the griffithsin anti-HIV microbicide and placebo gels on the rectal mucosal proteome and microbiome in non-human primates,” *Sci Rep*, vol. 8, no. 1, Art. no. 1, May 2018, doi: 10.1038/s41598-018-26313-8.
- [174] N. Teleshova *et al.*, “Results of a phase 1, randomized, placebo-controlled first-in-human trial of griffithsin formulated in a carrageenan vaginal gel,” *PLOS ONE*, vol. 17, no. 1, p. e0261775, Jan. 2022, doi: 10.1371/journal.pone.0261775.
- [175] J. S. Decker, R. Menacho-Melgar, and M. D. Lynch, “Low-Cost, Large-Scale Production of the Anti-viral Lectin Griffithsin,” *Frontiers in Bioengineering and Biotechnology*, vol. 8, 2020, Accessed: Oct. 12, 2023. [Online]. Available: <https://www.frontiersin.org/articles/10.3389/fbioe.2020.01020>
- [176] L. F. Kramzer *et al.*, “Preformulation Characterization of Griffithsin, a Biopharmaceutical Candidate for HIV Prevention,” *AAPS PharmSciTech*, vol. 22, no. 3, p. 83, Feb. 2021, doi: 10.1208/s12249-021-01931-0.
- [177] U. Sonnenborn, “Escherichia coli strain Nissle 1917—from bench to bedside and back: history of a special Escherichia coli strain with probiotic properties,” *FEMS Microbiology Letters*, vol. 363, no. 19, p. fnw212, Oct. 2016, doi: 10.1093/femsle/fnw212.
- [178] U. Sonnenborn and J. Schulze, “The non-pathogenic Escherichia coli strain Nissle 1917 – features of a versatile probiotic,” *Microbial Ecology in Health and Disease*, vol. 21, no. 3–4, pp. 122–158, Jan. 2009, doi: 10.3109/08910600903444267.
- [179] L. Grozdanov *et al.*, “Analysis of the Genome Structure of the Nonpathogenic Probiotic Escherichia coli Strain Nissle 1917,” *Journal of Bacteriology*, vol. 186, no. 16, pp. 5432–5441, Aug. 2004, doi: 10.1128/jb.186.16.5432-5441.2004.
- [180] L. Grozdanov *et al.*, “A Single Nucleotide Exchange in the wzy Gene Is Responsible for the Semirough O6 Lipopolysaccharide Phenotype and Serum Sensitivity of Escherichia coli Strain Nissle 1917,” *Journal of Bacteriology*, vol. 184, no. 21, pp. 5912–5925, Nov. 2002, doi: 10.1128/jb.184.21.5912-5925.2002.
- [181] S. Dubbert, B. Klinkert, M. Schimiczek, T. M. Wassenaar, and R. von Büнау, “No genotoxicity is detectable for Escherichia coli strain Nissle 1917 by standard in vitro and in vivo tests,” *European Journal of Microbiology and Immunology*, vol. 10, no. 1, pp. 11–19, Mar. 2020, doi: 10.1556/1886.2019.00025.
- [182] W. Kruis, E. Schütz, P. Fric, B. Fixa, G. Judmaier, and M. Stolte, “Double-blind comparison of an oral Escherichia coli preparation and mesalazine in maintaining remission of ulcerative colitis,” *Alimentary Pharmacology & Therapeutics*, vol. 11, no. 5, pp. 853–858, 1997, doi: 10.1046/j.1365-2036.1997.00225.x.
- [183] J. Henker, S. Müller, M. W. Laass, A. Schreiner, and J. Schulze, “Probiotic Escherichia coli Nissle 1917 (EcN) for Successful Remission Maintenance of Ulcerative Colitis in Children and Adolescents: an Open-Label Pilot Study,” *Z*

- Gastroenterol*, vol. 46, no. 09, pp. 874–875, Sep. 2008, doi: 10.1055/s-2008-1027463.
- [184] M. Möllenbrink and E. Bruckschen, “Behandlung der chronischen Obstipation mit physiologischen Escherichia-coli-Bakterien. Ergebnisse einer klinischen Studie zur Wirksamkeit und Verträglichkeit der mikrobiologischen Therapie mit dem E.-coli-Stamm Nissle 1917 (Mutaflor®),” *Med Klin*, vol. 89, no. 11, pp. 587–593, 1994.
 - [185] J. Henker *et al.*, “The probiotic Escherichia coli strain Nissle 1917 (EcN) stops acute diarrhoea in infants and toddlers,” *Eur J Pediatr*, vol. 166, no. 4, pp. 311–318, Apr. 2007, doi: 10.1007/s00431-007-0419-x.
 - [186] T. Danino *et al.*, “Programmable probiotics for detection of cancer in urine,” *Sci Transl Med*, vol. 7, no. 289, p. 289ra84, May 2015, doi: 10.1126/scitranslmed.aaa3519.
 - [187] R. L. Vincent *et al.*, “Probiotic-guided CAR-T cells for solid tumor targeting,” *Science*, vol. 382, no. 6667, pp. 211–218, Oct. 2023, doi: 10.1126/science.add7034.
 - [188] M. T. Nelson *et al.*, “Characterization of an engineered live bacterial therapeutic for the treatment of phenylketonuria in a human gut-on-a-chip,” *Nat Commun*, vol. 12, no. 1, Art. no. 1, May 2021, doi: 10.1038/s41467-021-23072-5.
 - [189] M. R. Charbonneau, V. M. Isabella, N. Li, and C. B. Kurtz, “Developing a new class of engineered live bacterial therapeutics to treat human diseases,” *Nat Commun*, vol. 11, no. 1, Art. no. 1, Apr. 2020, doi: 10.1038/s41467-020-15508-1.
 - [190] D. S. Leventhal *et al.*, “Immunotherapy with engineered bacteria by targeting the STING pathway for anti-tumor immunity,” *Nat Commun*, vol. 11, no. 1, Art. no. 1, Jun. 2020, doi: 10.1038/s41467-020-16602-0.
 - [191] D. Lubkowitz *et al.*, “An engineered bacterial therapeutic lowers urinary oxalate in preclinical models and in silico simulations of enteric hyperoxaluria,” *Molecular systems biology*, vol. 18, no. 3, p. e10539, 2022.
 - [192] F. Hayes, “Toxins-Antitoxins: Plasmid Maintenance, Programmed Cell Death, and Cell Cycle Arrest,” *Science*, vol. 301, no. 5639, pp. 1496–1499, Sep. 2003, doi: 10.1126/science.1088157.
 - [193] C. W. Kang, H. G. Lim, J. Yang, M. H. Noh, S. W. Seo, and G. Y. Jung, “Synthetic auxotrophs for stable and tunable maintenance of plasmid copy number,” *Metabolic Engineering*, vol. 48, pp. 121–128, Jul. 2018, doi: 10.1016/j.ymben.2018.05.020.
 - [194] G. Blum-Oehler *et al.*, “Development of strain-specific PCR reactions for the detection of the probiotic Escherichia coli strain Nissle 1917 in fecal samples,” *Research in Microbiology*, vol. 154, no. 1, pp. 59–66, Jan. 2003, doi: 10.1016/S0923-2508(02)00007-4.
 - [195] H. S. Zainuddin, Y. Bai, and T. J. Mansell, “CRISPR-based curing and analysis of metabolic burden of cryptic plasmids in Escherichia coli Nissle 1917,” *Engineering in Life Sciences*, vol. 19, no. 6, pp. 478–485, 2019, doi: 10.1002/elsc.201900003.
 - [196] Y. Jiang, B. Chen, C. Duan, B. Sun, J. Yang, and S. Yang, “Multigene Editing in the Escherichia coli Genome via the CRISPR-Cas9 System,” *Applied and Environmental Microbiology*, vol. 81, no. 7, pp. 2506–2514, Apr. 2015, doi: 10.1128/AEM.04023-14.

- [197] B. F. Cress *et al.*, “CRISPathBrick: Modular Combinatorial Assembly of Type II-A CRISPR Arrays for dCas9-Mediated Multiplex Transcriptional Repression in *E. coli*,” *ACS Synth. Biol.*, vol. 4, no. 9, pp. 987–1000, Sep. 2015, doi: 10.1021/acssynbio.5b00012.
- [198] A. Kan, I. Gelfat, S. Emani, P. Praveschotinunt, and N. S. Joshi, “Plasmid Vectors for in Vivo Selection-Free Use with the Probiotic *E. coli* Nissle 1917,” *ACS Synth. Biol.*, vol. 10, no. 1, pp. 94–106, Jan. 2021, doi: 10.1021/acssynbio.0c00466.
- [199] L. Cui, A. Vigouroux, F. Rousset, H. Varet, V. Khanna, and D. Bikard, “A CRISPRi screen in *E. coli* reveals sequence-specific toxicity of dCas9,” *Nat Commun*, vol. 9, no. 1, Art. no. 1, May 2018, doi: 10.1038/s41467-018-04209-5.
- [200] M. Schultz *et al.*, “Green fluorescent protein for detection of the probiotic microorganism *Escherichia coli* strain Nissle 1917 (EcN) in vivo,” *Journal of Microbiological Methods*, vol. 61, no. 3, pp. 389–398, Jun. 2005, doi: 10.1016/j.mimet.2005.01.007.
- [201] V. M. Isabella *et al.*, “Development of a synthetic live bacterial therapeutic for the human metabolic disease phenylketonuria,” *Nat Biotechnol*, vol. 36, no. 9, Art. no. 9, Oct. 2018, doi: 10.1038/nbt.4222.
- [202] F. R. Cianfanelli, O. Cunrath, and D. Bumann, “Efficient dual-negative selection for bacterial genome editing,” *BMC Microbiology*, vol. 20, no. 1, p. 129, May 2020, doi: 10.1186/s12866-020-01819-2.
- [203] R. E. Ahan, C. E. Ozcelik, I. N. Cagil, and U. O. S. Seker, “A Synthetic Protein Secretion System for Living Bacterial Therapeutics.” *bioRxiv*, p. 2023.06.14.544856, Jun. 15, 2023. doi: 10.1101/2023.06.14.544856.
- [204] N. Aggarwal, S. Kitano, G. R. Y. Puah, S. Kittelmann, I. Y. Hwang, and M. W. Chang, “Microbiome and Human Health: Current Understanding, Engineering, and Enabling Technologies,” *Chem. Rev.*, vol. 123, no. 1, pp. 31–72, Jan. 2023, doi: 10.1021/acs.chemrev.2c00431.
- [205] C. B. Kurtz *et al.*, “An engineered *E. coli* Nissle improves hyperammonemia and survival in mice and shows dose-dependent exposure in healthy humans,” *Science Translational Medicine*, vol. 11, no. 475, p. eaau7975, Jan. 2019, doi: 10.1126/scitranslmed.aau7975.
- [206] I. Y. Hwang *et al.*, “Engineered probiotic *Escherichia coli* can eliminate and prevent *Pseudomonas aeruginosa* gut infection in animal models,” *Nat Commun*, vol. 8, no. 1, Art. no. 1, Apr. 2017, doi: 10.1038/ncomms15028.
- [207] D. T. Riglar *et al.*, “Engineered bacteria can function in the mammalian gut long-term as live diagnostics of inflammation,” *Nat Biotechnol*, vol. 35, no. 7, Art. no. 7, Jul. 2017, doi: 10.1038/nbt.3879.
- [208] A. Cubillos-Ruiz, M. A. Alcantar, N. M. Donghia, P. Cárdenas, J. Avila-Pacheco, and J. J. Collins, “An engineered live biotherapeutic for the prevention of antibiotic-induced dysbiosis,” *Nat. Biomed. Eng*, vol. 6, no. 7, Art. no. 7, Jul. 2022, doi: 10.1038/s41551-022-00871-9.
- [209] E. Koh *et al.*, “Engineering probiotics to inhibit *Clostridioides difficile* infection by dynamic regulation of intestinal metabolism,” *Nat Commun*, vol. 13, no. 1, Art. no. 1, Jul. 2022, doi: 10.1038/s41467-022-31334-z.
- [210] R. E. Ahan, B. M. Kirpat, B. Saltepe, and U. Ö. Ş. Şeker, “A Self-Actuated Cellular Protein Delivery Machine,” *ACS Synth. Biol.*, vol. 8, no. 4, pp. 686–696, Apr. 2019, doi: 10.1021/acssynbio.9b00062.

- [211] T. Harimoto *et al.*, “A programmable encapsulation system improves delivery of therapeutic bacteria in mice,” *Nat Biotechnol*, vol. 40, no. 8, Art. no. 8, Aug. 2022, doi: 10.1038/s41587-022-01244-y.
- [212] T. Chien *et al.*, “Enhancing the tropism of bacteria via genetically programmed biosensors,” *Nat Biomed Eng*, vol. 6, no. 1, Art. no. 1, Jan. 2022, doi: 10.1038/s41551-021-00772-3.
- [213] T. R. D. Costa *et al.*, “Secretion systems in Gram-negative bacteria: structural and mechanistic insights,” *Nat Rev Microbiol*, vol. 13, no. 6, Art. no. 6, Jun. 2015, doi: 10.1038/nrmicro3456.
- [214] R. E. Dalbey and A. Kuhn, “Protein Traffic in Gram-negative bacteria – how exported and secreted proteins find their way,” *FEMS Microbiology Reviews*, vol. 36, no. 6, pp. 1023–1045, Nov. 2012, doi: 10.1111/j.1574-6976.2012.00327.x.
- [215] I. E. Brodsky *et al.*, “A Yersinia Effector Protein Promotes Virulence by Preventing Inflammasome Recognition of the Type III Secretion System,” *Cell Host & Microbe*, vol. 7, no. 5, pp. 376–387, May 2010, doi: 10.1016/j.chom.2010.04.009.
- [216] E. A. Miao *et al.*, “Innate immune detection of the type III secretion apparatus through the NLRC4 inflammasome,” *Proceedings of the National Academy of Sciences*, vol. 107, no. 7, pp. 3076–3080, Feb. 2010, doi: 10.1073/pnas.0913087107.
- [217] K. Kanonenberg, O. Spitz, I. N. Erenburg, T. Beer, and L. Schmitt, “Type I secretion system—it takes three and a substrate,” *FEMS Microbiology Letters*, vol. 365, no. 11, p. fny094, Jun. 2018, doi: 10.1093/femsle/fny094.
- [218] S. Wagner, I. Grin, S. Malmshemer, N. Singh, C. E. Torres-Vargas, and S. Westerhausen, “Bacterial type III secretion systems: a complex device for the delivery of bacterial effector proteins into eukaryotic host cells,” *FEMS Microbiology Letters*, vol. 365, no. 19, p. fny201, Oct. 2018, doi: 10.1093/femsle/fny201.
- [219] A. Sturm *et al.*, “The Cost of Virulence: Retarded Growth of Salmonella Typhimurium Cells Expressing Type III Secretion System 1,” *PLOS Pathogens*, vol. 7, no. 7, p. e1002143, Jul. 2011, doi: 10.1371/journal.ppat.1002143.
- [220] L. A. Burdette, S. A. Leach, H. T. Wong, and D. Tullman-Ercek, “Developing Gram-negative bacteria for the secretion of heterologous proteins,” *Microbial Cell Factories*, vol. 17, no. 1, p. 196, Dec. 2018, doi: 10.1186/s12934-018-1041-5.
- [221] C. A. Green *et al.*, “Engineering the flagellar type III secretion system: improving capacity for secretion of recombinant protein,” *Microbial Cell Factories*, vol. 18, no. 1, p. 10, Jan. 2019, doi: 10.1186/s12934-019-1058-4.
- [222] D. Büttner, “Protein Export According to Schedule: Architecture, Assembly, and Regulation of Type III Secretion Systems from Plant- and Animal-Pathogenic Bacteria,” *Microbiology and Molecular Biology Reviews*, vol. 76, no. 2, pp. 262–310, Jun. 2012, doi: 10.1128/mmbr.05017-11.
- [223] P. J. Bakkes, S. Jenewein, S. H. J. Smits, I. B. Holland, and L. Schmitt, “The Rate of Folding Dictates Substrate Secretion by the Escherichia coli Hemolysin Type 1 Secretion System,” *J Biol Chem*, vol. 285, no. 52, pp. 40573–40580, Dec. 2010, doi: 10.1074/jbc.M110.173658.
- [224] D. Gilzer, M. Schreiner, and H. H. Niemann, “Direct interaction of a chaperone-bound type III secretion substrate with the export gate,” *Nat*

- Commun*, vol. 13, no. 1, Art. no. 1, Jun. 2022, doi: 10.1038/s41467-022-30487-1.
- [225] T. Heel, G. F. Vogel, A. Lammirato, R. Schneider, and B. Auer, “FlgM as a Secretion Moiety for the Development of an Inducible Type III Secretion System,” *PLOS ONE*, vol. 8, no. 3, p. e59034, Mar. 2013, doi: 10.1371/journal.pone.0059034.
 - [226] M. O. Din *et al.*, “Synchronized cycles of bacterial lysis for in vivo delivery,” *Nature*, vol. 536, no. 7614, Art. no. 7614, Aug. 2016, doi: 10.1038/nature18930.
 - [227] C. R. Gurbatri *et al.*, “Engineered probiotics for local tumor delivery of checkpoint blockade nanobodies,” *Science Translational Medicine*, vol. 12, no. 530, p. eaax0876, Feb. 2020, doi: 10.1126/scitranslmed.aax0876.
 - [228] D. E. Cameron and J. J. Collins, “Tunable protein degradation in bacteria,” *Nat Biotechnol*, vol. 32, no. 12, Art. no. 12, Dec. 2014, doi: 10.1038/nbt.3053.
 - [229] J. Park, S. Bae, and J.-S. Kim, “Cas-Designer: a web-based tool for choice of CRISPR-Cas9 target sites,” *Bioinformatics*, vol. 31, no. 24, pp. 4014–4016, Dec. 2015, doi: 10.1093/bioinformatics/btv537.
 - [230] E. Kalyoncu, R. E. Ahan, C. E. Ozcelik, and U. O. S. Seker, “Genetic Logic Gates Enable Patterning of Amyloid Nanofibers,” *Advanced Materials*, vol. 31, no. 39, p. 1902888, 2019, doi: 10.1002/adma.201902888.
 - [231] K. N.-M. Daeffler *et al.*, “Engineering bacterial thiosulfate and tetrathionate sensors for detecting gut inflammation,” *Molecular Systems Biology*, vol. 13, no. 4, p. 923, Apr. 2017, doi: 10.15252/msb.20167416.
 - [232] A. Prindle, P. Samayoa, I. Razinkov, T. Danino, L. S. Tsimring, and J. Hasty, “A sensing array of radically coupled genetic ‘biopixels,’” *Nature*, vol. 481, no. 7379, Art. no. 7379, Jan. 2012, doi: 10.1038/nature10722.
 - [233] M. D’Huyvetter *et al.*, “¹³¹I-labeled Anti-HER2 Camelid sdAb as a Theranostic Tool in Cancer Treatment,” *Clinical Cancer Research*, vol. 23, no. 21, pp. 6616–6628, Oct. 2017, doi: 10.1158/1078-0432.CCR-17-0310.
 - [234] G. Zhang, S. Brokx, and J. H. Weiner, “Extracellular accumulation of recombinant proteins fused to the carrier protein YebF in *Escherichia coli*,” *Nat Biotechnol*, vol. 24, no. 1, Art. no. 1, Jan. 2006, doi: 10.1038/nbt1174.
 - [235] T. Selas Castiñeiras, S. G. Williams, A. G. Hitchcock, and D. C. Smith, “*E. coli* strain engineering for the production of advanced biopharmaceutical products,” *FEMS Microbiology Letters*, vol. 365, no. 15, p. fny162, Aug. 2018, doi: 10.1093/femsle/fny162.
 - [236] M. K. Puurunen *et al.*, “Safety and pharmacodynamics of an engineered *E. coli* Nissle for the treatment of phenylketonuria: a first-in-human phase 1/2a study,” *Nat Metab*, vol. 3, no. 8, Art. no. 8, Aug. 2021, doi: 10.1038/s42255-021-00430-7.
 - [237] R. A. Morgan, J. C. Yang, M. Kitano, M. E. Dudley, C. M. Laurencot, and S. A. Rosenberg, “Case Report of a Serious Adverse Event Following the Administration of T Cells Transduced With a Chimeric Antigen Receptor Recognizing ERBB2,” *Molecular Therapy*, vol. 18, no. 4, pp. 843–851, Apr. 2010, doi: 10.1038/mt.2010.24.
 - [238] Y. Wu *et al.*, “Control of the activity of CAR-T cells within tumours via focused ultrasound,” *Nat Biomed Eng*, vol. 5, no. 11, Art. no. 11, Nov. 2021, doi: 10.1038/s41551-021-00779-w.
 - [239] S. Mc Cafferty *et al.*, “In Vivo Validation of a Reversible Small Molecule-Based Switch for Synthetic Self-Amplifying mRNA Regulation,” *Molecular*

- Therapy*, vol. 29, no. 3, pp. 1164–1173, Mar. 2021, doi: 10.1016/j.ymthe.2020.11.010.
- [240] H.-S. Li *et al.*, “Multidimensional control of therapeutic human cell function with synthetic gene circuits,” *Science*, vol. 378, no. 6625, pp. 1227–1234, Dec. 2022, doi: 10.1126/science.ade0156.
 - [241] E. S. Shepherd, W. C. DeLoache, K. M. Pruss, W. R. Whitaker, and J. L. Sonnenburg, “An exclusive metabolic niche enables strain engraftment in the gut microbiota,” *Nature*, vol. 557, no. 7705, Art. no. 7705, May 2018, doi: 10.1038/s41586-018-0092-4.
 - [242] A. J. Meyer, T. H. Segall-Shapiro, E. Glassey, J. Zhang, and C. A. Voigt, “*Escherichia coli* ‘Marionette’ strains with 12 highly optimized small-molecule sensors,” *Nat Chem Biol*, vol. 15, no. 2, Art. no. 2, Feb. 2019, doi: 10.1038/s41589-018-0168-3.
 - [243] T. Fink *et al.*, “Design of fast proteolysis-based signaling and logic circuits in mammalian cells,” *Nat Chem Biol*, vol. 15, no. 2, Art. no. 2, Feb. 2019, doi: 10.1038/s41589-018-0181-6.
 - [244] F. Meng and T. Ellis, “The second decade of synthetic biology: 2010–2020,” *Nat Commun*, vol. 11, no. 1, Art. no. 1, Oct. 2020, doi: 10.1038/s41467-020-19092-2.
 - [245] M.-P. Robinson *et al.*, “Efficient expression of full-length antibodies in the cytoplasm of engineered bacteria,” *Nat Commun*, vol. 6, no. 1, Art. no. 1, Aug. 2015, doi: 10.1038/ncomms9072.
 - [246] H. Nausch, J. Huckauf, R. Koslowski, U. Meyer, I. Broer, and H. Mikschofsky, “Recombinant Production of Human Interleukin 6 in *Escherichia coli*,” *PLOS ONE*, vol. 8, no. 1, p. e54933, Jan. 2013, doi: 10.1371/journal.pone.0054933.
 - [247] R. E. Ahan *et al.*, “A Highly Potent SARS-CoV-2 Blocking Lectin Protein,” *ACS Infect. Dis.*, vol. 8, no. 7, pp. 1253–1264, Jul. 2022, doi: 10.1021/acsinfecdis.2c00006.
 - [248] E. S. Kehribar, M. E. Isilak, E. U. Bozkurt, J. Adamcik, R. Mezzenga, and U. O. S. Seker, “Engineering of biofilms with a glycosylation circuit for biomaterial applications,” *Biomater. Sci.*, vol. 9, no. 10, pp. 3650–3661, May 2021, doi: 10.1039/D0BM02192J.
 - [249] A. Sandomenico, J. P. Sivaccumar, and M. Ruvo, “Evolution of *Escherichia coli* Expression System in Producing Antibody Recombinant Fragments,” *Int J Mol Sci*, vol. 21, no. 17, p. 6324, Aug. 2020, doi: 10.3390/ijms21176324.
 - [250] C. Huttenhower *et al.*, “Structure, function and diversity of the healthy human microbiome,” *Nature*, vol. 486, no. 7402, Art. no. 7402, Jun. 2012, doi: 10.1038/nature11234.
 - [251] M. Mimeo, R. J. Citorik, and T. K. Lu, “Microbiome therapeutics — Advances and challenges,” *Advanced Drug Delivery Reviews*, vol. 105, pp. 44–54, Oct. 2016, doi: 10.1016/j.addr.2016.04.032.
 - [252] J. J. Faith *et al.*, “The Long-Term Stability of the Human Gut Microbiota,” *Science*, vol. 341, no. 6141, p. 1237439, Jul. 2013, doi: 10.1126/science.1237439.
 - [253] D. B. Pedrolli *et al.*, “Chapter 8 - Engineering microbial living therapeutics,” in *Precision Medicine for Investigators, Practitioners and Providers*, J. Faintuch and S. Faintuch, Eds., Academic Press, 2020, pp. 71–82. doi: 10.1016/B978-0-12-819178-1.00008-3.

- [254] J. Walker *et al.*, “Chemotherapy-induced weight gain in early-stage breast cancer: a prospective matched cohort study reveals associations with inflammation and gut dysbiosis,” *BMC Medicine*, vol. 21, no. 1, p. 178, May 2023, doi: 10.1186/s12916-023-02751-8.
- [255] A. Langdon, N. Crook, and G. Dantas, “The effects of antibiotics on the microbiome throughout development and alternative approaches for therapeutic modulation,” *Genome Medicine*, vol. 8, no. 1, p. 39, Apr. 2016, doi: 10.1186/s13073-016-0294-z.
- [256] K. G. Jameson and E. Y. Hsiao, “Linking the Gut Microbiota to a Brain Neurotransmitter,” *Trends in Neurosciences*, vol. 41, no. 7, pp. 413–414, Jul. 2018, doi: 10.1016/j.tins.2018.04.001.
- [257] I. Cho and M. J. Blaser, “The human microbiome: at the interface of health and disease,” *Nat Rev Genet*, vol. 13, no. 4, Art. no. 4, Apr. 2012, doi: 10.1038/nrg3182.
- [258] J. Amar, “Microbiota–Host Crosstalk: A Bridge Between Cardiovascular Risk Factors, Diet, and Cardiovascular Disease,” *American Journal of Hypertension*, vol. 31, no. 8, pp. 941–944, Jul. 2018, doi: 10.1093/ajh/hpy067.
- [259] L. R. Lopetuso, F. Scaldaferri, V. Petito, and A. Gasbarrini, “Commensal Clostridia: leading players in the maintenance of gut homeostasis,” *Gut Pathogens*, vol. 5, no. 1, p. 23, Aug. 2013, doi: 10.1186/1757-4749-5-23.
- [260] G. L. Hold, S. E. Pryde, V. J. Russell, E. Furrie, and H. J. Flint, “Assessment of microbial diversity in human colonic samples by 16S rDNA sequence analysis,” *FEMS Microbiology Ecology*, vol. 39, no. 1, pp. 33–39, Jan. 2002, doi: 10.1111/j.1574-6941.2002.tb00904.x.
- [261] F. Rivera-Chávez *et al.*, “Depletion of Butyrate-Producing Clostridia from the Gut Microbiota Drives an Aerobic Luminal Expansion of Salmonella,” *Cell Host & Microbe*, vol. 19, no. 4, pp. 443–454, Apr. 2016, doi: 10.1016/j.chom.2016.03.004.
- [262] D. Hatzioanou *et al.*, “Discovery of a novel lantibiotic nisin O from *Blautia obeum* A2-162, isolated from the human gastrointestinal tract,” *Microbiology*, vol. 163, no. 9, pp. 1292–1305, 2017, doi: 10.1099/mic.0.000515.
- [263] P. Van den Abbeele *et al.*, “Butyrate-producing Clostridium cluster XIVa species specifically colonize mucins in an in vitro gut model,” *ISME J*, vol. 7, no. 5, Art. no. 5, May 2013, doi: 10.1038/ismej.2012.158.
- [264] M. T. Sorbara *et al.*, “Functional and Genomic Variation between Human-Derived Isolates of Lachnospiraceae Reveals Inter- and Intra-Species Diversity,” *Cell Host & Microbe*, vol. 28, no. 1, pp. 134–146.e4, Jul. 2020, doi: 10.1016/j.chom.2020.05.005.
- [265] N. Studer *et al.*, “Functional Intestinal Bile Acid 7 α -Dehydroxylation by *Clostridium scindens* Associated with Protection from *Clostridium difficile* Infection in a Gnotobiotic Mouse Model,” *Frontiers in Cellular and Infection Microbiology*, vol. 6, 2016, Accessed: Nov. 09, 2023. [Online]. Available: <https://www.frontiersin.org/articles/10.3389/fcimb.2016.00191>
- [266] R. Abdugheni *et al.*, “Metabolite profiling of human-originated Lachnospiraceae at the strain level,” *iMeta*, vol. 1, no. 4, p. e58, 2022, doi: 10.1002/imt2.58.
- [267] N. Arpaia *et al.*, “Metabolites produced by commensal bacteria promote peripheral regulatory T-cell generation,” *Nature*, vol. 504, no. 7480, Art. no. 7480, Dec. 2013, doi: 10.1038/nature12726.

- [268] Y. Furusawa *et al.*, “Commensal microbe-derived butyrate induces the differentiation of colonic regulatory T cells,” *Nature*, vol. 504, no. 7480, Art. no. 7480, Dec. 2013, doi: 10.1038/nature12721.
- [269] R. E. Ahan, “A self-actuated cellular protein delivery machinery,” Bilkent University, 2018. Accessed: Nov. 09, 2023. [Online]. Available: <http://hdl.handle.net/11693/47738>
- [270] T. Seemann, “Prokka: rapid prokaryotic genome annotation,” *Bioinformatics*, vol. 30, no. 14, pp. 2068–2069, Jul. 2014, doi: 10.1093/bioinformatics/btu153.
- [271] F. Teufel *et al.*, “SignalP 6.0 predicts all five types of signal peptides using protein language models,” *Nat Biotechnol*, vol. 40, no. 7, Art. no. 7, Jul. 2022, doi: 10.1038/s41587-021-01156-3.
- [272] M. C. Frith, N. F. W. Saunders, B. Kobe, and T. L. Bailey, “Discovering Sequence Motifs with Arbitrary Insertions and Deletions,” *PLOS Computational Biology*, vol. 4, no. 5, p. e1000071, May 2008, doi: 10.1371/journal.pcbi.1000071.
- [273] M. T. Sorbara and E. G. Pamer, “Microbiome-based therapeutics,” *Nat Rev Microbiol*, vol. 20, no. 6, Art. no. 6, Jun. 2022, doi: 10.1038/s41579-021-00667-9.
- [274] M. P. McNerney, K. E. Doiron, T. L. Ng, T. Z. Chang, and P. A. Silver, “Theranostic cells: emerging clinical applications of synthetic biology,” *Nat Rev Genet*, vol. 22, no. 11, Art. no. 11, Nov. 2021, doi: 10.1038/s41576-021-00383-3.
- [275] D. T. Riglar and P. A. Silver, “Engineering bacteria for diagnostic and therapeutic applications,” *Nat Rev Microbiol*, vol. 16, no. 4, Art. no. 4, Apr. 2018, doi: 10.1038/nrmicro.2017.172.
- [276] K. Vandenbroucke *et al.*, “Orally administered *L. lactis* secreting an anti-TNF Nanobody demonstrate efficacy in chronic colitis,” *Mucosal Immunology*, vol. 3, no. 1, pp. 49–56, Jan. 2010, doi: 10.1038/mi.2009.116.
- [277] B. J. Russell *et al.*, “Intestinal transgene delivery with native *E. coli* chassis allows persistent physiological changes,” *Cell*, vol. 185, no. 17, pp. 3263–3277.e15, Aug. 2022, doi: 10.1016/j.cell.2022.06.050.
- [278] J. M. Ridlon, D.-J. Kang, and P. B. Hylemon, “Bile salt biotransformations by human intestinal bacteria,” *Journal of Lipid Research*, vol. 47, no. 2, pp. 241–259, Feb. 2006, doi: 10.1194/jlr.R500013-JLR200.
- [279] P. Louis and H. J. Flint, “Formation of propionate and butyrate by the human colonic microbiota,” *Environmental Microbiology*, vol. 19, no. 1, pp. 29–41, 2017, doi: 10.1111/1462-2920.13589.
- [280] S. Heilbronner, B. Krismer, H. Brötz-Oesterhelt, and A. Peschel, “The microbiome-shaping roles of bacteriocins,” *Nat Rev Microbiol*, vol. 19, no. 11, Art. no. 11, Nov. 2021, doi: 10.1038/s41579-021-00569-w.
- [281] K. E. Jones *et al.*, “Global trends in emerging infectious diseases,” *Nature*, vol. 451, no. 7181, Art. no. 7181, Feb. 2008, doi: 10.1038/nature06536.
- [282] M. Marani, G. G. Katul, W. K. Pan, and A. J. Parolari, “Intensity and frequency of extreme novel epidemics,” *Proceedings of the National Academy of Sciences*, vol. 118, no. 35, p. e2105482118, Aug. 2021, doi: 10.1073/pnas.2105482118.
- [283] O. K. Lawler *et al.*, “The COVID-19 pandemic is intricately linked to biodiversity loss and ecosystem health,” *The Lancet Planetary Health*, vol. 5, no. 11, pp. e840–e850, Nov. 2021, doi: 10.1016/S2542-5196(21)00258-8.

- [284] W. B. Karesh *et al.*, “Ecology of zoonoses: natural and unnatural histories,” *The Lancet*, vol. 380, no. 9857, pp. 1936–1945, Dec. 2012, doi: 10.1016/S0140-6736(12)61678-X.
- [285] Q. Chen, J. Zhang, P. Wang, and Z. Zhang, “The mechanisms of immune response and evasion by the main SARS-CoV-2 variants,” *iScience*, vol. 25, no. 10, p. 105044, Oct. 2022, doi: 10.1016/j.isci.2022.105044.
- [286] R. Otsubo and T. Yasui, “Monoclonal antibody therapeutics for infectious diseases: Beyond normal human immunoglobulin,” *Pharmacol Ther*, vol. 240, p. 108233, Dec. 2022, doi: 10.1016/j.pharmthera.2022.108233.
- [287] D. Gouglas *et al.*, “Estimating the cost of vaccine development against epidemic infectious diseases: a cost minimisation study,” *The Lancet Global Health*, vol. 6, no. 12, pp. e1386–e1396, Dec. 2018, doi: 10.1016/S2214-109X(18)30346-2.
- [288] Y. Li, D. Liu, Y. Wang, W. Su, G. Liu, and W. Dong, “The Importance of Glycans of Viral and Host Proteins in Enveloped Virus Infection,” *Frontiers in Immunology*, vol. 12, 2021, Accessed: Nov. 14, 2023. [Online]. Available: <https://www.frontiersin.org/articles/10.3389/fimmu.2021.638573>
- [289] A. F. S. Santos, M. D. C. da Silva, T. H. Napoleão, P. M. G. Paiva, M. T. S. Correia, and L. C. B. B. Coelho, “Lectins: Function, structure, biological properties and potential applications,” 2014, Accessed: Nov. 14, 2023. [Online]. Available: <https://repositorium.sdum.uminho.pt/handle/1822/43440>
- [290] R. J. Mersny, “Strategies for targeting protein therapeutics to selected tissues and cells,” *Expert Opinion on Biological Therapy*, vol. 4, no. 1, pp. 65–73, Jan. 2004, doi: 10.1517/14712598.4.1.65.
- [291] M. Sassone-Corsi *et al.*, “Microcins mediate competition among Enterobacteriaceae in the inflamed gut,” *Nature*, vol. 540, no. 7632, Art. no. 7632, Dec. 2016, doi: 10.1038/nature20557.
- [292] H. Chen *et al.*, “Advances in Escherichia coli Nissle 1917 as a customizable drug delivery system for disease treatment and diagnosis strategies,” *Mater Today Bio*, vol. 18, p. 100543, Jan. 2023, doi: 10.1016/j.mtbio.2023.100543.
- [293] G. L. Rosano and E. A. Ceccarelli, “Recombinant protein expression in Escherichia coli: advances and challenges,” *Frontiers in Microbiology*, vol. 5, 2014, Accessed: Nov. 14, 2023. [Online]. Available: <https://www.frontiersin.org/articles/10.3389/fmicb.2014.00172>
- [294] C. Tropini, K. A. Earle, K. C. Huang, and J. L. Sonnenburg, “The Gut Microbiome: Connecting Spatial Organization to Function,” *Cell Host & Microbe*, vol. 21, no. 4, pp. 433–442, Apr. 2017, doi: 10.1016/j.chom.2017.03.010.
- [295] M. K. Mimee, “Genetic technologies to engineer and understand the microbiome,” Thesis, Massachusetts Institute of Technology, 2018. Accessed: Nov. 14, 2023. [Online]. Available: <https://dspace.mit.edu/handle/1721.1/123071>

APPENDIX A

Primers Used in the Thesis

Table A1. List of primers used in the study.

Primer Name	Sequence (5' to 3')	Chapter the
-------------	---------------------	-------------

		primer used
pREA144	CTCTAGATTAACTTTAAGAAGGAGGGTACCATGGGTAGCAGCCACCATC	Chapter 2
pREA145	TGGTGGTGGTGGTCTCGAGTTATTAATACTGTTTCATAATAAATATCCAG GCTATCCAG	Chapter 2
pREA196	TGAGACCTGTCTCGGAAGCTCATAGGACTCGTTTTAGAGCTATGCTGTTT TGAATGGTCCCCAAAACCTAGCATAACCCCTTGGGG	Chapter 3
TR-PK- REV	ATTTTCGCGGGATCGAGATCT	Chapter 3 and Chapter 4
pREA140	CTCGATCCCGCGAAATGACGTCAATTGTGAGCGGATAACAATTGACATTG TGAGCGG	Chapter 3
pREA195	TCCTATGAGCTTCCGAGACAGGTCTCAGTTTTGGGACCATTCAAAACAGC ATAGCTCTAAAACGTGCTCAGTATCTTGTTATCCGCTC	Chapter 3
pREA161	CGGGATCTCGACGCTCTC	Chapter 3
SEQ22R	CCTCCTTTTCAGCAAAAAACCCC	Chapter 3
pREA197	ATGGTCTCAAAACGGTTGGACTCAAGACGATAGTTACCGGATAGTTTTAG AGCTATGCTGTTTTGAATGGTCCCCAAAACAGAGACCGC	Chapter 3
pREA198	GCGGTCTCTGTTTTGGGACCATTCAAAACAGCATAGCTCTAAAACATATCC GGTAACTATCGTCTTGAGTCCAACCGTTTTGAGACCAT	Chapter 3
pREA199	ATGGTCTCAAAACGGCGAAACCCGACAGGACTATAAAGATACCGTTTTA GAGCTATGCTGTTTTGAATGGTCCCCAAAACAGAGACCGC	Chapter 3
pREA200	GCGGTCTCTGTTTTGGGACCATTCAAAACAGCATAGCTCTAAAACGGTAT CTTTATAGTCCTGTCGGGTTTCGCCGTTTTGAGACCAT	Chapter 3
pREA211	AAACCTTGAATTTGATCCCCGAGCCCTGAAGGAAGTTTTAGAGCTATGCT GTTTTGAATGGTCCCA	Chapter 3
pREA212	GTTTTGGGACCATTCAAAACAGCATAGCTCTAAAACCTTCCTTCAGGGCTC GGGGATCAAATTCAAG	Chapter 3
pREA213	AAACGCCCACACTCACCATCAAAAACCCCGAGAAGTTTTAGAGCTATGCT GTTTTGAATGGTCCCA	Chapter 3
pREA214	GTTTTGGGACCATTCAAAACAGCATAGCTCTAAAACCTTCCTCGGGGTTTT GATGGTGAGTGTGGGC	Chapter 3
pREA215	GAGGATCGAGATCTCGATCCCGCGAAATGACGTCTTTGTCATCCCCGAC AATCC	Chapter 3
pREA216	ATTTGATGCCTTATAGCTCATCCAGTGACACC	Chapter 3
pREA217	GGTGCACTGGATGAGCTATAAGGCATCAATAAAACGAAAGGCTCAG	Chapter 3
pREA218	GTTATCCGCTCACAATGACGTCATCTACCTTCCGTAAGCTGAAAATGAA TACTGCAGGTCTAGGGCGGCGGATTTG	Chapter 3
pREA221	TCCATTCTCAGGCTGTCTCGTCTCGTCTCGTTTCAGTGGTGCGTACAATTA AG	Chapter 3
pREA222	GTTGACCAGACAAACCACGACCAGTGCTCTTGTGGGTCCGATTGCCAGA TGATAAGGAACGCGCTGAACGCGATTCTG	Chapter 3
pREA203	AGAGCACTGGTCGTGTTTTGTCTGGTCAACCACCGCGGTCTCAGTGGTGT ACGGTACAAACCCCGACCGCGGAACCCCTATTTGT	Chapter 3
pREA204	GCCCGGATGATCCTGACGACGGAGACCGCCGTCGTCGACAAGCCGGCCG AGAAGATCCTTTGATCTTTTCTACGG	Chapter 3
pREA205	GGCGGTCTCCGTCGTCAGGATCATCCGGGCACTAGTGACGTCCACAGCTA ACAC	Chapter 3
pREA206	GAGACGAGACGAGACAGCCTGAGAATGGATGCGAGTAATGTCTAGGGCG GCGGATTTGTC	Chapter 3
pREA219	ATCCATTCTCAGGCTGTCTCGTCTCGTCTCTGGCTTAGTACAGCATCTATC GTAC	Chapter 3
pREA220	CCAGACAAACCACGACCAGTGCTCTTGTGGGTCCGATTGCCAGATGATA AGGAACGTACACACAGCGTCCCCCTCTTC	Chapter 3
Pmut5	AACTGTGAAGCGATGAACCC	Chapter 3
Pmut6	GGACTGTTCAAGAGAGCTATC	Chapter 3
Pmut7	GACCAAGCGATAACCGGATG	Chapter 3
Pmut8	GTGAGATGATGGCCACGATT	Chapter 3
OApETF	TAGCATAACCCCTTGGGG	Chapter 4
pREA138	CCGTTTAGAGGCCCAAGGGGTTATGCTAGCTCGAGAAAAAAGGCCCC TGTTGAA	Chapter 4
pREA139	ATAACAATTGACATTGTGAGCGGATAACAAGATACTGAGCACATATGGC CGGAGCAGGAG	Chapter 4
pREA140	CTCGATCCCGCGAAATGACGTCAATTGTGAGCGGATAACAATTGACATTG TGAGCGG	Chapter 4

pREA161	CGGGATCTCGACGCTCTC	Chapter 4
pREA190	AAGCTTGGCATCAAATAAAACGAAAGGCTCAG	Chapter 4
pREA191	AGTCGCATAAGGGAGAGCGTCGAGATCCCGTTTAGCTTCCTTAGCTCCTG AAAATC	Chapter 4
JO22	AGCAGCAGCGGTGCGGACGAGGTATTTATGTTACCTCCTTCAGATCTT	Chapter 4
pREA119	AAGAATTCAAAAGATCTGAAGGAGGGTACCATGAAATACCTGCTGCCGA	Chapter 4
pREA192	GAGCCTTTCGTTTTATTTGATGCCAAGCTTTCATCATTTTCAGCCCCAGA	Chapter 4
pREA236	AAGAATTCAAAAGATCTGAAGGAGGGTACCATGAAAAAGATTGGCTGG C	Chapter 4
pREA237	CGTTTTATTTGATGCCAAGCTTTCATTAGTGGTGGTGGTGGTGGTGGGAT CCTTTTTTCTCGGACAGATATTTCACTG	Chapter 4
pREA238	AAGAATTCAAAAGATCTGAAGGAGGGTACCATGAAACTCGCCGCCTGTT T	Chapter 4
pREA239	CGTTTTATTTGATGCCAAGCTTTCATCAGTGGTGGTGGTGGTGGTGGGAT CCCTGTTGCCCTGCCCAGAC	Chapter 4
pREA240	AAGAATTCAAAAGATCTGAAGGAGGGTACCATGAAATTAATATATTTA CTAAATCTATGATTGGTATGG	Chapter 4
pREA241	CGTTTTATTTGATGCCAAGCTTTCATCAGTGGTGGTGGTGGTGGTGGGAT CCTTGCTTAGTAAACGGCGC	Chapter 4
pREA227	CTGCAGTCACTTGAATTCGATACCCAGCTGCAAAAAACCCCTCAAGACCC GT	Chapter 4
pREA228	TTTCCCTATCAGTGATAGAGATTGACATCCCTATCAGTGATAGAGATACT GAGCACATATGGCCGGAGCAGGAGT	Chapter 4
pREA229	CAATCTCTATCACTGATAGGGAACGTTTCGCAGAAGCTTCCGC	Chapter 4
pREA230	CAGCTGGGTATCGAATTCAGTG	Chapter 4
pREA231	AAAGTGGTACCTTGCGGAAGCTTCTGCGAAACGTTTCCCTATCAGTGATA GAGATTGACATCC	Chapter 4
pREA264	AAACTTGCTGTGGAAGCTGCCTGCACTAATGTTCTGTTTTAGAGCTATGCT GTTTTGAATGGTCCCA	Chapter 4
pREA265	GTTTTGGGACCATTCAAAACAGCATAGCTCTAAACGAACATTAGTGCAG GCAGCTTCCACAGCAA	Chapter 4
pREA223	GCCATACTGATGAGTCCGTGAGGACGAAACGAGTAAGCTCGTCTATGGC CGGAGCAGGAG	Chapter 4
pREA234	AAGCATGTTGCCAGCCGGCGCCAGCGAGGAGGCTGGGACCATGCCGGC CAAAAAAAGGCCCTGTTGAA	Chapter 4
pREA235	ATGAAAGCCGCTCTGGGGCTGAAATGATGAAAGCTTGCCATACTGATGA GTCCGTGAG	Chapter 4
pREA236	TCGACTGAGCCTTTCGTTTTATTTGATGCCACGCGTGTCCCATTGCCCATG CCGAAGCATGTTGCCAGCC	Chapter 4
pREA261	AGCAGCAGCGGTGCGCAGCAGGTATTTATGTTACCTCCTTTCTAGAGT CGTGGTTTGTCTGGTCAAC	Chapter 4
pREA262	ACATTTCCCCGAAAAGTGCCACCTGACGTCAAATAATAAAAAAGCCGGA TTAATAATCTGGC	Chapter 4
JO17	ATGAAATACCTGCTGCCGAC	Chapter 4
pBS216	TCGACTGAGCCTTTCGTTTTATTTGATGCCACGCGTTCATTTAGCCCC	Chapter 4
pREA284	CTAGTATTTCTCCTCTTTGTTAAATCTCCCAATG	Chapter 4
pREA285	GTATCAGCTCACTCAAAGGC	Chapter 4
pREA286	GGGAGATTTAACAAAGAGGAGAAATACTAGATGAAATACCTGCTGCCGA CC	Chapter 4
pREA287	CCGTATTACCGCCTTTGAGTGAGCTGATACTCGAACGACCGAGCGTAG	Chapter 4
pREA280	CCGTTTAGAGGCCCAAGGGGTTATGCTAGCTCGAGTTAATTTTTAAAGT ATGGGCAATCAATTGCTC	Chapter 4
pREA281	CGGACTCATCAGTATGGCAAGCTTTCATCACGCTGCAAGGGCGTAATT	Chapter 4
pREA282	TGATGAAAGCTTGCCATACTGATGAG	Chapter 4
pREA283	TAGAGGATCGAGATCTCGATCCCGCGAAATGACGTCTCGAACGACCGAG CGTAGC	Chapter 4
pREA291	GACGTCTCGAACGACCGA	Chapter 4
pREA292	TAGAGGATCGAGATCTCGATCCCGCGAAATATTTGTCCTACTCAGGAGAG CG	Chapter 4
pREA293	CGCTCGGTGCTTCGAGACGTCTTGACAGCTAGCTCAGTCCTAGGTATAAT GCTAGCACCATTAAAGAGGAGAAAGGTACCATGAC	Chapter 4
pREA119	AAGAATTCAAAAGATCTGAAGGAGGGTACCATGAAATACCTGCTGCCGA	Chapter 4
pREA242	TTTTATTTGATGCCAAGCTTTCATCAGTGGTGGTGGTGGTGGTGGAGCTCT GCTGCTGCGCTGCTCACGGTCACCT	Chapter 4
pREA243	CATGAAAAAGATTTGGCTGGCGCTGGCTGGTTTAGTTTTAGCGTTTAGCG	Chapter 4

	CATCGGCGG	
pREA244	GATCCCGCCGATGCGCTAAACGCTAAAACTAAACCAGCCAGCGCCAGCC AAATCTTTTTTCATGGTAC	Chapter 4
pREA245	CATGAAAAAATTATTATTCGCAATTCCTTTAGTTGTTCTTTCTATTCTCA CTCCG	Chapter 4
pREA246	GATCCGGAGTGAGAATAGAAAGGAACAATAAGGAATTGCGAATAATA ATTTTTTCATGGTAC	Chapter 4
pREA247	CATGAAAGCTACTAAACTGGTACTGGGCGCGGTAATCCTGGGTTCTACTC TGCTGGCAGGTG	Chapter 4
pREA248	GATCCACCTGCCAGCAGAGTAGAACCCAGGATTACCGCGCCCAGTACCA GTTTAGTAGCTTTCATGGTAC	Chapter 4
pREA249	CATGAATAAAAGTAAAATGTTATGTTTTATTTACGGCGTTACTATCCTCTCT ATATGCACACGGAG	Chapter 4
pREA250	GATCCTCCGTGTGCATATAGAGAGGATAGTAACGCCGTAAATAAAACAT AACATTTTACTTTTATTCATGGTAC	Chapter 4
pREA251	CATGAAAATAAAAAACAGGTGCACGCATCCTCGCATTATCCGCATTAACG ACGATGATGTTTTCCGCCCTCGGCTCTCGCCG	Chapter 4
pREA252	GATCCGGCGAGAGCCGAGGCGGAAAACATCATCGTCGTTAATGCGGATA ATGCGAGGATGCGTGCACCTGTTTTATTTTCATGGTAC	Chapter 4
pREA253	CATGAAAAAGACAGCTATCGCGATTGCAGTGGCACTGGCTGGTTTCGCTA CCGTAGCGCAGGCCG	Chapter 4
pREA254	GATCCGGCCTGCGCTACGGTAGCGAAACCAGCCAGTGCCACTGCAATCG CGATAGCTGTCTTTTCATGGTAC	Chapter 4
pREA255	AAGATCTGAAGGAGGGTACCATGAACAATAACGATCTCTTTCAGG	Chapter 4
pREA256	TGCAGCTGCACCTGGGATCCCGCAGTCGCACGTCGCG	Chapter 4
pREA294	CTCGCTGCCAGCCGGCGATGGCCGGATCCATGCGCCCGTCTGGCCGT	Chapter 4
pREA295	GAGCCTTCGTTTTATTTGATGCCAAGCTTTCATCACTCATCCTCCTGGAA G	Chapter 4
read3	AAAATTTTCTCCTTTACTGCAGGAGCT	Chapter 5
JHA.J1	GTATTCACCTTAGAGGACTTTGTGGGT	
read3	CTGAGCTCCTGCAGTAAAGGAGAAAAATTTATGAATAAGTATAAAAAAA TTGTTTCTATAGCAGC	Chapter 5
read4	GTCACCCACAAAGTCCTCTAAAGTGAATACTGTCTGGGCCTCTTGGGC	Chapter 5
read5	CTGAGCTCCTGCAGTAAAGGAGAAAAATTTATGTATAGGAAAAAGTTAA TGCGCGCATGC	Chapter 5
read6	GTCACCCACAAAGTCCTCTAAAGTGAATACTGTCTGGTGCTTCAGCTGC	Chapter 5
read17	CTGAGCTCCTGCAGTAAAGGAGAAAAATTTATGAGAAGATACAAAAAGA GAATCTCAAGA	Chapter 5
read18	GTCACCCACAAAGTCCTCTAAAGTGAATACTTGTGTCTTTTCCATAGCAT GTACCATATT	Chapter 5
read19	TGAGCTCCTGCAGTAAAGGAGAAAAATTTATGAGAAAAATTTTAAACG ATTTGCTTCAC	Chapter 5
read20	GTCACCCACAAAGTCCTCTAAAGTGAATACAGAAGTATTCTCCGATGCAT GGAC	Chapter 5
read21	CTGAGCTCCTGCAGTAAAGGAGAAAAATTTATGAAAAAGAAAGTTTATAG CAGTGTTACTT	Chapter 5
read22	CTGTCTCCAGTCACCCACAAAGTCCTCTAAAGTGAATACTCCTGCTAACA ATCCTGCTAC	Chapter 5
read23	TGAGCTCCTGCAGTAAAGGAGAAAAATTTATGAAAATGAAAAAGTACG AAGTGTAAGTC	Chapter 5
read24	GTCACCCACAAAGTCCTCTAAAGTGAATACTGTATCATCTGGTGCCGC	Chapter 5
read25	CTGAGCTCCTGCAGTAAAGGAGAAAAATTTATGAAGAAGAAGAAAAAGT TGCTTGCACTA	Chapter 5
read26	ACAATGTTAATACCATAGCCAATGCTAAGAATAGTGCAAGCAACTTTTTT TTCTTCTTCA	Chapter 5
read27	TCTTAGCATTTGGCTATGGTATTAACATTGTTGCCGGTAACACCTGTGTTG CG	Chapter 5
read28	GTCACCCACAAAGTCCTCTAAAGTGAATACTTCTGCTGCGAACACAGGTG TTACCGGCA	Chapter 5
JHA.C08	AAGCTTGGCACTGGCCGTGCG	Chapter 5
read7	AGCCAAAATCCTCTCGCACAA	Chapter 5
read8	CGTTGTAAAACGACGGCCAGTGCCAAGCTTTTATTTGAATTTTTTACTAA TACTGCAAGGATTG	Chapter 5
read9	GGTTGGAGATTGTGCGAGAGGATTTTGGCTGGTGGTGGTAGTGGTAAGA AAGATGTAGAGATCAGAGAAG	Chapter 5
read10	GGTTGGAGATTGTGCGAGAGGATTTTGGCTGGTGGTGGTAGTGGTCAGG	Chapter 5

	AAGTTGGAGATTTC	
read11	GGTTGGAGATTGTGCGAGAGGATTTGGCTGGTGGTAGTGGTGACG ACAAAAAGAGGATACTAAGA	Chapter 5
read12	GGTTGGAGATTGTGCGAGAGGATTTGGCTGGTGGTAGTGGTTCTCC AAAGACCGGAGATGA	Chapter 5

APPENDIX B

Genes Used in the Thesis

Table B-1. Protein or DNA sequences of genes in the thesis

Gene name (amino acids or DNA)	Sequence
Griffithsin	MGSSHHHHHSSGLVPRGSLTHRKFGGSGGSPFSGLSSIAV RSGSYLDAIIIDGVHHGGSGGNLSPTFTFGSGEYISNMTIRSG DYIDNISFETNMGRRFPGPYGGSGGSANTLSNVKVIQINGSA GDYLDSDIYYEQY
Mock- CRISPR array	GTTTATAGAGCTATGCTGTTTTGAATGGTCCCAAACTGA GACCTGTCTCGGAAGCTCATAGGACTCGTTTTAGAGCTA TGCTGTTTTGAATGGTCCCAAAAC
E1 gRNA with repeat	GGTTGGACTCAAGACGATAGTTACCGGATAGTTTTAGAG CTATGCTGTTTTGAATGGTCCCAAAAC
E2 gRNA with repeat	GGCGAAACCCGACAGGACTATAAAGATACCGTTTTAGAG CTATGCTGTTTTGAATGGTCCCAAAAC
g9 gRNA with repeat	CTTGAATTTGATCCCCGAGCCCTGAAGGAAGTTTTAGAG CTATGCTGTTTTGAATGGTCCCAAAAC
g10 rRNA with repeat	GGCGAAACCCGACAGGACTATAAAGATACCGTTTTAGAG CTATGCTGTTTTGAATGGTCCCAAAAC
relB	MMATINIRIDDELKRSRYAALEKLGVTPEVLRTLEYVAQ SGRLPFQVLLTEDDADLMAIVRDRLNPQAGVKVSLDEL
sfGFP	ATGCGTAAAGGCGAAGAGCTGTTCACTGGTGTCTGTCCT ATTCTGGTGGAACTGGATGGTGTATGTCAACGGTCATAAG TTTTCCGTGCGTGGCGAGGGTGAAGGTGACGCAACTAAT GGTAACTGACGCTGAAGTTCATCTGTACTACTGGTAAA CTGCCGGTACCTTGCCGACTCTGGTAACGACGCTGACT TATGGTGTTCAGTGCTTTGCTCGTTATCCGGACCATATGA AGCAGCATGACTTCTTCAAGTCCGCCATGCCGGAAGGCT ATGTGCAGGAACGCACGATTTCTTTAAGGATGACGGCA CGTACAAAACGCGTGCAGGAAGTGAAATTTGAAGGCGAT ACCCTGGTAAACCGCATTGAGCTGAAAGGCATTGACTTT AAAGAAGACGGCAATATCCTGGGCCATAAGCTGGAATA CAATTTTAACAGCCACAATGTTTACATCACCGCCGATAA ACAAAAAATGGCATTAAAGCGAATTTTAAAATTCGCCA CAACGTGGAGGATGGCAGCGTGCAGCTGGCTGATCACTA CCAGCAAAACACTCCAATCGGTGATGGTCTGTTCTGCT GCCAGACAATCACTATCTGAGCACGCAAAGCGTTCTGTC TAAAGATCCGAACGAGAAACGCGATCATATGGTTCTGCT GGAGTTCGTAACCGCAGCGGGCATCACGCATGGTATGGA TGAAGTGTACAAATGA
proD	CACAGCTAACACCACGTCGTCCTATCTGCTGCCCTAGG

	TCTATGAGTGGTTGCTGGATAACTTTACGGGCATGCATA AGGCTCGTATAATATATTCAGGGAGACCAAATAATTTTG TTAACTTT
micL	TATGGCCGGAGCAGGAGTCCGTGCAGAAAACCTGCACC ACTTCCTCGATGCCGGAGTGCTGGAAGTCCATAGCTCCG CGGGAGCGTGGCAAGCCTCACCGATGCGTTATCGTAATC AAGGATTGTCCATGTCATCAGATGAACACGCGGACGAGT ATTCGCGTTATATCGTAGACGGGGCGGCGGTTGCTGAAA TGAAAGGAATCATTGAACGCCATCAGGCCAAATGATTTT TACCGTTGCATCATGTGCGCCAATATGATGCTTGCTCGTA CCAGGCCCCTGCAATTTCAACAGGGGCCTTTTTTT
Compact- micL	GCCATACTGATGAGTCCGTGAGGACGAAACGAGTAAGCT CGTCTATGGCCGGAGCAGGAGTCCGTGCAGAAAACCTGC ACCACTTCCTCGATGCCGGAGTGCTGGAAGTCCATAGCT CCGCGGGAGCGTGGCAAGCCTCACCGATGCGTTATCGTA ATCAAGGATTGTCCATGTCATCAGATGAACACGCGGACG AGTATTCGCGTTATATCGTAGACGGGGCGGCGGTTGCTG AAATGAAAGGAATCATTGAACGCCATCAGGCCAAATGA TTTTTACCGTTGCATCATGTGCGCCAATATGATGCTTGCT CGTACCAGGCCCCTGCAATTTCAACAGGGGCCTTTTTTTG GCCGGCATGGTCCCAGCCTCCTCGCTGGCGCCGGCTGGG CAACATGCTTCGGCATGGCGAATGGGAC
bALP	MKYLLPTAAAGLLLLAAQPAMAMGHHHHHHGGGGSRTPE MPVLENRAAQGDITAPGGARRLTGDQTAALRDSLSDKPAK NIILLIGDGMGDSEITAARNYAEGAGGFFKGIDALPLTGQYT HYALNKKTGKPDYVTDSAASATAWSTGVKTYNGALGVDI HEKDHPTILEMAKAAGLATGNVSTAEQDATPAALVAHVT SRKCYGPSATSEKCPGNALEKGGKGSITEQLLNARADVTLG GGAKTFAETATAGEWQGKTLREQAQARGYQLVSDAASLN SVTEANQQKPLLGLFADGNMPVRWLGPATYHGNIDKPA VTCTPNPQRNDSVPTLAQMTDKAIELLSKNEKGFFLQVEGA SIDKQDHAANPCGQIGETVDLDEAVQRALEFAKKEGNTLVI VTADHAHASQIVAPDTKAPGLTQALNTKDGAVMVMSYGN SEEDSQEHTGSQLRIAAAYGPHAANVVGLTDQTDLFYTMKA ALGLK
DsbA	MKKIWLALAGLVLAFSASAAQYEDGKQYTTLEKPVAGAP QVLEFFSFFCPHCYQFEEVLHISDNVKKKLPEGVKMTKYHV NFMGGDLGKDLTQAWAVAMALGVEDKVTVPLFEGVQKT QTIRSASDIRDVFINAGIKGEEYDAAWNSFVVKSLVAQQEK AAADVQLRGVPAMFVNGKYQLNPQGMDTSNMDVFVQQY ADTVKYLSEKKGSHHHHHH
MalS	MKLAACFLTLLPGFAVAASWTSPGFPAFSEQGTGTFVSHAQ LPKGTRPLTLNFDQQCWQPADAIKLNQMLSLQPCSNTPPQ WRLFRDGEYTLQIDTRSGTPTLMISIQNAEPVASLVRECPK WDGLPLTVDVSAFPEGAAVRDYYSQQIAIVKNGQIMLQP AATSNGLLLLERAETDTSAPFDWHNATVYFVLTDRFENGD PSNDQSYGRHKDGM AEIGTFHGGDLRGLTNKLDYLQQLG VNALWISAPFEQIHGWVGGGTKGDFPHYAYHGYTQDWT NLDANMGNEADLRTLVD SAHQRGIRILFDVVMNHTGYATL ADMQEYQFGALYLSGDEVKKS LGERWSDWKPAAGQTDWH

	SFNDYINFSDKTGWDKWWGKNWIRTDIGDYDNPGFDDL MSLAFLPDIKTESTTASGLPVFYKNKMDTHAKAIDGYTPRD YLTHWLSQWVRDYGIDGFRVDTAKHVELPAWQQLKTEAS AALREWKKANPDKALDDKPFWMTGEAWGHGVMQSDYY RHGFDAMINFDYQEQAQAVDCLAQMDTTWQQMAEKLQ GFNVLSYLSSHDTRLFREGGDKAAELLLLAPGAVQIFYGDE SSRPFGPTGSDPLQGTRSDMNWQDVSGKSAASVAHWQKIS QFRARHPAIGAGKQTTLLLKQGYGFVREHGDDKVLVVA GQQGSHHHHHH
ChiA	MKLNIFTKSMIGMGLVCSALPALAMEAWNQQGGNKYQV IFDGKIYENAWWVSSTNCPGKAKANDATNPWRLKRTATA AEISQFGNTLSCEKSGSSSSSSNSNTPASNTPANGGSATPAQG TVPSNSSVVAWNKQQGGQTWYVVFNGAVYKNAWWVASS NCPGDAKSNDASNPWRYVRAATATEISETSNPQSCTSAQP SPDVKPAPDVKPAPDVQPAPADKSNDNYAVVAWKGQEGS STWYVIYNGGIYKNAWWVGAANCPGDAKENDASNPWRY VRAATATEISQYGNPGSCSVKPDNNGGAVTPVDPTPETPVT PTPDNSEPSTPADSVNDYSLQAWSGQEGSEIYHVIFNGNVY KNAWWVGSKDCPRGTSANSENPNPWLRLERTATAAELSQYG NPTTCEIDNGGVIVADGFQASKAYSADSIVDYNDAHYKTSV DQDAWGFVPGGDNPWKKYEPAKAWSASTVYVKGDRVVV DGQAYEALFWTQSDNPALVANQNATGSNSRPWKPLGKAQ SYSNEELNNAPQFNPETLYASDTLIRFNGVNYISQSKVQKVS PSDSNPWRVFVDWTGTERVGTGPKAWPKHVIYAPYVDFT LNTIPDLAALAKNHNVNHFLLAFVVS KDANTCLPTWGTAY GMQNYAQYSKIKALREAGGDVMLSIGGANNAPLAASCKN VDDLMMQHYDYDIVDNLNLKVLDFDIEGTWVADQASIERNL AVKKVQDKWKSEGKDIAIWYTLPIPTGLTPEGMNVLSDA KAKGVELAGVNVMTMDYGNAICQSANTEGQNIHGKCATS AIANLHSQLKGLHPNKSDAEIDAMMGTTMPVGVNDVQGE VFYLS DARLVMQDAQKRNLGMVGIWSIARDLPGGTNLSPE FHGLTKEQAPKYAFSEIFAPFTKQGSHHHHHH
pBAD	TGCATAATGTGCCTGTCAAATGGACGAAGCAGGGATTCT GCAAACCCTATGCTACTCCGTCAAGCCGTCAATTGTCTG ATTCGTTACCAATTATGACAACTTGACGGCTACATCATTC ACTTTTTCTTCACAACCGGCACGGAAGCTCGCTCGGGCTG GCCCCGGTGCATTTTTTAAATACCCGCGAGAAATAGAGT TGATCGTCAAAACCAACATTGCGACCGACGGTGCGGATA GGCATCCGGGTGGTGCTCAAAAGCAGCTTCGCCTGGCTG ATACGTTGGTCCTCGCGCCAGCTTAAGACGCTAATCCCT AACTGCTGGCGGAAAAGATGTGACAGACGCGACGGCGA CAAGCAAACATGCTGTGCGACGCTGGCGATATCAAAATT GCTGTCTGCCAGGTGATCGCTGATGTACTGACAAGCCTC GCGTACCCGATTATCCATCGGTGGATGGAGCGACTCGTT AATCGCTTCCATGCGCCGCAGTAACAATTGCTCAAGCAG ATTTATCGCCAGCAGCTCCGAATAGCGCCCTTCCCCTTGC CCGGCGTTAATGATTTGCCCAAACAGGTCGCTGAAATGC GGCTGGTGCGCTTCATCCGGGCGAAAGAACCCCGTATTG GCAAATATTGACGGCCAGTTAAGCCATTCATGCCAGTAG GCGCGCGGACGAAAGTAAACCCACTGGTGATACCATTG

	CGAGCCTCCGGATGACGACCGTAGTGATGAATCTCTCCT GGCGGGAACAGCAAAATATCACCCGGTCGGCAAACAAA TTCTCGTCCCTGATTTTTTCACCACCCCCTGACCGCGAATG GTGAGATTGAGAATATAACCTTTCATTCCCAGCGGTCTGG TCGATAAAAAAATCGAGATAACCGTTGGCCTCAATCGGC GTTAAACCCGCCACCAGATGGGCATTAAACGAGTATCCC GGCAGCAGGGGATCATTTTTGCGCTTCAGCCATACTTTTC ATACTCCCGCCATTTCAGAGAAGAAACCAATTGTCCATAT TGCATCAGACATTGCCGTCACCTGCGTCTTTTACTGGCTCT TCTCGCTAACCAAACCGGTAACCCCGCTTATTAAAAGCA TTCTGTAACAAAGCGGGACCAAAGCCATGACAAAAACG CGTAACAAAAGTGTCTATAATCACGGCAGAAAAAGTCCAC ATTGATTATTTGCACGGCGTCACACTTTGCTATGCCATAG CATTTTTATCCATAAGATTAGCGGATCCTACCTGACGCTT TTTATCGCAACTCTCTACTGTTTCTCCATACCCGTTTTTTG GG
pNTetO	TTTAGCTTCCTTAGCTCCTGAAAATCTCGATAACTCAAAA AATACGCCCGGTAGTGATCTTATTTTCATTATGGTGAAAG TTGGAACCTCTTACGTGCCGATCAACGTCTCATTTTCGCC AGATATCGACGTCTTAAGACCCACTTTCACATTTAAGTTG TTTTTCTAATCCGCATATGATCAATTCAAGGCCGAATAA GAAGGCTGGCTCTGCACCTTGGTGATCAAATAATTTCGAT AGCTTGTCTGTAATAATGGCGGCATACTATCAGTAGTAGG TGTTTCCCTTTCTTCTTTAGCGACTTGATGCTCTTGATCTT CCAATACGCAACCTAAAGTAAAAATGCCCCACAGCGCTGA GTGCATATAATGCATTCTCTAGTGAAAAACCTTGTTGGC ATAAAAAGGCTAATTGATTTTCGAGAGTTTCATACTGTTT TTCTGTAGGCCGTGTACCTAAATGTACTTTTGCTCCATCG CGATGACTTAGTAAAGCACATCTAAAACCTTTTAGCGTTA TTACGTAAAAAATCTTGCCAGCTTTCCCCTTCTAAAGGGC AAAAGTGAGTATGGTGCCTATCTAACATCTCAATGGCTA AGGCGTCGAGCAAAGCCCGCTTATTTTTTACATGCCAAT ACAATGTAGGCTGCTCTACACCTAGCTTCTGGGCGAGTT TACGGGTTGTAAACCTTCGATTCCGACCTCATTAAGCA GCTCTAATGCGCTGTTAATCACTTTACTTTTATCTAATCT AGACATCATTAATTCCTAATTTTTGTTGACACTCTATCGT TGATAGAGTTATTTTACCACTCCCTATCAGTGATAGAGA AAAGAATTCAAAAGATCT
pTetO	TCCCTATCAGTGATAGAGATTGACATCCCTATCAGTGAT AGAGATACTGAGCACATCAGCAGGACGCACTGACC
pLacO	TCCCTATCAGTGATAGAGATTGACATCCCTATCAGTGAT AGAGATACTGAGCACA
pLux	TAAAACGGTAATGGATTGACATTTGATTCTAATAAATTG GATTTTTGTACACTATTGTATCGCTGGGAATACAATTAC TTAACATAAGCACCTGTAGGATCGTACAGGTTTACGCAA GAAAATGGTTTGTATAGTCGAATGAATTCATT
LuxR	ATGATATATAACACGCAAAACTTGCGACAAACAATAGGT AAGGATAAAGAGATGGGTATGAAAAACATAAATGCCGA CGACACATACAGAATAATTAATAAAATTAAAGCTTGATG AAGCAATAATGATATTAATCAATGCTTATCTGATATGAC

	TAAAATGGTACATTGTGAATATTATTTACTCGCGATCATT TATCCTCATTCTATGGTTAAATCTGATATTTCAATTCTAG ATAATTACCCTAAAAAATGGAGGCAATATTATGATGACG CTAATTTAATAAAATATGATCCTATAGTAGATTATTCTAA CTCCAATCATTACCAATTAATTGGAATATATTTGAAAA CAATGCTGTAAATAAAAAATCTCCAAATGTAATTAAGA AGCGAAAACATCAGGTCTTATCACTGGGTTTAGTTTCCCT ATTCATACGGCTAACAATGGCTTCGGAATGCTTAGTTTTG CACATTCAGAAAAAGACAACCTATATAGATAGTTTATTTT TACATGCGTGTATGAACATAACCATTAAATTGTTCTTCTCT AGTTGATAATTATCGAAAAATAAATATAGCAAATAATAA ATCAAACAACGATTTAACCAAAAAGAGAAAAAGAATGTT TAGCGTGGGCATGCGAAGGAAAAAGCTCTTGGGATATTT CAAAAATATTAGGCTGCAGTGAGCGTACTGTCACTTTCC ATTTAACCAATGCGCAAATGAACTCAATACAACAAACC GCTGCCAAAGTATTTCTAAAGCAATTTTAACAGGAGCAA TTGATTGCCCATACTTTAAAAATTAA
LuxI	ATGACTATAATGATAAAAAAATCGGATTTTTTGGCAATT CCATCGGAGGAGTATAAAGGTATTCTAAGTCTTCGTTAT CAAGTGTTTAAGCAAAGACTTGAGTGGGACTTAGTTGTA GAAAATAACCTTGAATCAGATGAGTATGATAACTCAAAT GCAGAATATATTTATGCTTGTGATGATACTGAAAATGTA AGTGGATGCTGGCGTTTATTACCTACAACAGGTGATTAT ATGCTGAAAAGTGTTTTTCTGAATTGCTTGGTCAACAG AGTGCTCCCAAAGATCCTAATATAGTCGAATTAAGTCGT TTTGCTGTAGGTAAAAATAGCTCAAAGATAAATAACTCT GCTAGTGAAATTACAATGAACTATTTGAAGCTATATAT AAACACGCTGTTAGTCAAGGTATTACAGAATATGTAACA GTAACATCAACAGCAATAGAGCGATTTTTTAAAGCGTATT AAAGTTCCTTGTCATCGTATTGGAGACAAAGAAATTCAT GTATTAGGTGATACTAAATCGGTTGTATTGTCTATGCCTA TTAATGAACAGTTTAAAAAAGCAGTCTTAAATGCAGCGA ACGACGAAAATTACGCCCTTGCAGCGTAA
Inverted proD with BxbI sites	TCGGCCGGCTTGTCGACGACGGCGGTCTCCGTCGTCAGG ATCATCCGGGCCTTAAAAAGTTAAACAAAATTATTTGTA GAGGGAAACCGTTGTGGTCTCCCTGAATATATTATACGA GCCTTATGCATGCCCGTAAAGTTATCCAGCAACCACTCA TAGACCTAGGGCAGCAGATAGGGACGACGTGGTGTTAG CTGTGGTACCGTCGGGGTTTGTACCGTACACCACTGAGA CCGCGGTGGTTGACCAGACAAACCACGAC
BxbI	ATGGGAACGGTGGCGCAGATGGAATTAGAAGCGATCAA AGAGCGGAACCGTTCGGCTGCGCATTTCATATCCGCGC CGGGAAATACCGAGGATCCCTGCCGCCGTGGGGATACCT GCCTACGCGCGTGGACGGGGAGTGGCGGCTGGTGCCGG ACCCTGTGCAGCGAGAGCGCATCCTCGAGGTGTATCACC GCGTCGTCGACAACCACGAGCCGCTGCATCTGGTGGCCC ACGACCTGAACCGGCGTGGTGTCTGTGCGCGAAGGACT ACTTCGCGCAGCTGCAAGGCCGCGAGCCGCAGGGCCGG GAGTGGTCGGCTACCGCGCTGAAGCGATCGATGATCTCC GAGGCGATGCTCGGGTACGCGACTCTGAACGGTAAGACC

	GTCCGAGACGACGACGGAGCCCCGCTGGTGCGGGCTGA GCGGATCCTGACCCGTGAGCAGCTGGAGGCGCTGCGCGC CGAGCTCGTGAAGACCTCCCGGGCGAAGCCCGCGGTGTC TACCCCGTCGCTGCTGCTGCGGGTGTTGTTCTGCGCGGTG TGCGGGGAGCCCGCGTACAAGTTCGCCGGGGGAGGACG TAAGCACCCGCGCTACCGCTGCCGCTCGATGGGGTTCCC GAAGCACTGCGGGAACGGCACGGTGGCGATGGCCGAGT GGGACGCGTTCTGCGAGGAGCAGGTACTGGATCTGCTCG GGGACGCGGAGCGTCTGGAGAAAGTCTGGGTAGCGGGC TCGGACTCCGCGGTCTGAACCTCGCGGAGGTGAACGCGGA GCTGGTGGACCTGACGTCGCTGATCGGCTCCCCGGCCTA CCGGGCGGGCTCTCCGCAGCGAGAAGCACTGGATGCCCG TATTGCGGCGCTGGCCGCGCGGCAAGAGGAGCTGGAGG GCCTGGAGGCTCGCCCGTCTGGCTGGGAGTGGCGCGAGA CCGGGCAGCGGTTCTGGGGACTGGTGGCGGGAGCAGGAC ACCGCGGCAAAGAACACCTGGCTTCGGTCGATGAACGTT CGGCTGACGTTTCGACGTCCGCGGCGGGCTGACTCGCACG ATCGACTTCGGGGATCTTCAGGAGTACGAGCAGCATCTC AGGCTCGGCAGCGTGGTCTGAACGGCTACACACCGGGAT GTCGTAA
IL1-Ra	MKYLLPTAAAGLLLLAAQPAMAGSMRPSGRKSSKMQAfri WDVNQKTFYLRNNQLVAGYLQGPVNLLEEKIDVVPiEPHA LFLGIHGGKMCLSCVKSGDETRLQLEAVNITDLSENrkQDK RFAFIRSDSGPTTSFESAACPGWFLCTAMEADQPVSLTNMP DEGVMVTKFYFQEDE
2Rs15d nanobody	MKYLLPTAAAGLLLLAAQPAMAGSQVQLQESGGGSVQAG GSLKLTCAASGYIFNSCGMGWYRQSPGRERELVSRISGDGD TWHKESVKGRFTISQDNVKKTLYLQMNSLKPEDTAVYFCA VCYNLETYWGGQTQVTVSSAAELHHHHH
NanoLuc	GTATTCACCTTTAGAGGACTTTGTGGGTGACTGGAGACAG ACTGCAGGATACAATTTGGACCAGGTGCTTGAGCAGGGT GGTGTAAGTAGCCTATTTCAAACTTAGGAGTTAGCGTG ACTCCTATACAGAGGATAGTTTTGAGCGGAGAAAATGGT TTGAAAATTGATATTCATGTAATTATTCATACGAGGGTT TATCAGGTGACCAAATGGGTCAGATTGAAAAGATATTCA AAGTAGTATATCCTGTGGACGATCACCATTTCAAGGTAA TATTACACTATGGAACCTCTAGTGATAGACGGTGTGACAC CTAATATGATTGACTATTTTGGTAGACCTTACGAGGGTAT TGCAGTGTTGATGGAAAAAAGATAACTGTGACTGGAAC TCTTTGGAACGGTAACAAAATAATAGACGAGAGGTTAAT AAATCCAGACGGTAGTCTACTTTTCAGGGTGACTATTAA TGGAGTAACAGGTTGGAGATTGTGCGAGAGGATTTTGGC TTAG
IbpA	MYRKKLMRACLTGLMVVTVATANVAPVFAAEAPAEETVQ EPGAGAVGADVTVKFVDERGNVIDEKNLSFDEGIYNYEML DQYLPAGYERKSMGDFHALQGAVIEVRVMQKAPVEEVQN VKINYDEAAAGKQVAEVPVQVSSNASVVDMAILTRYLPAG YTLEGSDCIIRGGYVYVSVAPAEVQNVKINYDEAAKKQ VAEVPVQVSDTSYVDMAILTRYLPEGYTLEGSDCIIRGGYV YVSVAPAEVQNVKINYDEAAKKQVAEVPVQVSDTSYV

DMAILTRYLPKGYTIEGTDCIIRDGYVYVSVAPAEVQNVK
 INYYDEDAEKQVAEVPVQVSIDTSCVNMAILTRYMPEGYA
 LVSSDCIIRDGYVYVSVKKDVEIREAVLHITFETPNGEVVT
 ETVTAEGADGEDAVFRLGVDFNLPTGYKLSNDRDQVTEITI
 PFGSTGGHTMVVEKGDLSIVKIQFVDAENNDEVVAGGDY
 FVDGDGDGIFHTREITEWVPEGYELQEVDGFQVELYKETPL
 QLSVTKIKEDKPDTPDPEEPNKPEDPKEDTNNKDDKKEDT
 KKEDKKKNSPKTGDETSAAAAALPAGVSLAAILAVLVKKF
 K

APPENDIX C

Plasmid Maps and Sequencing Results

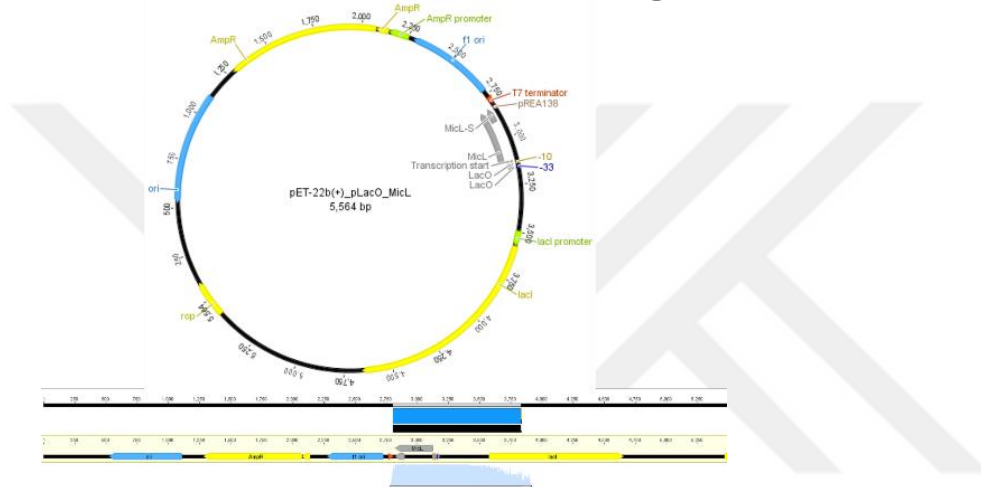


Figure C-1. Plasmid map of IPTG inducible micL sRNA in pET22b (+) vector and sequence verification of the vector with sanger sequencing.

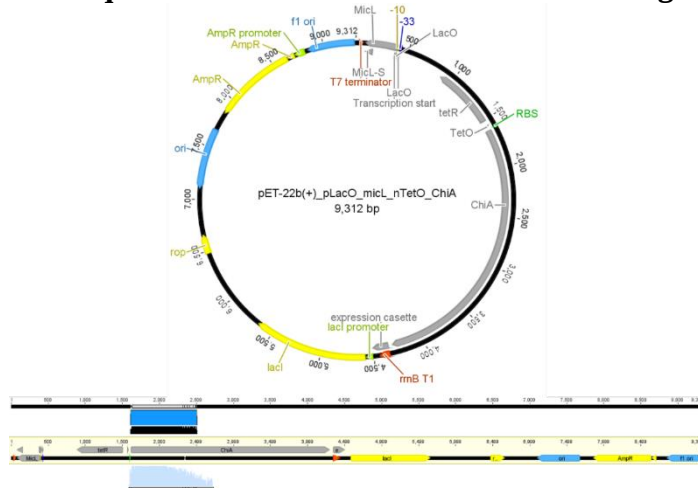


Figure C-2. Plasmid map of IPTG inducible micL sRNA and aTc inducible ChiA gene in pET22b (+) vector and sequence verification of the vector with sanger sequencing.

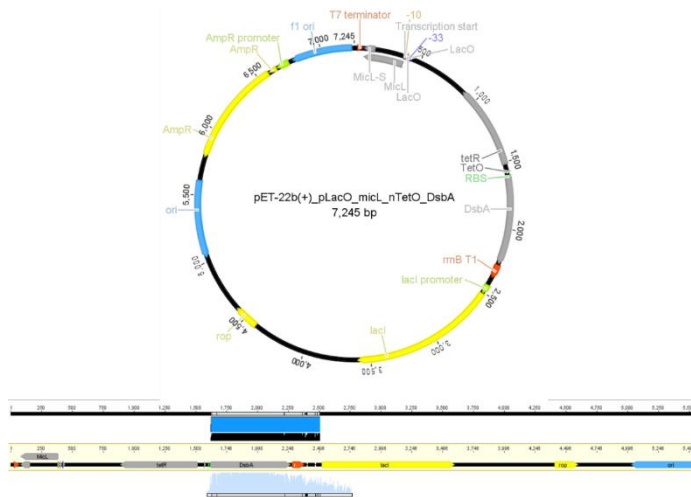


Figure C-3. Plasmid map of IPTG inducible micL sRNA and aTc inducible DsbA gene in pET22b (+) vector and sequence verification of the vector with sanger sequencing.

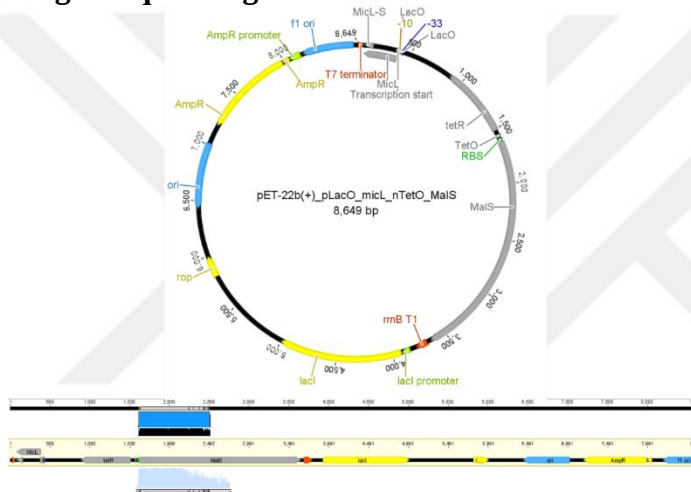


Figure C-4. Plasmid map of IPTG inducible micL sRNA and aTc inducible MalS gene in pET22b (+) vector and sequence verification of the vector with sanger sequencing.

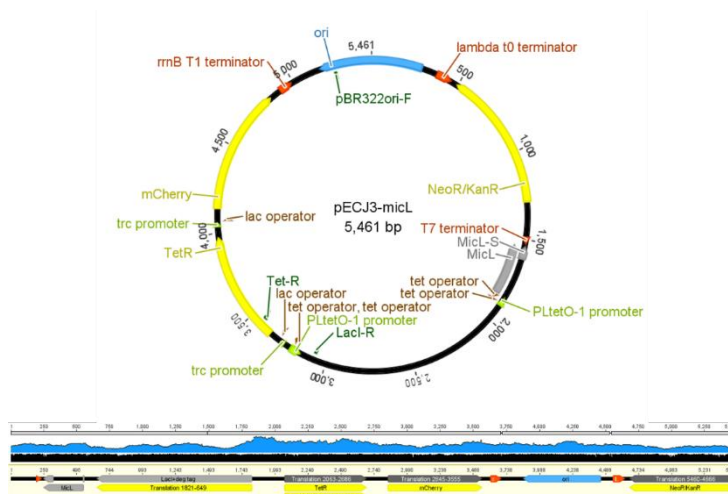


Figure C-5. Plasmid map of toggle switch circuit with micL in pECJ3 vector and sequence verification of the vector with NGS.

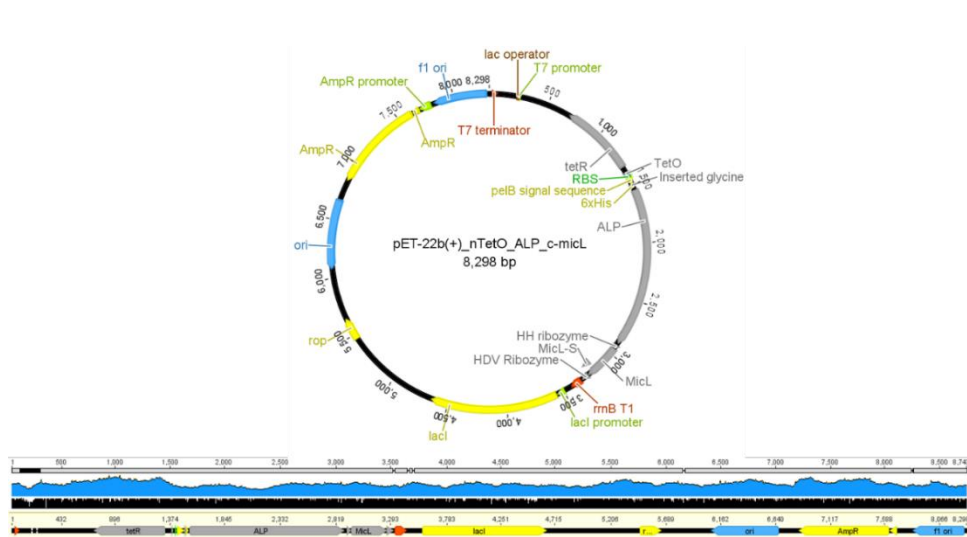


Figure C-6. Plasmid map of aTc inducible ALP (PhoA) gene fused with c-micL in pET22b (+) vector and sequence verification of the vector with NGS.

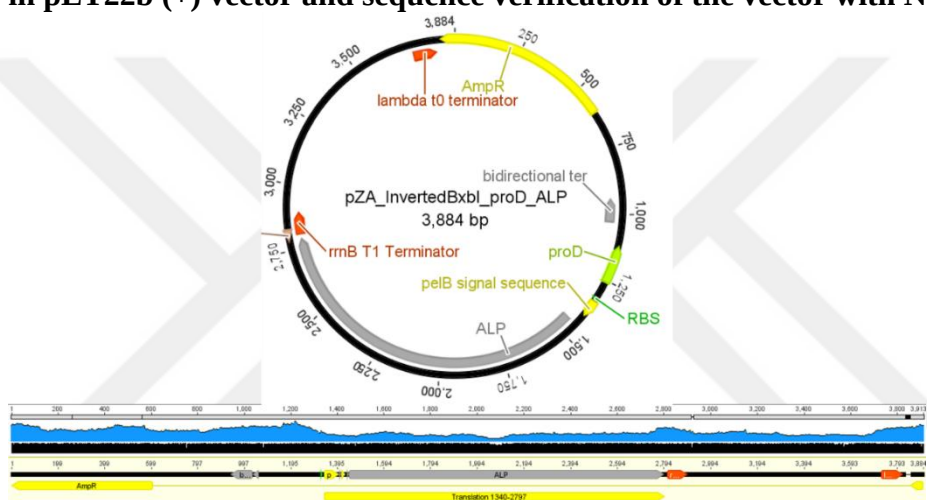


Figure C-7. Plasmid map of ALP (PhoA) gene located at inverted proD flanked by attB and attP sites of BxbI and sequence verification of the vector with NGS.

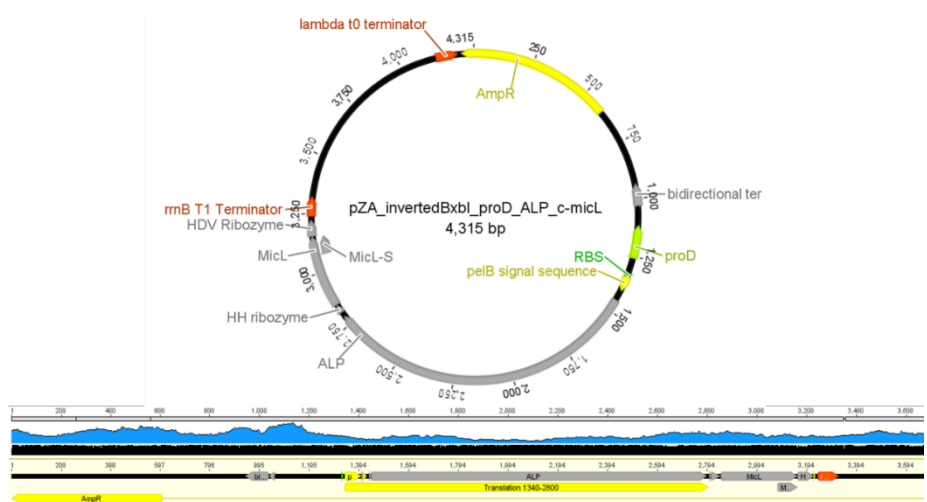


Figure C-8. Plasmid map of ALP (PhoA) gene fused with c-micL located at inverted proD flanked by attB and attP sites of BxbI and sequence verification of the vector with NGS.

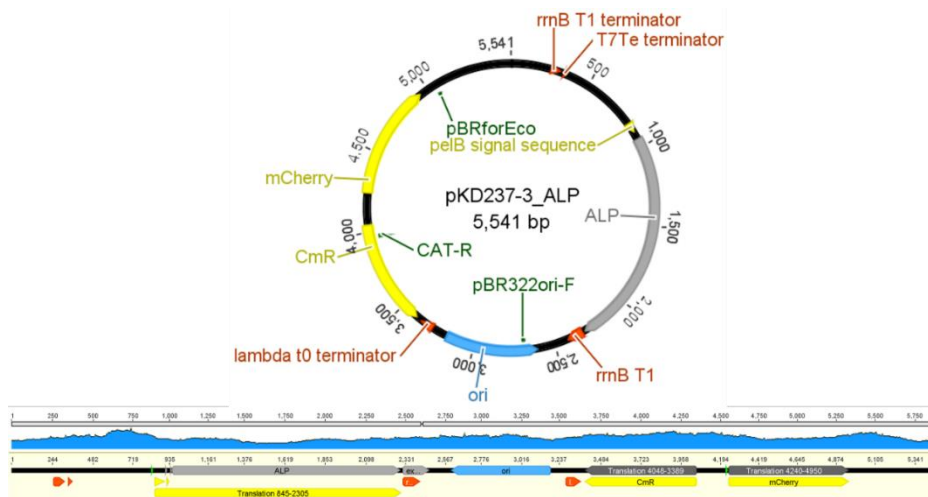


Figure C-9. Plasmid map of ALP (PhoA) cloned at the downstream of phsA in pKD237 vector and sequence verification of the vector with NGS.

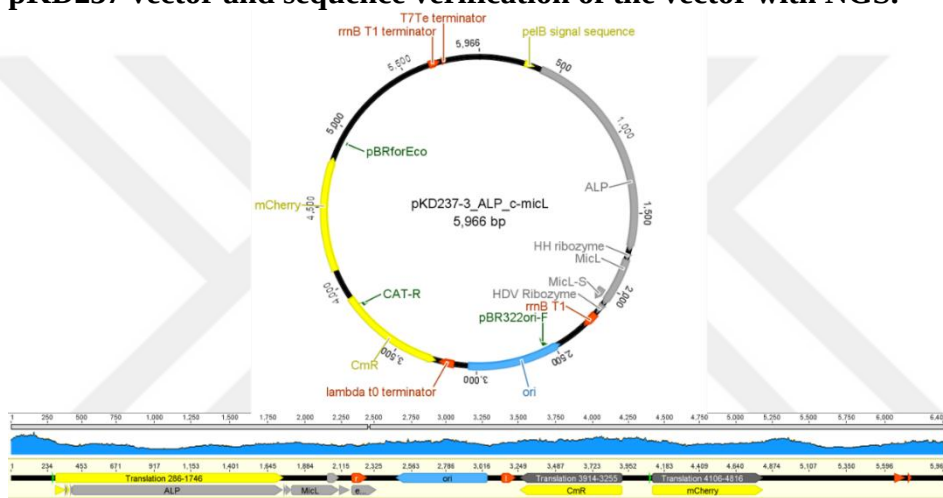


Figure C-10. Plasmid map of ALP (PhoA) fused with c-micL cloned at the downstream of phsA in pKD237 vector and sequence verification of the vector with NGS.

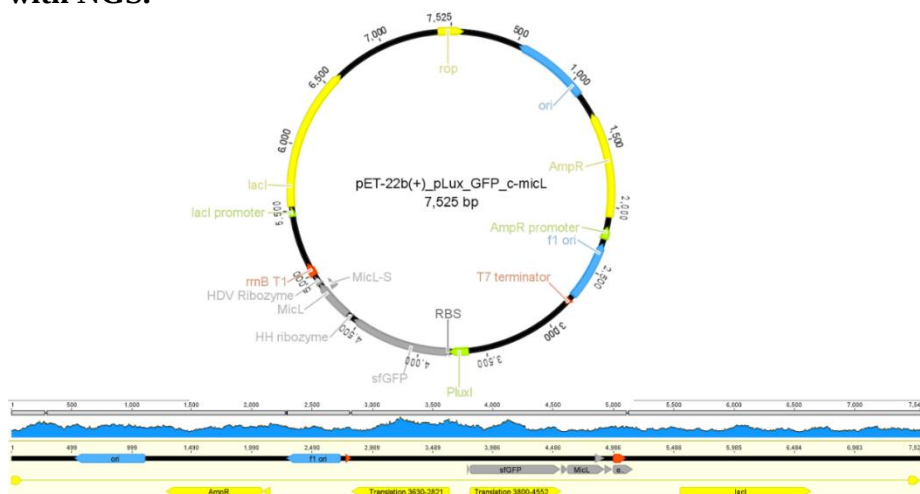


Figure C-11. Plasmid map of GFP fused with c-micL cloned downstream of pLux in pET22b (+) vector and sequence verification of the vector with NGS

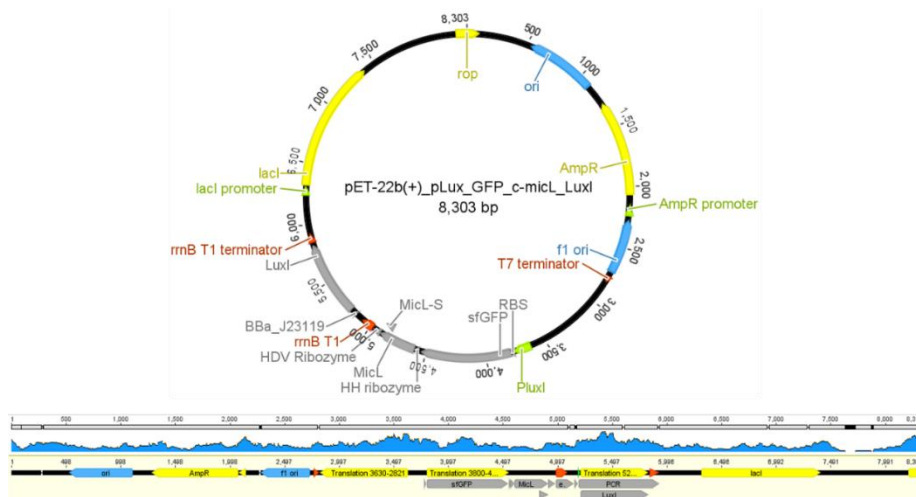


Figure C-12. Plasmid map of GFP fused with c-micL cloned at the downstream of pLux and constitutively expressed LuxI in pET22b (+) vector and sequence verification of the vector with NGS.

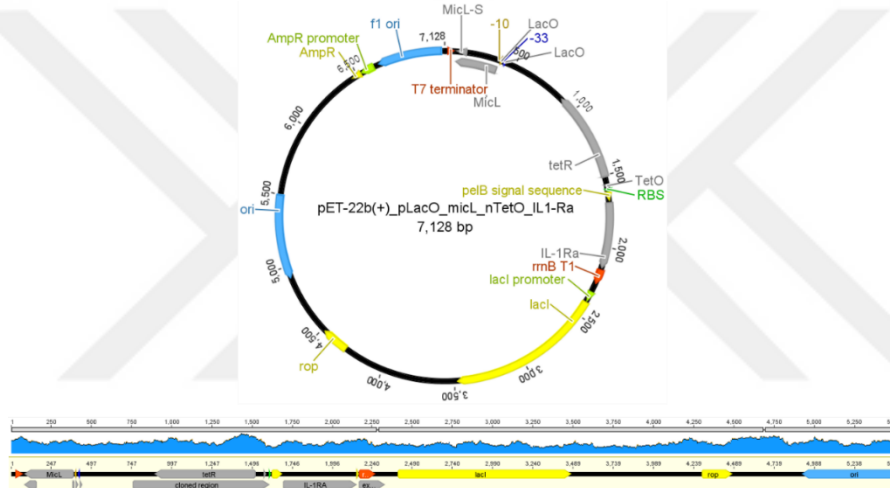


Figure C-13. Plasmid map of IPTG inducible micL sRNA and aTc inducible IL-1Ra gene in pET22b (+) vector and sequence verification of the vector with NGS.

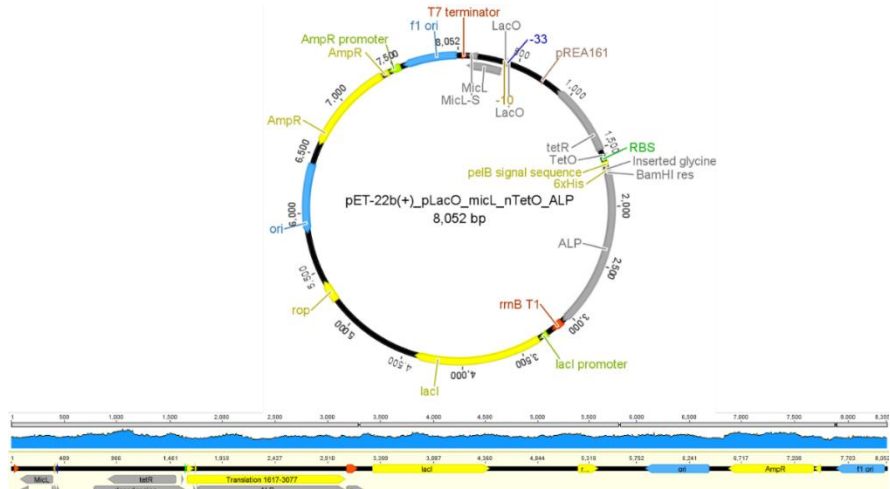


Figure C-14. Plasmid map of IPTG inducible micL sRNA and aTc inducible ALP (PhoA) gene in pET22b (+) vector and sequence verification of the vector with NGS.

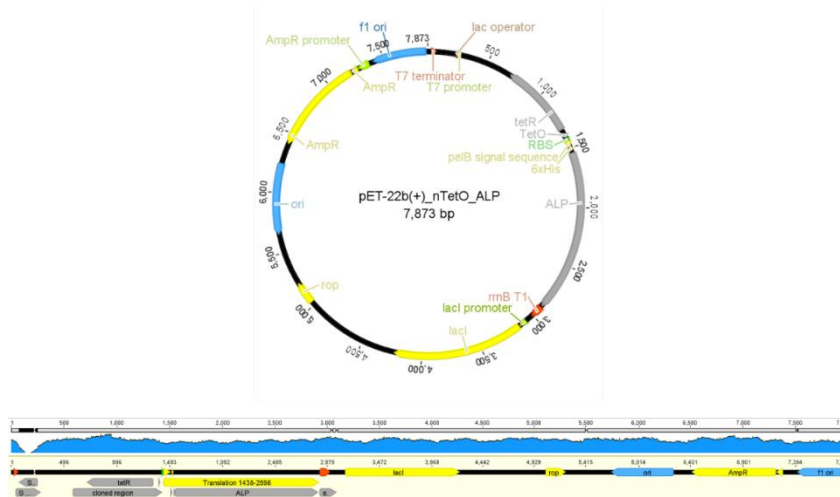


Figure C-15. Plasmid map of aTc inducible ALP (PhoA) gene in pET22b (+) vector and sequence verification of the vector with NGS.

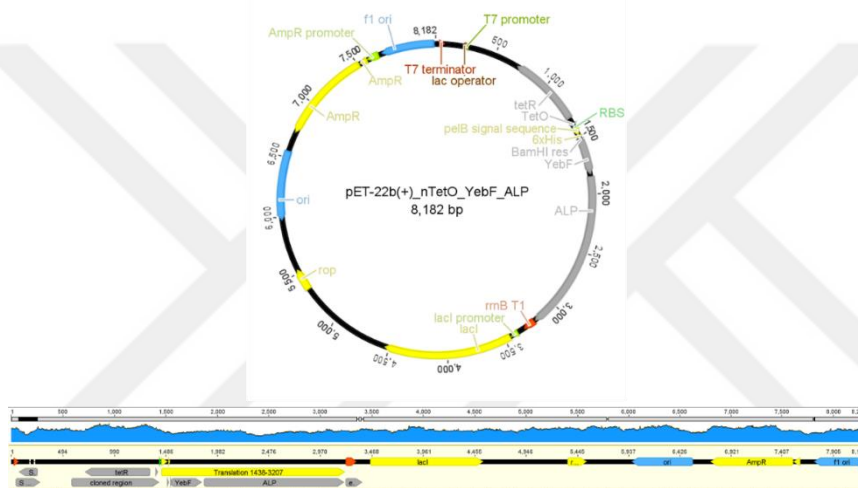


Figure C-16. Plasmid map of aTc inducible ALP (PhoA) gene fused with YebF gene in pET22b (+) vector and sequence verification of the vector with NGS.

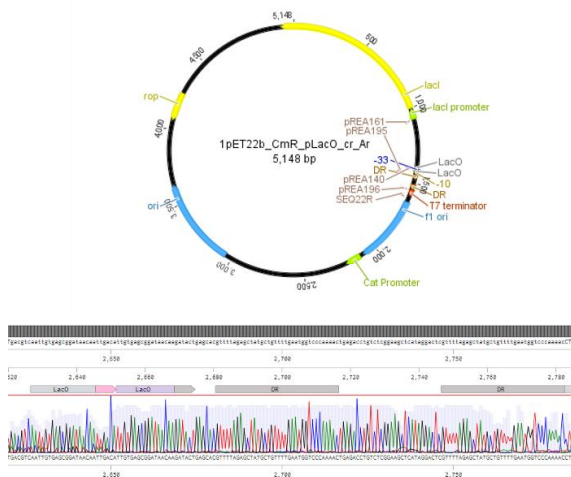


Figure C-17. Plasmid map of IPTG inducible CRISPR Array in pET22b (+) vector and sequence verification of the vector with Sanger.

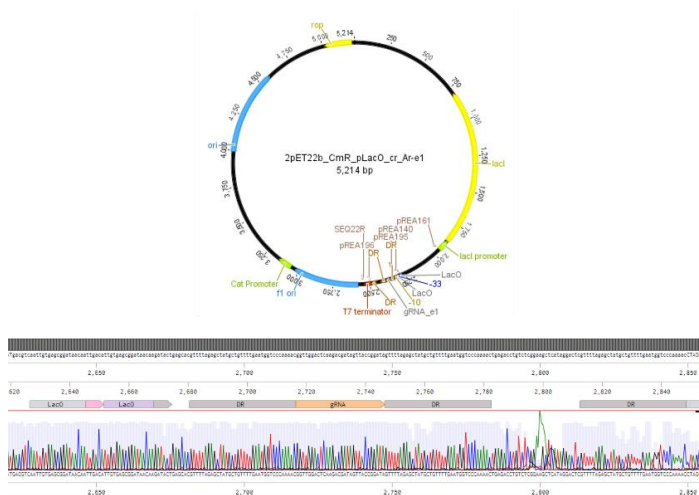


Figure C-18. Plasmid map of IPTG inducible CRISPR Array and E1 gRNA in pET22b (+) vector and sequence verification of the vector with Sanger.



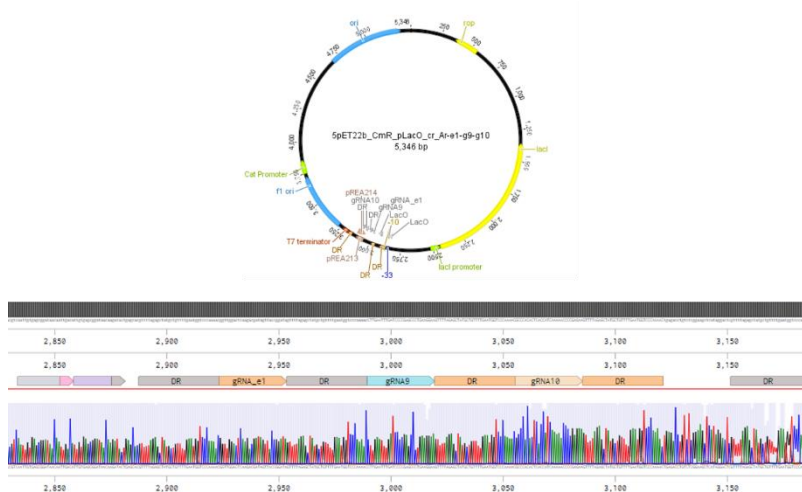


Figure C-21. Plasmid map of IPTG inducible CRISPR Array, and E1, g9 and g10 gRNA in pET22b (+) vector and sequence verification of the vector with Sanger.

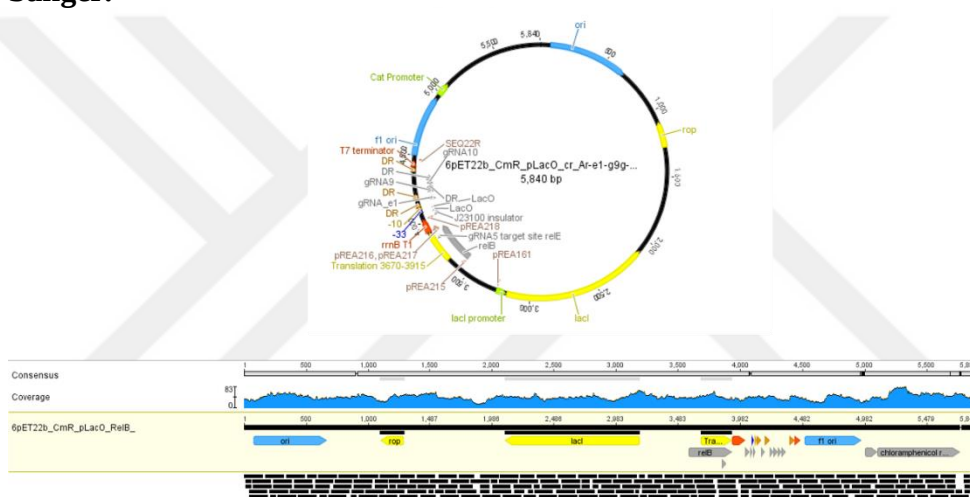
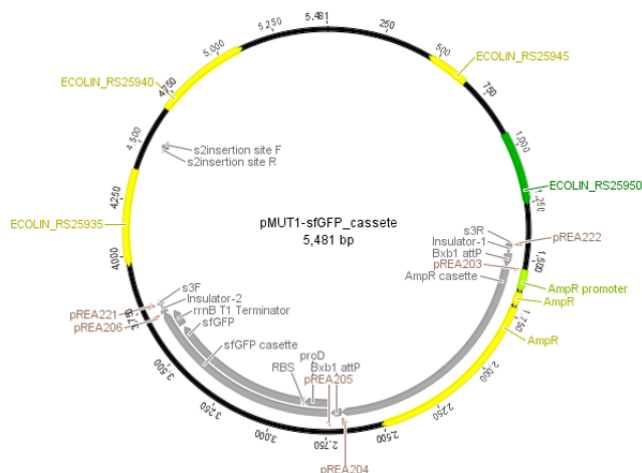
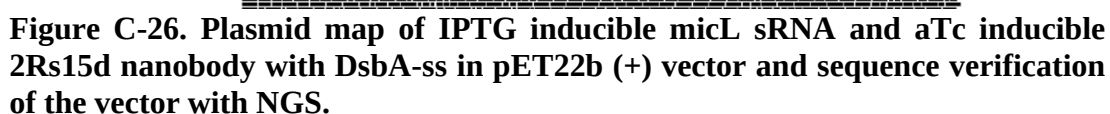
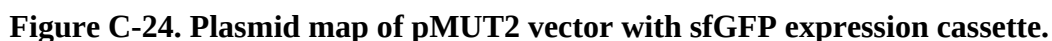


Figure C-22. Plasmid map of IPTG inducible CRISPR Array, relB and E1, g9 and g10 gRNA in pET22b (+) vector and sequence verification of the vector with NGS.





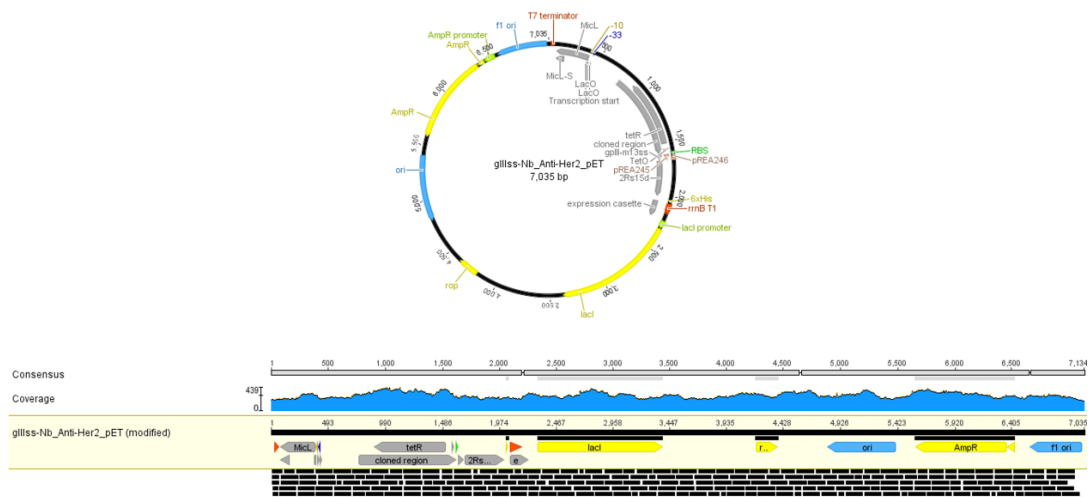


Figure C-27. Plasmid map of IPTG inducible micL sRNA and aTc inducible 2Rs15d nanobody with gIII-ss in pET22b (+) vector and sequence verification of the vector with NGS.

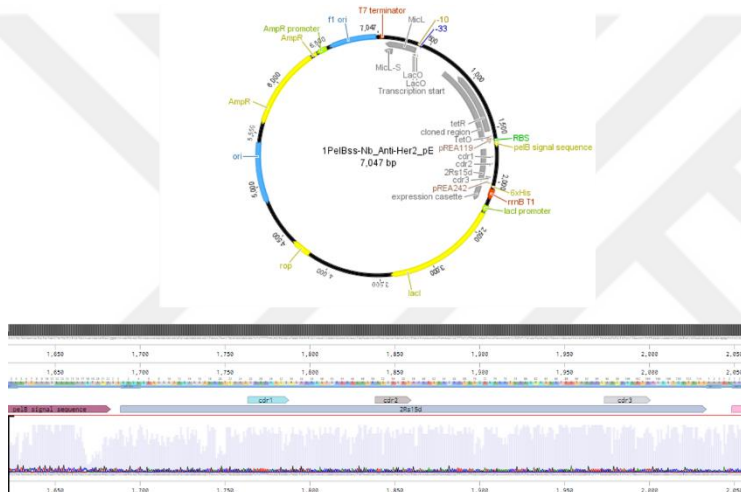


Figure C-28. Plasmid map of IPTG inducible micL sRNA and aTc inducible 2Rs15d nanobody with pelB-ss in pET22b (+) vector and sequence verification of the vector with Sanger.

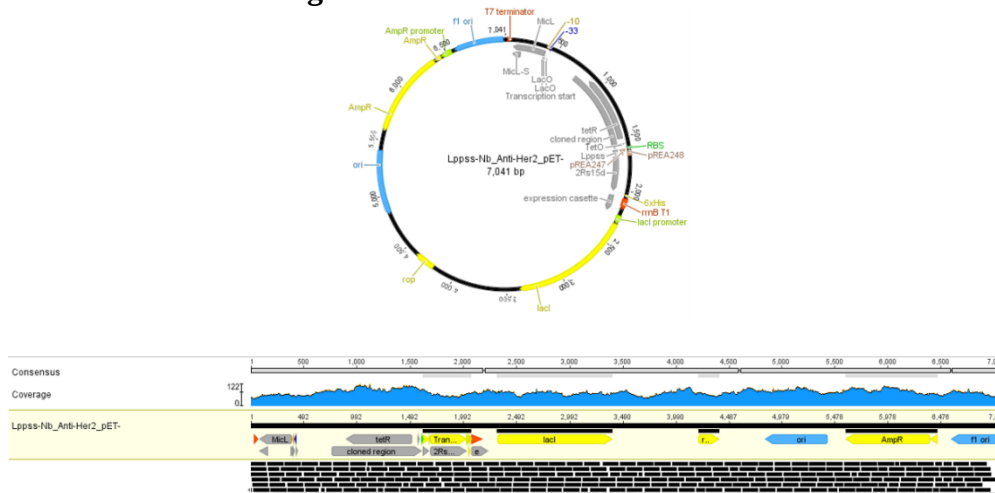


Figure C-29. Plasmid map of IPTG inducible micL sRNA and aTc inducible 2Rs15d nanobody with lpp-ss in pET22b (+) vector and sequence verification of the vector with NGS.

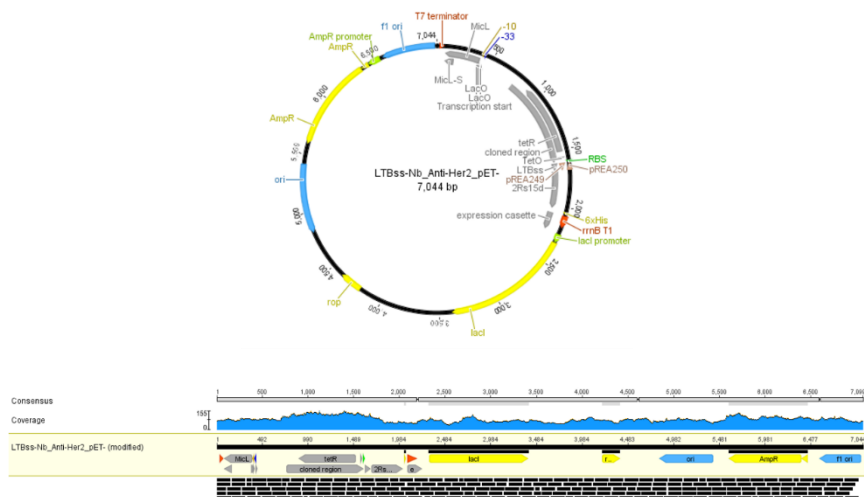


Figure C-30. Plasmid map of IPTG inducible micL sRNA and aTc inducible 2Rs15d nanobody with LTB-ss in pET22b (+) vector and sequence verification of the vector with NGS.

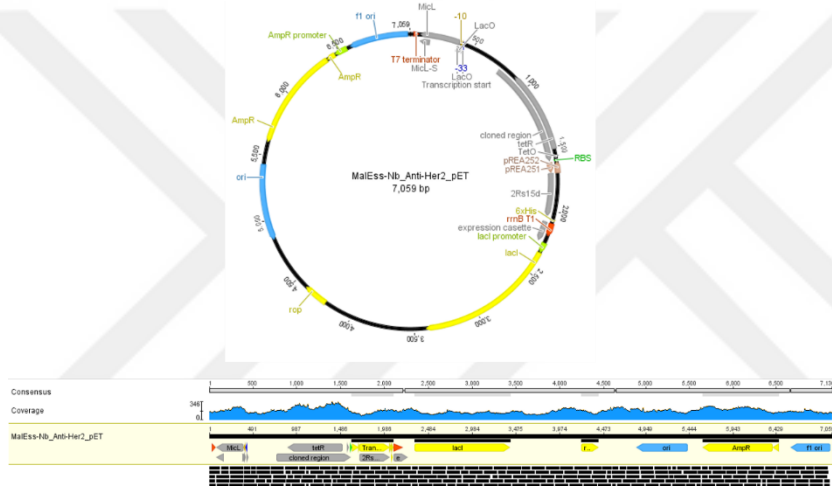


Figure C-31. Plasmid map of IPTG inducible micL sRNA and aTc inducible 2Rs15d nanobody with MalE-ss in pET22b (+) vector and sequence verification of the vector with NGS.

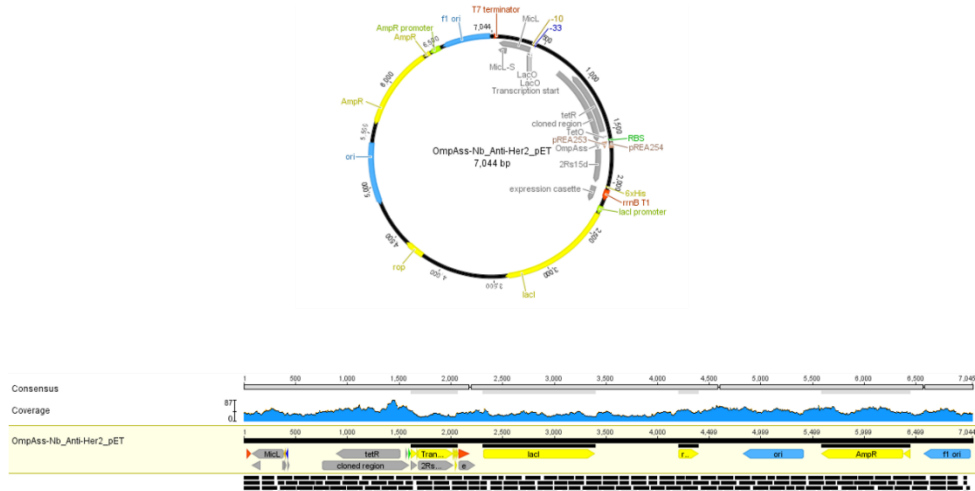


Figure C-32. Plasmid map of IPTG inducible micL sRNA and aTc inducible 2Rs15d nanobody with ompA-ss in pET22b (+) vector and sequence verification of the vector with NGS.

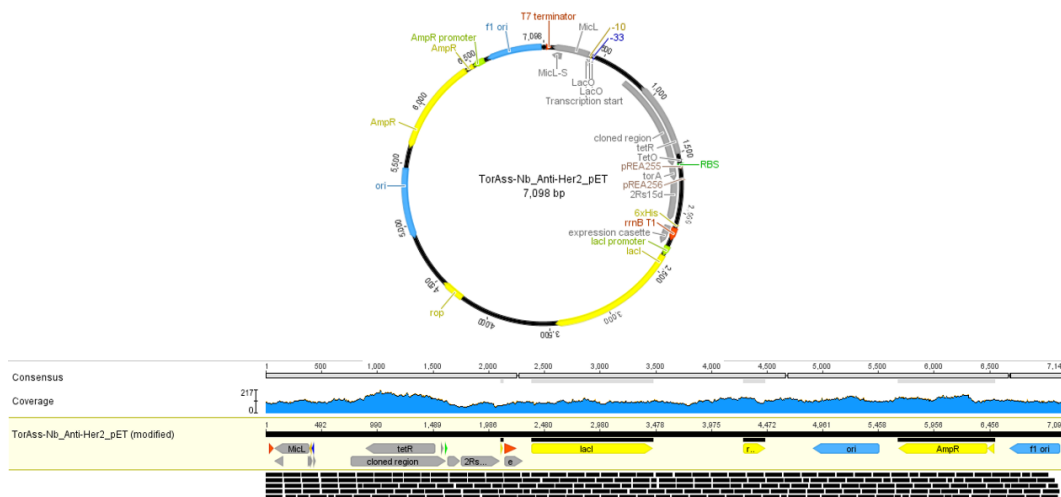


Figure C-33. Plasmid map of IPTG inducible micL sRNA and aTc inducible 2Rs15d nanobody with TorA-ss in pET22b (+) vector and sequence verification of the vector with NGS.

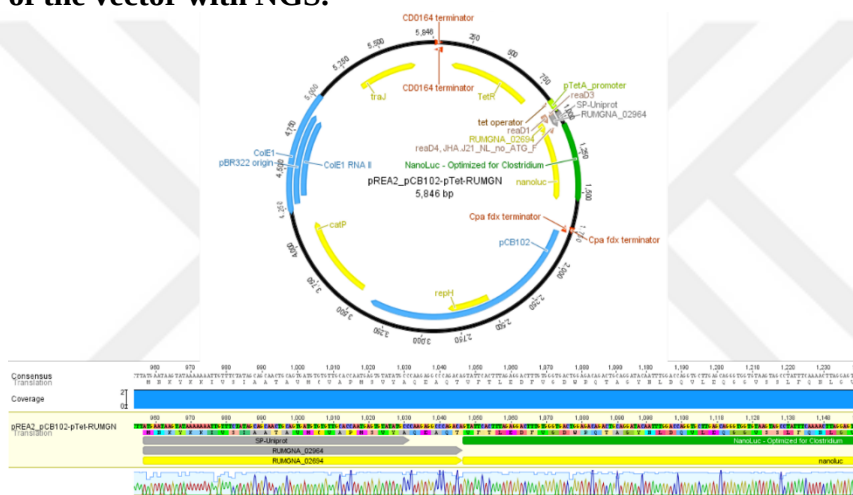


Figure C-34. Plasmid map of aTc inducible Nanoluciferase with ssREA-2 signal peptide in Lachnospiraceae shuttle vector and sequence verification of the vector with Sanger.



Figure C-35. Plasmid map of aTc inducible Nanoluciferase with ssREA-3 signal peptide in Lachnospiraceae shuttle vector and sequence verification of the vector with Sanger.

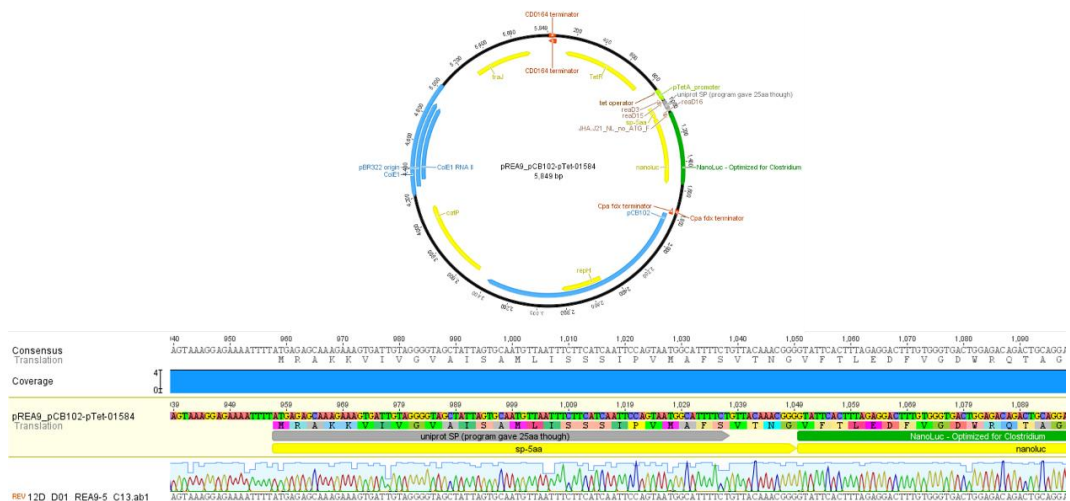


Figure C-36. Plasmid map of aTc inducible Nanoluciferase with ssREA-9 signal peptide in Lachnospiraceae shuttle vector and sequence verification of the vector with Sanger.

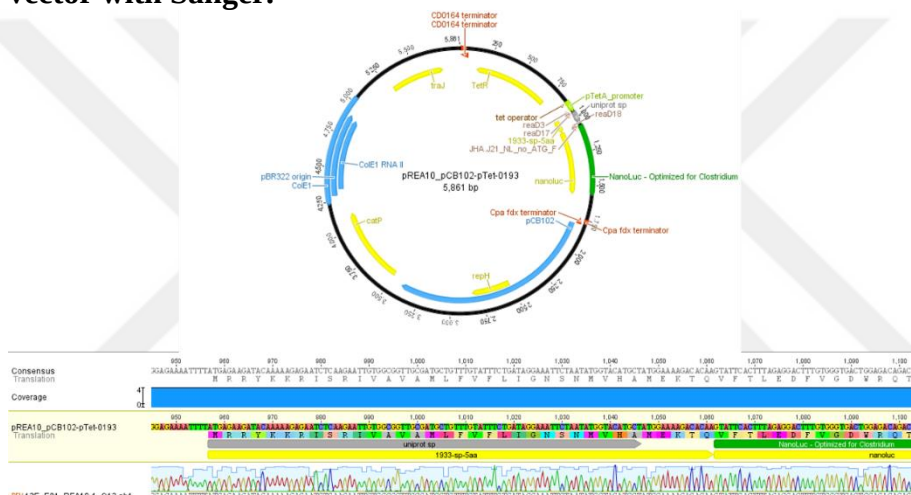


Figure C-37. Plasmid map of aTc inducible Nanoluciferase with ssREA-10 signal peptide in Lachnospiraceae shuttle vector and sequence verification of the vector with Sanger.

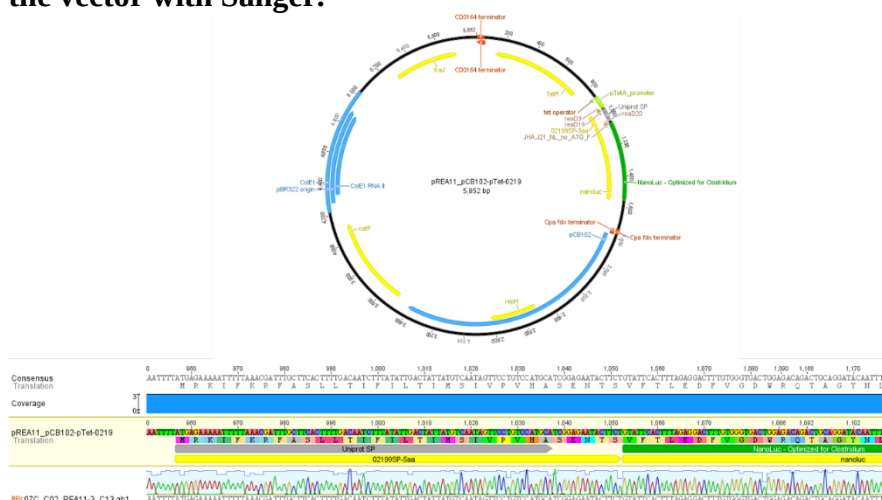


Figure C-38. Plasmid map of aTc inducible Nanoluciferase with ssREA-11 signal peptide in Lachnospiraceae shuttle vector and sequence verification of the vector with Sanger.



Figure C-39. Plasmid map of aTc inducible Nanoluciferase with ssREA-12 signal peptide in Lachnospiraceae shuttle vector and sequence verification of the vector with Sanger.

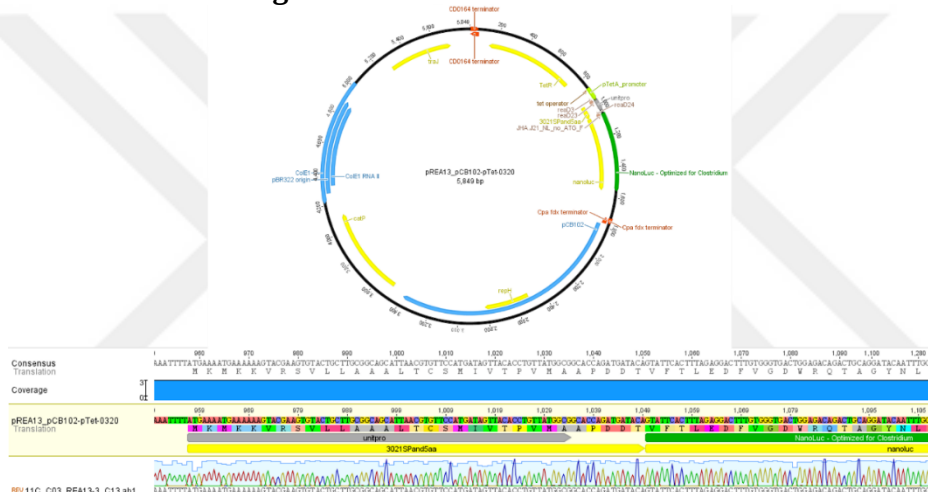


Figure C-40. Plasmid map of aTc inducible Nanoluciferase with ssREA-13 signal peptide in Lachnospiraceae shuttle vector and sequence verification of the vector with Sanger.

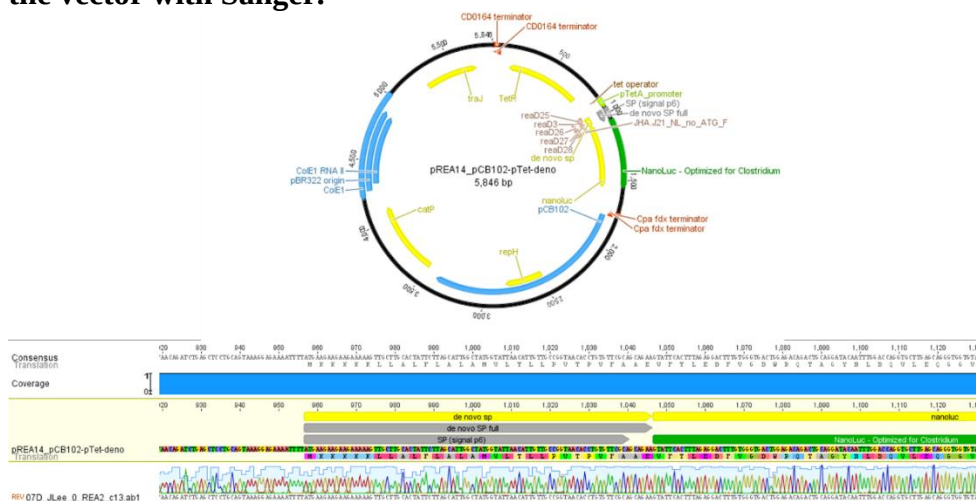


Figure C-41. Plasmid map of aTc inducible Nanoluciferase with ssREA-14 signal peptide in Lachnospiraceae shuttle vector and sequence verification of the vector with Sanger.

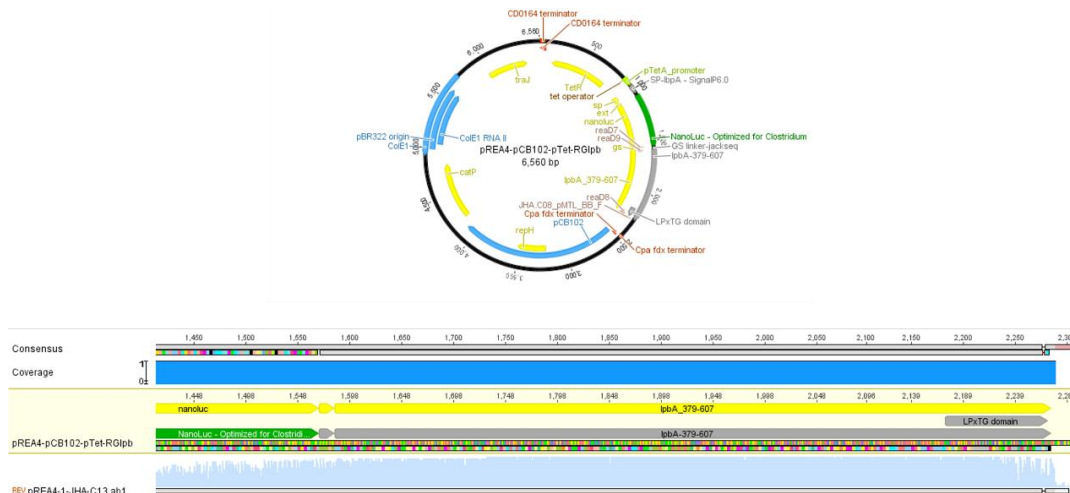


Figure C-42. Plasmid map of aTc inducible IbpA display shuttle 379-607 variant with ssREA-3 signal peptide and Nanoluciferase in Lachnospiraceae shuttle vector and sequence verification of the vector with Sanger.

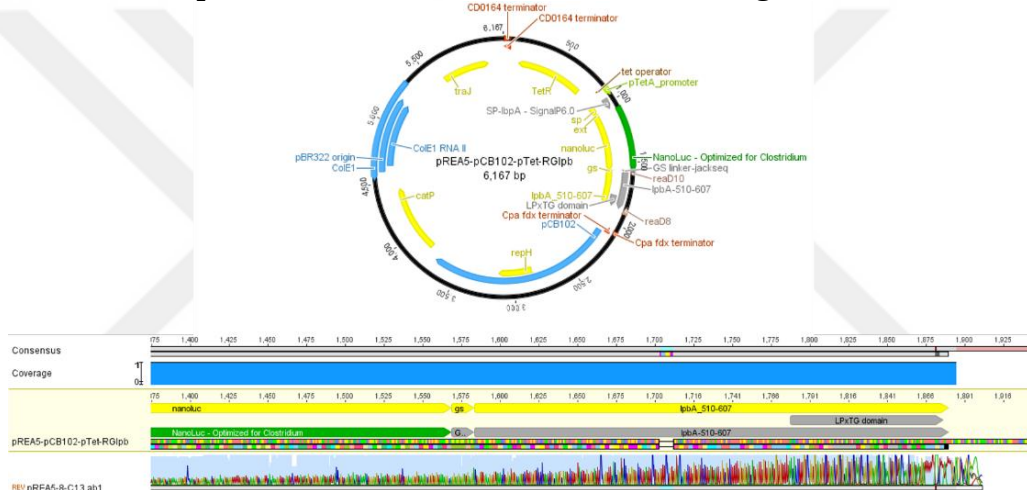


Figure C-43. Plasmid map of aTc inducible IbpA display shuttle 510-607 variant with ssREA-3 signal peptide and Nanoluciferase in Lachnospiraceae shuttle vector and sequence verification of the vector with Sanger.

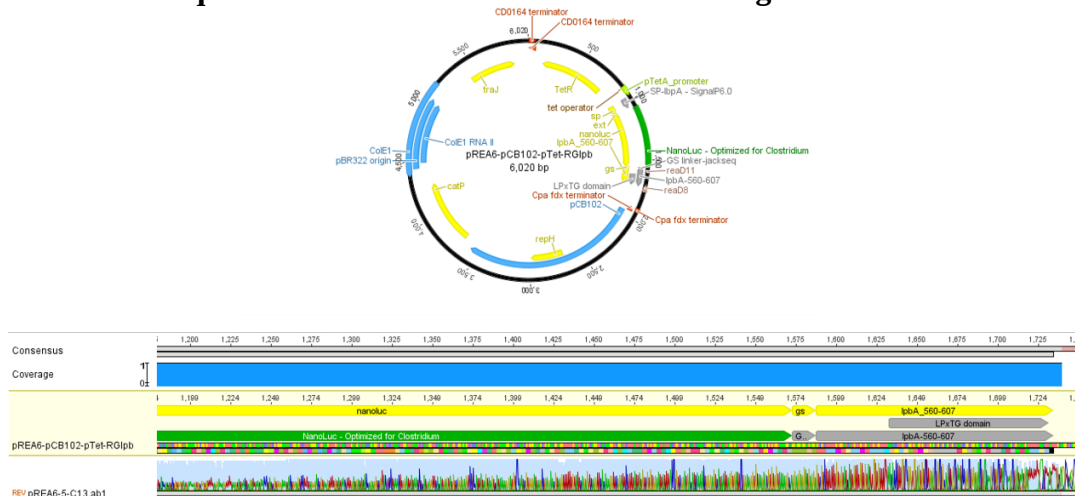


Figure C-44. Plasmid map of aTc inducible IbpA display shuttle 560-607 variant with ssREA-3 signal peptide and Nanoluciferase in Lachnospiraceae shuttle vector and sequence verification of the vector with Sanger.

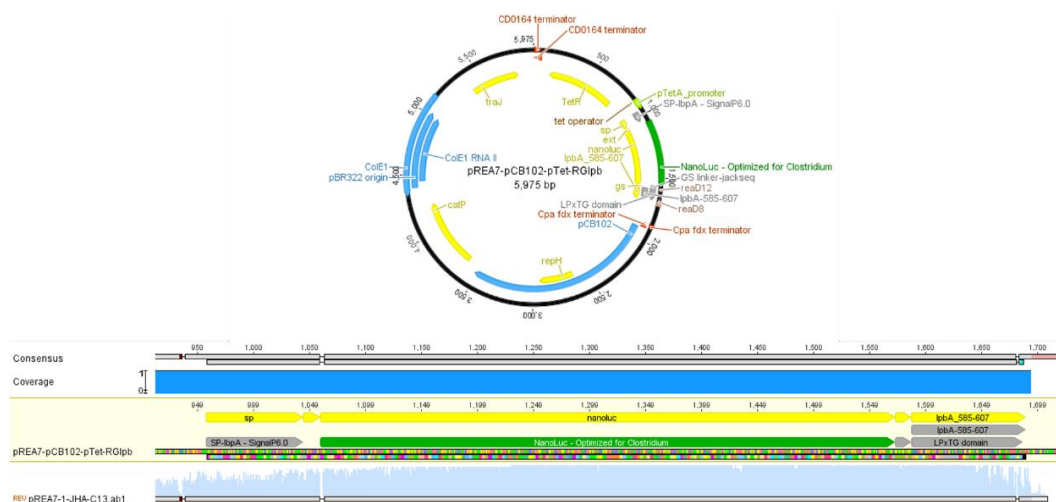


Figure C-45. Plasmid map of aTc inducible IbpA display shuttle 585-607 variant with ssREA-3 signal peptide and Nanoluciferase in Lachnospiraceae shuttle vector and sequence verification of the vector with Sanger.

APPENDIX D

General Recipes and Protocols

Table D-1. Q5 polymerase PCR conditions

Step	Temperature	Time
Initial Denaturation	98°C	30 seconds
Polymerization 25–35 Cycles	98°C 50–72°C 72°C	5–10 seconds 10–30 seconds 20–30 seconds/kb
Final Extension	72°C	2 minutes
Hold	4–10°C	

Table D-2. Gibson Assembly mix recipe (1.33x)

Substance	Amount
Taq ligase (40u/μl)	50μl
5x isothermal buffer*	100μl
T5 exonuclease (1u/μl)	2μl
Phusion polymerase (2u/μl)	6.25μl
Nuclease-free water	216.75μl

Table D-3. 5x isothermal buffer recipe

Substance	Amount
PEG-8000	25%
Tris-HCl	500mM, pH 7.5
MgCl ₂	50mM
DTT	50mM
NAD	5mM
dNTPs	1mM each

Table D-4. T4 ligation reaction recipe

Component	20µl reaction
T4 DNA Ligase Buffer (10X)	2µl
Vector DNA	50ng
Insert DNA	equal molar to vector DNA
Nuclease-free water	to 20µl
T4 DNA Ligase	1µl

Table D-5. Restriction Digestion Reaction

Component	20µl reaction
CutSmart (10X)	2µl
Vector DNA	1000-2000ng
Enzymes (each)	0.75µl
Nuclease-free water	to 20µl

APPENDIX F

Additional Data

Time Course Microscopy Images of Caco-2 Cells During Wound Healing

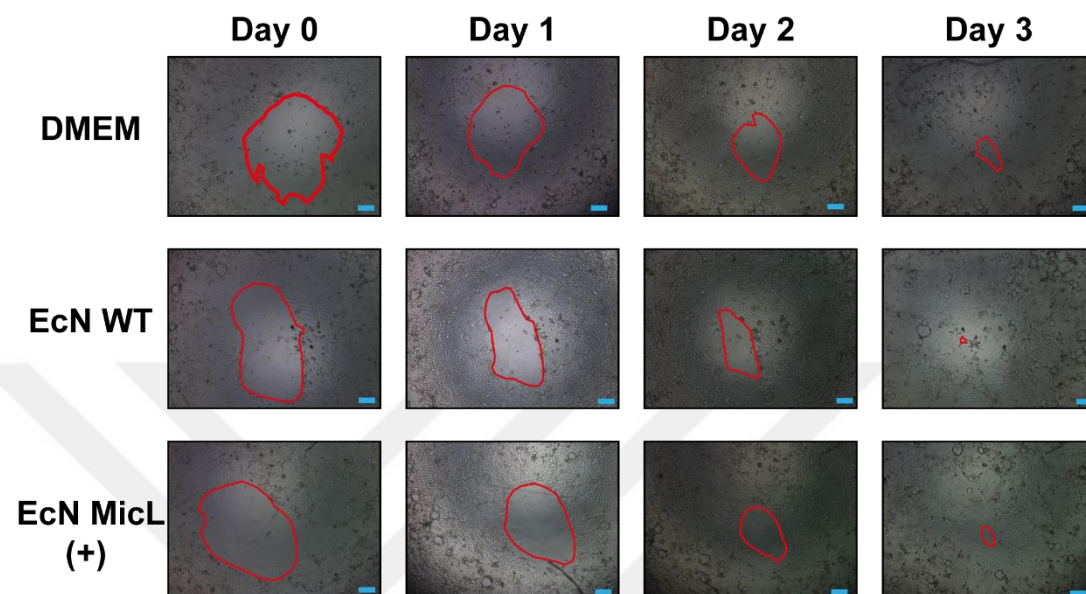


Figure F-1. Microscope images of holes in the migration assay. Holes were imaged in each day for three days. For each image, the area of hole is calculated using ImageJ software.

Intracellular Proteome and Secretome of *R. gnavus*

Table F-1. Proteins detected only in the intracellular soluble fraction

Accession	Description
A0A2N5NN11	4-hydroxy-tetrahydrodipicolinate synthase OS=Ruminococcus gnavus OX=33038 GN=dapA PE=3 SV=1
A0A2N5NK39	Thioredoxin OS=Ruminococcus gnavus OX=33038 GN=trxA PE=3 SV=1
A0A2N5NJ95	Aldose 1-epimerase OS=Ruminococcus gnavus OX=33038 GN=CDL18_06910 PE=3 SV=1
A0A829NSQ1	Co-chaperonin GroES OS=[Ruminococcus] gnavus CC55_001C OX=1073375 GN=groES PE=3 SV=1
A0A829NFV6	Deoxyribose-phosphate aldolase OS=[Ruminococcus] gnavus CC55_001C OX=1073375 GN=deoC PE=3 SV=1
A0A2N5NW99	Acetyl-CoA carboxylase biotin carboxyl carrier protein subunit OS=Ruminococcus gnavus OX=33038 GN=CDL20_11835 PE=4 SV=1
A0A2N5NS10	M3 family oligoendopeptidase OS=Ruminococcus gnavus OX=33038 GN=CDL23_07190 PE=3 SV=1
A0A2N5NRB4	Zinc ribbon domain-containing protein OS=Ruminococcus gnavus OX=33038 GN=CDL27_08945 PE=4 SV=1
A0A2N5Q349	PTS galactitol transporter subunit IIC OS=Ruminococcus gnavus OX=33038 GN=CDL20_02205 PE=4 SV=1
A0A829NX90	Trigger factor OS=[Ruminococcus] gnavus CC55_001C OX=1073375 GN=tig PE=3 SV=1
A7AZS1	Diaminopimelate epimerase OS=Ruminococcus gnavus (strain ATCC 29149 / VPI C7-9) OX=411470 GN=dapF PE=3 SV=1
A0A2N5NMI4	Threonine--tRNA ligase OS=Ruminococcus gnavus OX=33038 GN=thrS PE=3 SV=1
A0A2N5NMS0	S1 RNA-binding domain-containing protein OS=Ruminococcus gnavus OX=33038 GN=CDL18_00995 PE=4 SV=1
A0A2N5P2M5	Phosphopentomutase OS=Ruminococcus gnavus OX=33038 GN=deoB PE=3 SV=1
A0A396G6J8	Amino acid ABC transporter substrate-binding protein OS=Ruminococcus gnavus OX=33038 GN=DWZ70_05210 PE=4 SV=1
A0A2N5NN01	Aspartokinase OS=Ruminococcus gnavus OX=33038 GN=CCY17_00570 PE=3 SV=1
A0A2N5NMF3	Glutamate synthase (NADPH), homotetrameric OS=Ruminococcus gnavus OX=33038 GN=gltA PE=4 SV=1
A7AZR8	30S ribosomal protein S9 OS=Ruminococcus gnavus (strain ATCC 29149 / VPI C7-9) OX=411470 GN=rpsI PE=3 SV=1
A7B6S7	ATPase_AAA_core domain-containing protein OS=Ruminococcus gnavus (strain ATCC 29149 / VPI C7-9) OX=411470

	GN=RUMGNA_03291 PE=4 SV=1
A0A414P769	UDP-glucose 4-epimerase OS=Ruminococcus gnavus OX=33038 GN=galE PE=3 SV=1
A7B7K8	DJ-1 family protein OS=Ruminococcus gnavus (strain ATCC 29149 / VPI C7-9) OX=411470 GN=RUMGNA_03574 PE=4 SV=1
A0A3E4KA45	Aminopeptidase OS=Ruminococcus gnavus OX=33038 GN=DW270_07245 PE=3 SV=1
A0A2N5NL74	DUF3887 domain-containing protein OS=Ruminococcus gnavus OX=33038 GN=CDL18_03400 PE=4 SV=1
A0A2N5PTL1	50S ribosomal protein L18 OS=Ruminococcus gnavus OX=33038 GN=rplR PE=3 SV=1
A7B1B8	Polya polymerase OS=Ruminococcus gnavus (strain ATCC 29149 / VPI C7-9) OX=411470 GN=RUMGNA_01344 PE=4 SV=1
A0A2N5P2P5	Dihydroorotase OS=Ruminococcus gnavus OX=33038 GN=pyrC PE=3 SV=1
A7B8B9	Polyribonucleotide nucleotidyltransferase OS=Ruminococcus gnavus (strain ATCC 29149 / VPI C7-9) OX=411470 GN=pnp PE=3 SV=1
A7AZT8	carbamate kinase OS=Ruminococcus gnavus (strain ATCC 29149 / VPI C7-9) OX=411470 GN=RUMGNA_00812 PE=3 SV=1
A0A2N5NN12	Glucose-1-phosphate thymidyltransferase OS=Ruminococcus gnavus OX=33038 GN=rfaA PE=3 SV=1
A0A2N5PJS3	NUDIX domain-containing protein OS=Ruminococcus gnavus OX=33038 GN=CDL23_08165 PE=4 SV=1
A0A2N5PUG8	Transcription termination/antitermination protein NusG OS=Ruminococcus gnavus OX=33038 GN=nusG PE=3 SV=1
A0A2N5NKA8	dITP/XTP pyrophosphatase OS=Ruminococcus gnavus OX=33038 GN=CDL18_05165 PE=3 SV=1
A0A2N5PLS4	30S ribosomal protein S2 OS=Ruminococcus gnavus OX=33038 GN=rpsB PE=3 SV=1
A0A396G968	Class I SAM-dependent methyltransferase OS=Ruminococcus gnavus OX=33038 GN=DW142_04690 PE=4 SV=1
A0A2N5NIN7	50S ribosomal protein L13 OS=Ruminococcus gnavus OX=33038 GN=rplM PE=3 SV=1
A0A6N3A3H3	Glutamate dehydrogenase OS=Ruminococcus gnavus OX=33038 GN=gdhA PE=3 SV=1
A7B493	Acetylornithine aminotransferase OS=Ruminococcus gnavus (strain ATCC 29149 / VPI C7-9) OX=411470 GN=argD PE=3 SV=1
A0A2N5P6Z4	Pyruvate kinase OS=Ruminococcus gnavus OX=33038 GN=pyk PE=3 SV=1
A0A2N5P6Y4	Anaerobic ribonucleoside-triphosphate reductase-activating protein OS=Ruminococcus gnavus OX=33038 GN=nrdG PE=3 SV=1
A7B230	Ribose-5-phosphate isomerase B OS=Ruminococcus gnavus (strain ATCC 29149 / VPI C7-9) OX=411470 GN=rpiB PE=3 SV=1
A0A2N5PNX4	Phosphate-specific transport system accessory protein PhoU OS=Ruminococcus gnavus OX=33038 GN=phoU PE=3 SV=1
A0A2N5NFU7	Cytidylate kinase OS=Ruminococcus gnavus OX=33038 GN=CDL18_11855 PE=4 SV=1
A0A2N5NIP8	Sugar ABC transporter ATP-binding protein OS=Ruminococcus gnavus OX=33038 GN=CCY17_04900 PE=4 SV=1
A0A2N5NN24	Cofactor-independent phosphoglycerate mutase OS=Ruminococcus gnavus OX=33038 GN=CDL18_01545 PE=3 SV=1
A0A2N5NHX8	DNA-directed RNA polymerase subunit beta' OS=Ruminococcus gnavus OX=33038 GN=rpoC PE=3 SV=1
A0A2N5P963	Alanine--tRNA ligase OS=Ruminococcus gnavus OX=33038 GN=alaS PE=3 SV=1
A0A2N5PQD4	HU family DNA-binding protein OS=Ruminococcus gnavus OX=33038 GN=CCY17_10035 PE=3 SV=1
A7B7R4	Uridylate kinase OS=Ruminococcus gnavus (strain ATCC 29149 / VPI C7-9) OX=411470 GN=pyrH PE=3 SV=1
A0A829NTQ7	HPr domain-containing protein OS=[Ruminococcus] gnavus CC55_001C OX=1073375 GN=HMPREF1201_01490 PE=4 SV=1
A7B4K0	ATPase OS=Ruminococcus gnavus (strain ATCC 29149 / VPI C7-9) OX=411470 GN=RUMGNA_02484 PE=4 SV=1
A0A2N5PR79	50S ribosomal protein L35 OS=Ruminococcus gnavus OX=33038 GN=rplI PE=3 SV=1
A0A2N5NK61	Phosphoesterase OS=Ruminococcus gnavus OX=33038 GN=CDL18_05160 PE=3 SV=1
A0A6N3F8I5	Orotidine 5'-phosphate decarboxylase OS=Ruminococcus gnavus OX=33038 GN=pyrF PE=3 SV=1
A7AYI2	Creatinase OS=Ruminococcus gnavus (strain ATCC 29149 / VPI C7-9) OX=411470 GN=RUMGNA_00347 PE=4 SV=1
A0A2N5P8A9	Carboxyl transferase OS=Ruminococcus gnavus OX=33038 GN=CDL26_12410 PE=4 SV=1
A7B4R1	Purine nucleoside phosphorylase OS=Ruminococcus gnavus (strain ATCC 29149 / VPI C7-9) OX=411470 GN=RUMGNA_02545 PE=3 SV=1
A0A829NUI0	FAD-binding FR-type domain-containing protein OS=[Ruminococcus] gnavus CC55_001C OX=1073375 GN=HMPREF1201_00491 PE=4 SV=1
A7B6Z6	Threonine synthase OS=Ruminococcus gnavus (strain ATCC 29149 / VPI C7-9) OX=411470 GN=thrC PE=3 SV=1
A0A2N5NJ67	ATP synthase gamma chain OS=Ruminococcus gnavus OX=33038 GN=atpG PE=3 SV=1
A0A2N5NEL1	Imidazole glycerol phosphate synthase subunit HisF OS=Ruminococcus gnavus OX=33038 GN=hisF PE=3 SV=1
A0A2N5NIY3	Ornithine carbamoyltransferase OS=Ruminococcus gnavus OX=33038 GN=argF PE=3 SV=1
A0A2N5PWJ9	Argininosuccinate synthase OS=Ruminococcus gnavus OX=33038 GN=argG PE=3 SV=1
A0A829NTH2	FeS cluster assembly scaffold protein NifU OS=[Ruminococcus] gnavus CC55_001C OX=1073375 GN=HMPREF1201_02562 PE=4 SV=1

A0A2N5NII5	Glutamate synthase large subunit OS=Ruminococcus gnavus OX=33038 GN=CDL18_07225 PE=3 SV=1
A0A2N5PDU3	Oxidoreductase OS=Ruminococcus gnavus OX=33038 GN=CDL26_06405 PE=4 SV=1
A0A2N5NHV1	Asparagine synthase (Glutamine-hydrolyzing) OS=Ruminococcus gnavus OX=33038 GN=asnB PE=3 SV=1
A7B393	UPF0297 protein RUMGNA_02021 OS=Ruminococcus gnavus (strain ATCC 29149 / VPI C7-9) OX=411470 GN=RUMGNA_02021 PE=3 SV=1
A7B370	PTS system, glucose-like IIB component OS=Ruminococcus gnavus (strain ATCC 29149 / VPI C7-9) OX=411470 GN=RUMGNA_01998 PE=4 SV=1
A0A829NWL3	50S ribosomal protein L17 OS=[Ruminococcus] gnavus CC55_001C OX=1073375 GN=rplQ PE=3 SV=1
A0A2N5PIV0	Amidase OS=Ruminococcus gnavus OX=33038 GN=CCY17_00890 PE=4 SV=1
A0A2N5P9F8	Ketol-acid reductoisomerase (NADP(+)) OS=Ruminococcus gnavus OX=33038 GN=ilvC PE=3 SV=1
A0A2N5PH98	adenosylhomocysteine nucleosidase OS=Ruminococcus gnavus OX=33038 GN=CCY17_01750 PE=4 SV=1
A0A2N5NMN9	Cysteine synthase OS=Ruminococcus gnavus OX=33038 GN=cysK PE=3 SV=1
A0A829NPB5	Uracil phosphoribosyltransferase OS=[Ruminococcus] gnavus CC55_001C OX=1073375 GN=upp PE=3 SV=1
A7B2C5	DUF5717_N domain-containing protein OS=Ruminococcus gnavus (strain ATCC 29149 / VPI C7-9) OX=411470 GN=RUMGNA_01703 PE=4 SV=1
A0A829NL85	Cytidylate kinase-like family protein OS=[Ruminococcus] gnavus CC55_001C OX=1073375 GN=HMPREF1201_00242 PE=4 SV=1
A0A412C459	Transketolase OS=Ruminococcus gnavus OX=33038 GN=DWY88_07680 PE=4 SV=1

Table F-2. Proteins detected only in the cell free supernatant

T: Accession	T: Description
A0A2N5NIT6	LacI family transcriptional regulator OS=Ruminococcus gnavus OX=33038 GN=CDL18_07680 PE=4 SV=1
A0A412NNZ5	Glyco_hyd_65N_2 domain-containing protein OS=Ruminococcus gnavus OX=33038 GN=DWX36_01985 PE=4 SV=1
A0A829NUE0	Bacteriocin OS=[Ruminococcus] gnavus CC55_001C OX=1073375 GN=HMPREF1201_01301 PE=4 SV=1
A0A2N5P1J7	F5/8 type C domain-containing protein OS=Ruminococcus gnavus OX=33038 GN=CDL25_08230 PE=4 SV=1
A0A2N5PGN9	LacI family transcriptional regulator OS=Ruminococcus gnavus OX=33038 GN=CCY17_03375 PE=4 SV=1
A0A2N5PTZ2	BMP family ABC transporter substrate-binding protein OS=Ruminococcus gnavus OX=33038 GN=CCY17_06225 PE=3 SV=1
A0A2N5NNV7	ABC transporter substrate-binding protein OS=Ruminococcus gnavus OX=33038 GN=CCY17_09825 PE=4 SV=1
A0A2N5NIV1	Sulfatase domain-containing protein OS=Ruminococcus gnavus OX=33038 GN=CDL18_07720 PE=4 SV=1
A0A2N5NI63	Ribosome-recycling factor OS=Ruminococcus gnavus OX=33038 GN=frr PE=3 SV=1
A0A396G2D2	FctA domain-containing protein OS=Ruminococcus gnavus OX=33038 GN=DWZ70_12385 PE=4 SV=1
A0A2N5P1D8	Fibronectin type-III domain-containing protein OS=Ruminococcus gnavus OX=33038 GN=CDL25_08605 PE=4 SV=1
A0A2N5P7A7	LytR_cpsA_psr domain-containing protein OS=Ruminococcus gnavus OX=33038 GN=CDL25_01400 PE=3 SV=1
A7B520	30S ribosomal protein S13 OS=Ruminococcus gnavus (strain ATCC 29149 / VPI C7-9) OX=411470 GN=rpsM PE=3 SV=1
A0A2N5P333	F5/8 type C domain-containing protein (Fragment) OS=Ruminococcus gnavus OX=33038 GN=CDL25_05755 PE=4 SV=1
A0A2N5PGS7	Predicted outer membrane protein OS=Ruminococcus gnavus OX=33038 GN=CDL23_10050 PE=4 SV=1
A0A2N5NUL4	peptidylprolyl isomerase OS=Ruminococcus gnavus OX=33038 GN=CDL27_02905 PE=4 SV=1
A0A2N5NKD8	30S ribosomal protein S8 OS=Ruminococcus gnavus OX=33038 GN=rpsH PE=3 SV=1
A0A2N5NK74	UDP-galactopyranose mutase OS=Ruminococcus gnavus OX=33038 GN=glf PE=3 SV=1
A0A2N5NM00	SrtB family sortase OS=Ruminococcus gnavus OX=33038 GN=srtB PE=4 SV=1
A0A2N5Q0I3	50S ribosomal protein L24 OS=Ruminococcus gnavus OX=33038 GN=rplX PE=3 SV=1
A7B1I8	AB hydrolase-1 domain-containing protein OS=Ruminococcus gnavus (strain ATCC 29149 / VPI C7-9) OX=411470 GN=RUMGNA_01414 PE=4 SV=1

Table F-3. Proteins detected both in the intracellular soluble fraction and cell free supernatant

Accession	Description
A0A6N3D9I2	Pesticidal crystal protein cry22Aa OS=Ruminococcus gnavus OX=33038 GN=RGLFYP36_00965 PE=4 SV=1
A0A414S751	DUF5011 domain-containing protein (Fragment) OS=Ruminococcus gnavus OX=33038 GN=DW270_16380 PE=4 SV=1
A7B1Y0	Phosphoenolpyruvate carboxykinase (ATP) OS=Ruminococcus gnavus (strain ATCC 29149 / VPI C7-9) OX=411470 GN=pckA

	PE=3 SV=1
A0A412N DK9	NlpC/P60 domain-containing protein OS=Ruminococcus gnavus OX=33038 GN=DWX36_13205 PE=3 SV=1
A0A829N S27	SBP_bac_5 domain-containing protein OS=[Ruminococcus] gnavus CC55_001C OX=1073375 GN=HMPREF1201_01106 PE=4 SV=1
A0A415S D75	PDZ domain-containing protein OS=Ruminococcus gnavus OX=33038 GN=DWZ50_02355 PE=4 SV=1
A0A2N5P 545	GRAM_POS_ANCHORING domain-containing protein OS=Ruminococcus gnavus OX=33038 GN=CDL25_02540 PE=4 SV=1
A0A2N5P BX4	NlpC/P60 domain-containing protein OS=Ruminococcus gnavus OX=33038 GN=CDL23_12900 PE=3 SV=1
A0A2N5P DR7	Sialidase OS=Ruminococcus gnavus OX=33038 GN=CDL26_06400 PE=3 SV=1
A0A396G7 69	GRAM_POS_ANCHORING domain-containing protein OS=Ruminococcus gnavus OX=33038 GN=DWZ70_12785 PE=4 SV=1
A7B7G5	N-acetylmuramoyl-L-alanine amidase OS=Ruminococcus gnavus (strain ATCC 29149 / VPI C7-9) OX=411470 GN=RUMGNA_03531 PE=4 SV=1
A0A2N5P DC8	Pyruvate, phosphate dikinase OS=Ruminococcus gnavus OX=33038 GN=CDL26_07225 PE=3 SV=1
A0A2N5Q 1K7	Formate acetyltransferase OS=Ruminococcus gnavus OX=33038 GN=pflB PE=3 SV=1
A0A2N5P ZN6	Elongation factor Tu OS=Ruminococcus gnavus OX=33038 GN=tuf PE=3 SV=1
A7B798	Pyruvate synthase OS=Ruminococcus gnavus (strain ATCC 29149 / VPI C7-9) OX=411470 GN=nifJ PE=3 SV=1
A0A3E4K 4I5	phosphate acetyltransferase OS=Ruminococcus gnavus OX=33038 GN=pta PE=3 SV=1
A0A2N5P PH7	Peptidase M23 OS=Ruminococcus gnavus OX=33038 GN=CDL18_04840 PE=4 SV=1
A7B4U9	Meso-diaminopimelate D-dehydrogenase OS=Ruminococcus gnavus (strain ATCC 29149 / VPI C7-9) OX=411470 GN=RUMGNA_02583 PE=3 SV=1
A7B5T3	Stannin unstructured linker OS=Ruminococcus gnavus (strain ATCC 29149 / VPI C7-9) OX=411470 GN=RUMGNA_02929 PE=4 SV=1
A0A2N5P QT9	Glyceraldehyde-3-phosphate dehydrogenase OS=Ruminococcus gnavus OX=33038 GN=gap PE=3 SV=1
A0A829N SC6	Hemerythrin domain-containing protein OS=[Ruminococcus] gnavus CC55_001C OX=1073375 GN=HMPREF1201_01216 PE=3 SV=1
A0A2N5N J56	GRAM_POS_ANCHORING domain-containing protein OS=Ruminococcus gnavus OX=33038 GN=CDL18_06875 PE=4 SV=1
A7B3G8	Chaperonin GroEL OS=Ruminococcus gnavus (strain ATCC 29149 / VPI C7-9) OX=411470 GN=groL PE=3 SV=1
A0A2N5P 4N4	3-dehydro-L-gulonate 2-dehydrogenase OS=Ruminococcus gnavus OX=33038 GN=CDL25_03360 PE=4 SV=1
A7B342	Hydantoinase/oxoprolinase OS=Ruminococcus gnavus (strain ATCC 29149 / VPI C7-9) OX=411470 GN=RUMGNA_01970 PE=4 SV=1
A7B1R8	LPXTG-motif cell wall anchor domain protein OS=Ruminococcus gnavus (strain ATCC 29149 / VPI C7-9) OX=411470 GN=RUMGNA_01494 PE=4 SV=1
A0A829N GW6	Peripla_BP_4 domain-containing protein OS=[Ruminococcus] gnavus CC55_001C OX=1073375 GN=HMPREF1201_02375 PE=4 SV=1
A7B313	2,3-bisphosphoglycerate-independent phosphoglycerate mutase OS=Ruminococcus gnavus (strain ATCC 29149 / VPI C7-9) OX=411470 GN=gpmI PE=3 SV=1
A0A2N5N IT6	LacI family transcriptional regulator OS=Ruminococcus gnavus OX=33038 GN=CDL18_07680 PE=4 SV=1
A0A829N KQ3	AAA family ATPase OS=[Ruminococcus] gnavus CC55_001C OX=1073375 GN=HMPREF1201_02928 PE=4 SV=1
A7B6C5	Peptidase_C39_2 domain-containing protein OS=Ruminococcus gnavus (strain ATCC 29149 / VPI C7-9) OX=411470 GN=RUMGNA_03138 PE=4 SV=1
A0A2N5P 4R7	Triosephosphate isomerase OS=Ruminococcus gnavus OX=33038 GN=tpiA PE=3 SV=1
A0A396G DC8	Aldehyde-alcohol dehydrogenase OS=Ruminococcus gnavus OX=33038 GN=DW270_01310 PE=3 SV=1
A0A2N5P 1A7	NADH peroxidase OS=Ruminococcus gnavus OX=33038 GN=CCY17_12460 PE=4 SV=1
A0A6N3C LS4	Reverse rubrerythrin-1 OS=Ruminococcus gnavus OX=33038 GN=rbr3A PE=4 SV=1
A0A396G ES8	Cellulosome anchor protein OS=Ruminococcus gnavus OX=33038 GN=DW142_03975 PE=4 SV=1
A7B512	Cell Wall Hydrolase OS=Ruminococcus gnavus (strain ATCC 29149 / VPI C7-9) OX=411470 GN=RUMGNA_02649 PE=4 SV=1
A7B247	Fructose-1,6-bisphosphate aldolase, class II OS=Ruminococcus gnavus (strain ATCC 29149 / VPI C7-9) OX=411470 GN=fba PE=4 SV=1
A0A412N NZ5	Glyco_hyd_65N_2 domain-containing protein OS=Ruminococcus gnavus OX=33038 GN=DWX36_01985 PE=4 SV=1
A7B066	CARDB domain-containing protein OS=Ruminococcus gnavus (strain ATCC 29149 / VPI C7-9) OX=411470 GN=RUMGNA_00942 PE=4 SV=1
A0A414V0 R9	Phosphoglycerate kinase OS=Ruminococcus gnavus OX=33038 GN=pgk PE=3 SV=1
A7B375	Pyruvate formate-lyase-activating enzyme OS=Ruminococcus gnavus (strain ATCC 29149 / VPI C7-9) OX=411470 GN=pflA PE=3 SV=1
A0A412C AZ7	Chaperone protein DnaK OS=Ruminococcus gnavus OX=33038 GN=dnaK PE=2 SV=1
A0A829NI F0	YhcH/YjgK/YiaL family protein OS=[Ruminococcus] gnavus CC55_001C OX=1073375 GN=HMPREF1201_02524 PE=4 SV=1

A0A2N5P5R5	GGtGRT protein OS=Ruminococcus gnavus OX=33038 GN=CDL25_01480 PE=4 SV=1
A0A2N5NL02	Acetate kinase OS=Ruminococcus gnavus OX=33038 GN=ackA PE=3 SV=1
A0A829NL11	ATP synthase subunit alpha OS=[Ruminococcus] gnavus CC55_001C OX=1073375 GN=atpA PE=3 SV=1
A7B1Y8	50S ribosomal protein L7/L12 OS=Ruminococcus gnavus (strain ATCC 29149 / VPI C7-9) OX=411470 GN=rplL PE=3 SV=1
A0A2N5NW70	30S ribosomal protein S1 OS=Ruminococcus gnavus OX=33038 GN=CDL27_00895 PE=4 SV=1
A0A829NUE0	Bacteriocin OS=[Ruminococcus] gnavus CC55_001C OX=1073375 GN=HMPREF1201_01301 PE=4 SV=1
A0A2N5P1J7	F5/8 type C domain-containing protein OS=Ruminococcus gnavus OX=33038 GN=CDL25_08230 PE=4 SV=1
A0A829NP57	Peptidyl-prolyl cis-trans isomerase OS=[Ruminococcus] gnavus CC55_001C OX=1073375 GN=HMPREF1201_01829 PE=3 SV=1
A0A2N5NN11	4-hydroxy-tetrahydronicotinate synthase OS=Ruminococcus gnavus OX=33038 GN=dapA PE=3 SV=1
A0A2N5PGV8	Cell division protein FtsZ OS=Ruminococcus gnavus OX=33038 GN=ftsZ PE=3 SV=1
A7B1Y6	50S ribosomal protein L1 OS=Ruminococcus gnavus (strain ATCC 29149 / VPI C7-9) OX=411470 GN=rplA PE=3 SV=1
A0A396G8V2	Rubrythrin family protein OS=Ruminococcus gnavus OX=33038 GN=DW142_02205 PE=4 SV=1
A0A2N5P6S2	Butanol dehydrogenase OS=Ruminococcus gnavus OX=33038 GN=CDL25_00310 PE=4 SV=1
A0A2N5PGN9	LacI family transcriptional regulator OS=Ruminococcus gnavus OX=33038 GN=CCY17_03375 PE=4 SV=1
A0A2N5PE79	Mannosate dehydratase OS=Ruminococcus gnavus OX=33038 GN=uxuA PE=3 SV=1
A0A829NLX4	30S ribosomal protein S3 OS=[Ruminococcus] gnavus CC55_001C OX=1073375 GN=rpsC PE=3 SV=1
A0A2N5NK39	Thioredoxin OS=Ruminococcus gnavus OX=33038 GN=trxA PE=3 SV=1
A0A2N5NSV6	30S ribosomal protein S5 OS=Ruminococcus gnavus OX=33038 GN=rpsE PE=3 SV=1
A7B537	50S ribosomal protein L29 OS=Ruminococcus gnavus (strain ATCC 29149 / VPI C7-9) OX=411470 GN=rpmC PE=3 SV=1
A0A2N5NHX4	Elongation factor G OS=Ruminococcus gnavus OX=33038 GN=fusA PE=3 SV=1
A0A2N5NL70	Aminoacyl-histidine dipeptidase OS=Ruminococcus gnavus OX=33038 GN=CDL18_03445 PE=4 SV=1
A7B7Y2	Redoxin family protein OS=Ruminococcus gnavus (strain ATCC 29149 / VPI C7-9) OX=411470 GN=RUMGNA_03698 PE=4 SV=1
A0A2N5NJ95	Aldose 1-epimerase OS=Ruminococcus gnavus OX=33038 GN=CDL18_06910 PE=3 SV=1
A0A2N5NKW0	Phosphoglucosyltransferase OS=Ruminococcus gnavus OX=33038 GN=CDL18_04580 PE=3 SV=1
A0A2N5NZG4	50S ribosomal protein L5 OS=Ruminococcus gnavus OX=33038 GN=rplE PE=3 SV=1
A7B726	branched-chain-amino-acid transaminase OS=Ruminococcus gnavus (strain ATCC 29149 / VPI C7-9) OX=411470 GN=ilvE PE=3 SV=1
A0A829NSQ1	Co-chaperonin GroES OS=[Ruminococcus] gnavus CC55_001C OX=1073375 GN=groES PE=3 SV=1
A0A2N5PIR9	dTDP-glucose 4,6-dehydratase OS=Ruminococcus gnavus OX=33038 GN=rfbB PE=3 SV=1
A0A2N5PTZ2	BMP family ABC transporter substrate-binding protein OS=Ruminococcus gnavus OX=33038 GN=CCY17_06225 PE=3 SV=1
A0A2N5PNP0	Phosphate ABC transporter substrate-binding protein OS=Ruminococcus gnavus OX=33038 GN=CCY17_05020 PE=4 SV=1
A0A829NFV6	Deoxyribose-phosphate aldolase OS=[Ruminococcus] gnavus CC55_001C OX=1073375 GN=deoC PE=3 SV=1
A0A2N5NSP2	30S ribosomal protein S7 OS=Ruminococcus gnavus OX=33038 GN=rpsG PE=3 SV=1
A0A2N5NW99	Acetyl-CoA carboxylase biotin carboxyl carrier protein subunit OS=Ruminococcus gnavus OX=33038 GN=CDL20_11835 PE=4 SV=1
A0A2N5PQF2	Transketolase OS=Ruminococcus gnavus OX=33038 GN=CCY17_10130 PE=4 SV=1
A0A2N5PSP7	50S ribosomal protein L6 OS=Ruminococcus gnavus OX=33038 GN=rplF PE=3 SV=1
A0A2N5PD47	Ribulose-phosphate 3-epimerase OS=Ruminococcus gnavus OX=33038 GN=rpe PE=3 SV=1
A7AZS5	CobW/P47K family protein OS=Ruminococcus gnavus (strain ATCC 29149 / VPI C7-9) OX=411470 GN=RUMGNA_00799 PE=4 SV=1
A0A2N5NS10	M3 family oligoendopeptidase OS=Ruminococcus gnavus OX=33038 GN=CDL23_07190 PE=3 SV=1
A0A2N5NRB4	Zinc ribbon domain-containing protein OS=Ruminococcus gnavus OX=33038 GN=CDL27_08945 PE=4 SV=1
A0A2N5NGT5	Poly(Hydroxyalcanoate) granule associated protein (Phasin) OS=Ruminococcus gnavus OX=33038 GN=CCY17_03480 PE=4 SV=1
A7B545	50S ribosomal protein L3 OS=Ruminococcus gnavus (strain ATCC 29149 / VPI C7-9) OX=411470 GN=rplC PE=3 SV=1
A0A2N5NKK7	Glucose-6-phosphate isomerase OS=Ruminococcus gnavus OX=33038 GN=pgi PE=3 SV=1

A0A2N5Q349	PTS galactitol transporter subunit IIC OS=Ruminococcus gnavus OX=33038 GN=CDL20_02205 PE=4 SV=1
A0A414UW10	TlpA family protein disulfide reductase OS=Ruminococcus gnavus OX=33038 GN=DW243_07695 PE=4 SV=1
A0A2N5PUX0	Hypoxanthine phosphoribosyltransferase OS=Ruminococcus gnavus OX=33038 GN=hpt PE=3 SV=1
A0A2N5NSW1	50S ribosomal protein L4 OS=Ruminococcus gnavus OX=33038 GN=rpID PE=3 SV=1
A0A2N5PAG5	50S ribosomal protein L11 OS=Ruminococcus gnavus OX=33038 GN=rpIK PE=3 SV=1
A0A2N5NHP5	Copper-translocating P-type ATPase OS=Ruminococcus gnavus OX=33038 GN=CDL18_09375 PE=3 SV=1
A0A829NVX5	NifU homolog involved in Fe-S cluster formation OS=[Ruminococcus] gnavus CC55_001C OX=1073375 GN=HMPREF1201_00315 PE=4 SV=1
A0A2N5PC53	Transcription elongation factor GreA OS=Ruminococcus gnavus OX=33038 GN=greA PE=3 SV=1
A0A2N5P2W4	50S ribosomal protein L19 OS=Ruminococcus gnavus OX=33038 GN=rpIS PE=3 SV=1
A0A412NGH6	Hydratase OS=Ruminococcus gnavus OX=33038 GN=DWX36_09975 PE=4 SV=1
A0A829NX90	Trigger factor OS=[Ruminococcus] gnavus CC55_001C OX=1073375 GN=tig PE=3 SV=1
A0A2N5NII2	LL-diaminopimelate aminotransferase OS=Ruminococcus gnavus OX=33038 GN=dapL PE=3 SV=1
A0A829NQF7	50S ribosomal protein L10 OS=[Ruminococcus] gnavus CC55_001C OX=1073375 GN=rpIJ PE=3 SV=1
A0A2N5PG65	Endosialidase OS=Ruminococcus gnavus OX=33038 GN=CCY17_05075 PE=4 SV=1
A0A2N5PPY0	Aminoglycoside phosphotransferase OS=Ruminococcus gnavus OX=33038 GN=CCY17_06535 PE=4 SV=1
A0A2N5NNV7	ABC transporter substrate-binding protein OS=Ruminococcus gnavus OX=33038 GN=CCY17_09825 PE=4 SV=1
A7AZS1	Diaminopimelate epimerase OS=Ruminococcus gnavus (strain ATCC 29149 / VPI C7-9) OX=411470 GN=dapF PE=3 SV=1
A0A2N5P039	Citrate transporter OS=Ruminococcus gnavus OX=33038 GN=CDL25_10245 PE=4 SV=1
A0A2N5PBU1	Inosine monophosphate cyclohydrolase OS=Ruminococcus gnavus OX=33038 GN=CCY17_10625 PE=4 SV=1
A0A396G851	peptidylprolyl isomerase OS=Ruminococcus gnavus OX=33038 GN=tig PE=4 SV=1
A0A2N5NMI4	Threonine--tRNA ligase OS=Ruminococcus gnavus OX=33038 GN=thrS PE=3 SV=1
A0A2N5NMS0	S1 RNA-binding domain-containing protein OS=Ruminococcus gnavus OX=33038 GN=CDL18_00995 PE=4 SV=1
A0A2N5NJ82	ATP synthase subunit beta OS=Ruminococcus gnavus OX=33038 GN=atpD PE=3 SV=1
A0A2N5NIV1	Sulfatase domain-containing protein OS=Ruminococcus gnavus OX=33038 GN=CDL18_07720 PE=4 SV=1
A0A2N5NFA6	S-adenosylmethionine synthase OS=Ruminococcus gnavus OX=33038 GN=metK PE=3 SV=1
A0A2N5Q0G5	30S ribosomal protein S10 OS=Ruminococcus gnavus OX=33038 GN=rpsJ PE=3 SV=1
A0A3E4K4N9	LPXTG cell wall anchor domain-containing protein OS=Ruminococcus gnavus OX=33038 GN=DW243_13490 PE=4 SV=1
A0A2N5P2M5	Phosphopentomutase OS=Ruminococcus gnavus OX=33038 GN=deoB PE=3 SV=1
A7B248	Glutamate--ammonia ligase, catalytic domain protein OS=Ruminococcus gnavus (strain ATCC 29149 / VPI C7-9) OX=411470 GN=glnA PE=3 SV=1
A7B777	Peptidase M16 inactive domain protein OS=Ruminococcus gnavus (strain ATCC 29149 / VPI C7-9) OX=411470 GN=RUMGNA_03443 PE=4 SV=1
A0A396G6J8	Amino acid ABC transporter substrate-binding protein OS=Ruminococcus gnavus OX=33038 GN=DWZ70_05210 PE=4 SV=1
A0A2N5NN01	Aspartokinase OS=Ruminococcus gnavus OX=33038 GN=CCY17_00570 PE=3 SV=1
A0A2N5PIP1	NAD-dependent malic enzyme OS=Ruminococcus gnavus OX=33038 GN=CDL23_08780 PE=3 SV=1
A0A2N5NMF3	Glutamate synthase (NADPH), homotetrameric OS=Ruminococcus gnavus OX=33038 GN=gltA PE=4 SV=1
A0A6N3FAC3	DNA-binding protein OS=Ruminococcus gnavus OX=33038 GN=RGLFYP19_02472 PE=4 SV=1
A7AZR8	30S ribosomal protein S9 OS=Ruminococcus gnavus (strain ATCC 29149 / VPI C7-9) OX=411470 GN=rpsI PE=3 SV=1
A7B1C0	ATP-dependent 6-phosphofructokinase OS=Ruminococcus gnavus (strain ATCC 29149 / VPI C7-9) OX=411470 GN=pfkA PE=3 SV=1
A7B6S7	ATPase_AAA_core domain-containing protein OS=Ruminococcus gnavus (strain ATCC 29149 / VPI C7-9) OX=411470 GN=RUMGNA_03291 PE=4 SV=1
A7B896	30S ribosomal protein S20 OS=Ruminococcus gnavus (strain ATCC 29149 / VPI C7-9) OX=411470 GN=rpsT PE=3 SV=1
A0A2N5P4D2	Carbamoyl-phosphate synthase large chain OS=Ruminococcus gnavus OX=33038 GN=carB PE=3 SV=1
A0A414P769	UDP-glucose 4-epimerase OS=Ruminococcus gnavus OX=33038 GN=galE PE=3 SV=1
A7B7K8	DJ-1 family protein OS=Ruminococcus gnavus (strain ATCC 29149 / VPI C7-9) OX=411470 GN=RUMGNA_03574 PE=4 SV=1

A0A3E4K A45	Aminopeptidase OS=Ruminococcus gnavus OX=33038 GN=DW270_07245 PE=3 SV=1
A0A2N5N I63	Ribosome-recycling factor OS=Ruminococcus gnavus OX=33038 GN=frr PE=3 SV=1
A7B3J7	Elongation factor Ts OS=Ruminococcus gnavus (strain ATCC 29149 / VPI C7-9) OX=411470 GN=tsf PE=3 SV=1
A0A2N5N MS3	Phosphoenolpyruvate-protein phosphotransferase OS=Ruminococcus gnavus OX=33038 GN=ptsP PE=3 SV=1
A0A2N5N L74	DUF3887 domain-containing protein OS=Ruminococcus gnavus OX=33038 GN=CDL18_03400 PE=4 SV=1
A0A829N H06	Phosphoribosylaminoimidazole-succinocarboxamide synthase OS=[Ruminococcus] gnavus CC55_001C OX=1073375 GN=purC PE=3 SV=1
A0A2N5P 984	30S ribosomal protein S21 OS=Ruminococcus gnavus OX=33038 GN=rpsU PE=3 SV=1
A0A2N5P TL1	50S ribosomal protein L18 OS=Ruminococcus gnavus OX=33038 GN=rplR PE=3 SV=1
A7B1B8	Polya polymerase OS=Ruminococcus gnavus (strain ATCC 29149 / VPI C7-9) OX=411470 GN=RUMGNA_01344 PE=4 SV=1
A0A2N5N ST4	DNA-directed RNA polymerase subunit alpha OS=Ruminococcus gnavus OX=33038 GN=rpoA PE=3 SV=1
A0A2N5P 2P5	Dihydroorotase OS=Ruminococcus gnavus OX=33038 GN=pyrC PE=3 SV=1
A7B527	50S ribosomal protein L30 OS=Ruminococcus gnavus (strain ATCC 29149 / VPI C7-9) OX=411470 GN=rpmD PE=3 SV=1
A0A2N5N I07	DNA-directed RNA polymerase subunit beta OS=Ruminococcus gnavus OX=33038 GN=rpoB PE=3 SV=1
A7B8B9	Polyribonucleotide nucleotidyltransferase OS=Ruminococcus gnavus (strain ATCC 29149 / VPI C7-9) OX=411470 GN=pnp PE=3 SV=1
A0A396G2 D2	FctA domain-containing protein OS=Ruminococcus gnavus OX=33038 GN=DWZ70_12385 PE=4 SV=1
A0A829N MK6	dTDP-4-dehydrorhamnose 3,5-epimerase OS=[Ruminococcus] gnavus CC55_001C OX=1073375 GN=HMPREF1201_00720 PE=3 SV=1
A7AZT8	carbamate kinase OS=Ruminococcus gnavus (strain ATCC 29149 / VPI C7-9) OX=411470 GN=RUMGNA_00812 PE=3 SV=1
A0A2N5P 1D8	Fibronectin type-III domain-containing protein OS=Ruminococcus gnavus OX=33038 GN=CDL25_08605 PE=4 SV=1
A0A2N5N N12	Glucose-1-phosphate thymidyltransferase OS=Ruminococcus gnavus OX=33038 GN=rfbA PE=3 SV=1
A0A2N5N MM3	NAD(P)H nitroreductase OS=Ruminococcus gnavus OX=33038 GN=CDL18_00660 PE=4 SV=1
A0A2N5PJ S3	NUDIX domain-containing protein OS=Ruminococcus gnavus OX=33038 GN=CDL23_08165 PE=4 SV=1
A0A2N5P 7A7	LytR_cpsA_psr domain-containing protein OS=Ruminococcus gnavus OX=33038 GN=CDL25_01400 PE=3 SV=1
A0A2N5Q 248	Phosphoribosylamine--glycine ligase OS=Ruminococcus gnavus OX=33038 GN=purD PE=3 SV=1
A0A2N5P UG8	Transcription termination/antitermination protein NusG OS=Ruminococcus gnavus OX=33038 GN=nusG PE=3 SV=1
A0A2N5N KA8	dITP/XTP pyrophosphatase OS=Ruminococcus gnavus OX=33038 GN=CDL18_05165 PE=3 SV=1
A0A2N5P LS4	30S ribosomal protein S2 OS=Ruminococcus gnavus OX=33038 GN=rpsB PE=3 SV=1
A0A396G9 68	Class I SAM-dependent methyltransferase OS=Ruminococcus gnavus OX=33038 GN=DW142_04690 PE=4 SV=1
A0A2N5N JN7	50S ribosomal protein L13 OS=Ruminococcus gnavus OX=33038 GN=rplM PE=3 SV=1
A0A2N5P QZ5	Adenylosuccinate synthetase OS=Ruminococcus gnavus OX=33038 GN=purA PE=3 SV=1
A0A6N3A 3H3	Glutamate dehydrogenase OS=Ruminococcus gnavus OX=33038 GN=gdhA PE=3 SV=1
A0A2N5N QM2	dTDP-4-dehydrorhamnose reductase OS=Ruminococcus gnavus OX=33038 GN=rfbD PE=3 SV=1
A7B493	Acetylmornithine aminotransferase OS=Ruminococcus gnavus (strain ATCC 29149 / VPI C7-9) OX=411470 GN=argD PE=3 SV=1
A0A2N5N LF4	50S ribosomal protein L21 OS=Ruminococcus gnavus OX=33038 GN=rplU PE=3 SV=1
A0A2N5Q 451	50S ribosomal protein L27 OS=Ruminococcus gnavus OX=33038 GN=rpmA PE=3 SV=1
A0A2N5P 6Z4	Pyruvate kinase OS=Ruminococcus gnavus OX=33038 GN=pyk PE=3 SV=1
A0A2N5P 6Y4	Anaerobic ribonucleoside-triphosphate reductase-activating protein OS=Ruminococcus gnavus OX=33038 GN=nrdG PE=3 SV=1
A7B230	Ribose-5-phosphate isomerase B OS=Ruminococcus gnavus (strain ATCC 29149 / VPI C7-9) OX=411470 GN=rpiB PE=3 SV=1
A0A2N5N SX9	30S ribosomal protein S19 OS=Ruminococcus gnavus OX=33038 GN=rpsS PE=3 SV=1
A0A2N5N GI5	Dihydroxyacetone kinase OS=Ruminococcus gnavus OX=33038 GN=CDL18_10950 PE=4 SV=1
A0A2N5P NX4	Phosphate-specific transport system accessory protein PhoU OS=Ruminococcus gnavus OX=33038 GN=phoU PE=3 SV=1
A0A2N5P T08	Single-stranded DNA-binding protein OS=Ruminococcus gnavus OX=33038 GN=ssb PE=3 SV=1
A0A2N5N FU7	Cytidylate kinase OS=Ruminococcus gnavus OX=33038 GN=CDL18_11855 PE=4 SV=1

A0A2N5N IP8	Sugar ABC transporter ATP-binding protein OS=Ruminococcus gnavus OX=33038 GN=CCY17_04900 PE=4 SV=1
A0A2N5N N24	Cofactor-independent phosphoglycerate mutase OS=Ruminococcus gnavus OX=33038 GN=CDL18_01545 PE=3 SV=1
A0A829N K78	50S ribosomal protein L14 OS=[Ruminococcus] gnavus CC55_001C OX=1073375 GN=rplN PE=3 SV=1
A0A2N5N H02	Peptidyl-prolyl cis-trans isomerase OS=Ruminococcus gnavus OX=33038 GN=CDL18_10100 PE=3 SV=1
A0A2N5N HX8	DNA-directed RNA polymerase subunit beta' OS=Ruminococcus gnavus OX=33038 GN=rpoC PE=3 SV=1
A0A2N5P 963	Alanine--tRNA ligase OS=Ruminococcus gnavus OX=33038 GN=alaS PE=3 SV=1
A0A2N5N SU5	Adenylate kinase OS=Ruminococcus gnavus OX=33038 GN=adk PE=3 SV=1
A0A2N5N I20	ABC transporter OS=Ruminococcus gnavus OX=33038 GN=CDL18_08890 PE=4 SV=1
A7B520	30S ribosomal protein S13 OS=Ruminococcus gnavus (strain ATCC 29149 / VPI C7-9) OX=411470 GN=rpsM PE=3 SV=1
A0A2N5P QD4	HU family DNA-binding protein OS=Ruminococcus gnavus OX=33038 GN=CCY17_10035 PE=3 SV=1
A0A2N5P 333	F5/8 type C domain-containing protein (Fragment) OS=Ruminococcus gnavus OX=33038 GN=CDL25_05755 PE=4 SV=1
A0A2N5N GK8	Oxaloacetate decarboxylase OS=Ruminococcus gnavus OX=33038 GN=CDL18_11120 PE=4 SV=1
A0A2N5P 400	30S ribosomal protein S12 OS=Ruminococcus gnavus OX=33038 GN=rpsL PE=3 SV=1
A7B7R4	Uridylate kinase OS=Ruminococcus gnavus (strain ATCC 29149 / VPI C7-9) OX=411470 GN=pyrH PE=3 SV=1
A7B212	Proline--tRNA ligase OS=Ruminococcus gnavus (strain ATCC 29149 / VPI C7-9) OX=411470 GN=proS PE=3 SV=1
A0A829NI W9	30S ribosomal protein S16 OS=[Ruminococcus] gnavus CC55_001C OX=1073375 GN=rpsP PE=3 SV=1
A0A829N TQ7	HPr domain-containing protein OS=[Ruminococcus] gnavus CC55_001C OX=1073375 GN=HMPREF1201_01490 PE=4 SV=1
A0A6N3E VA2	Protein RecA OS=Ruminococcus gnavus OX=33038 GN=recA PE=3 SV=1
A0A2N5N LA0	Pyridoxal phosphate homeostasis protein OS=Ruminococcus gnavus OX=33038 GN=CDL18_03600 PE=3 SV=1
A7B4K0	ATPase OS=Ruminococcus gnavus (strain ATCC 29149 / VPI C7-9) OX=411470 GN=RUMGNA_02484 PE=4 SV=1
A0A2N5P R79	50S ribosomal protein L35 OS=Ruminococcus gnavus OX=33038 GN=rpmI PE=3 SV=1
A0A2N5P LS3	Elongation factor P OS=Ruminococcus gnavus OX=33038 GN=efp PE=3 SV=1
A0A2N5P GS7	Predicted outer membrane protein OS=Ruminococcus gnavus OX=33038 GN=CDL23_10050 PE=4 SV=1
A0A2N5N K61	Phosphoesterase OS=Ruminococcus gnavus OX=33038 GN=CDL18_05160 PE=3 SV=1
A0A6N3F 8I5	Orotidine 5'-phosphate decarboxylase OS=Ruminococcus gnavus OX=33038 GN=pyrF PE=3 SV=1
A0A3E4U XK0	SLATT domain-containing protein OS=Ruminococcus gnavus OX=33038 GN=DXC31_14950 PE=4 SV=1
A0A396G3 Y8	histidine kinase OS=Ruminococcus gnavus OX=33038 GN=DW142_15370 PE=4 SV=1
A0A2N5N HK8	Phenylalanine--tRNA ligase beta subunit OS=Ruminococcus gnavus OX=33038 GN=pheT PE=3 SV=1
A7AYI2	Creatinase OS=Ruminococcus gnavus (strain ATCC 29149 / VPI C7-9) OX=411470 GN=RUMGNA_00347 PE=4 SV=1
A0A2N5P 8A9	Carboxyl transferase OS=Ruminococcus gnavus OX=33038 GN=CDL26_12410 PE=4 SV=1
A7B4R1	Purine nucleoside phosphorylase OS=Ruminococcus gnavus (strain ATCC 29149 / VPI C7-9) OX=411470 GN=RUMGNA_02545 PE=3 SV=1
A0A2N5N LA4	Penicillin-binding protein OS=Ruminococcus gnavus OX=33038 GN=CCY17_01560 PE=3 SV=1
A0A829N UI0	FAD-binding FR-type domain-containing protein OS=[Ruminococcus] gnavus CC55_001C OX=1073375 GN=HMPREF1201_00491 PE=4 SV=1
A7B6Z6	Threonine synthase OS=Ruminococcus gnavus (strain ATCC 29149 / VPI C7-9) OX=411470 GN=thrC PE=3 SV=1
A0A2N5N J67	ATP synthase gamma chain OS=Ruminococcus gnavus OX=33038 GN=atpG PE=3 SV=1
A0A2N5N EL1	Imidazole glycerol phosphate synthase subunit HisF OS=Ruminococcus gnavus OX=33038 GN=hisF PE=3 SV=1
A0A2N5N FL2	RNA-binding protein OS=Ruminococcus gnavus OX=33038 GN=CDL18_12660 PE=4 SV=1
A0A2N5N IY3	Ornithine carbamoyltransferase OS=Ruminococcus gnavus OX=33038 GN=argF PE=3 SV=1
A0A2N5N UL4	peptidylprolyl isomerase OS=Ruminococcus gnavus OX=33038 GN=CDL27_02905 PE=4 SV=1
A0A2N5P WJ9	Argininosuccinate synthase OS=Ruminococcus gnavus OX=33038 GN=argG PE=3 SV=1
A0A829N TH2	FeS cluster assembly scaffold protein NifU OS=[Ruminococcus] gnavus CC55_001C OX=1073375 GN=HMPREF1201_02562 PE=4 SV=1
A0A2N5N I15	Glutamate synthase large subunit OS=Ruminococcus gnavus OX=33038 GN=CDL18_07225 PE=3 SV=1

A0A2N5P DU3	Oxidoreductase OS=Ruminococcus gnavus OX=33038 GN=CDL26_06405 PE=4 SV=1
A0A2N5N HW2	Leucine--tRNA ligase OS=Ruminococcus gnavus OX=33038 GN=leuS PE=3 SV=1
A0A2N5N PV6	Sugar ABC transporter substrate-binding protein OS=Ruminococcus gnavus OX=33038 GN=CDL25_11860 PE=4 SV=1
A0A2N5N HV1	Asparagine synthase (Glutamine-hydrolyzing) OS=Ruminococcus gnavus OX=33038 GN=asnB PE=3 SV=1
A0A2N5P N50	Hydroxylamine reductase OS=Ruminococcus gnavus OX=33038 GN=hcp PE=3 SV=1
A7B393	UPF0297 protein RUMGNA_02021 OS=Ruminococcus gnavus (strain ATCC 29149 / VPI C7-9) OX=411470 GN=RUMGNA_02021 PE=3 SV=1
A7B370	PTS system, glucose-like IIB component OS=Ruminococcus gnavus (strain ATCC 29149 / VPI C7-9) OX=411470 GN=RUMGNA_01998 PE=4 SV=1
A0A829N WL3	50S ribosomal protein L17 OS=[Ruminococcus] gnavus CC55_001C OX=1073375 GN=rplQ PE=3 SV=1
A0A2N5N IH6	Glutamate synthase OS=Ruminococcus gnavus OX=33038 GN=CDL18_07230 PE=4 SV=1
A7AXY2	Cell division ATP-binding protein FtsE OS=Ruminococcus gnavus (strain ATCC 29149 / VPI C7-9) OX=411470 GN=ftsE PE=3 SV=1
A0A2N5N KD8	30S ribosomal protein S8 OS=Ruminococcus gnavus OX=33038 GN=rpsH PE=3 SV=1
A0A2N5PI V0	Amidase OS=Ruminococcus gnavus OX=33038 GN=CCY17_00890 PE=4 SV=1
A0A2N5P 9F8	Ketol-acid reductoisomerase (NADP(+)) OS=Ruminococcus gnavus OX=33038 GN=ilvC PE=3 SV=1
A0A2N5N Q37	Pyrophosphate--fructose 6-phosphate 1-phosphotransferase OS=Ruminococcus gnavus OX=33038 GN=pfp PE=3 SV=1
A0A2N5P H98	adenosylhomocysteine nucleosidase OS=Ruminococcus gnavus OX=33038 GN=CCY17_01750 PE=4 SV=1
A0A2N5N K74	UDP-galactopyranose mutase OS=Ruminococcus gnavus OX=33038 GN=glf PE=3 SV=1
A0A2N5N M00	SrtB family sortase OS=Ruminococcus gnavus OX=33038 GN=srtB PE=4 SV=1
A0A2N5N MN9	Cysteine synthase OS=Ruminococcus gnavus OX=33038 GN=cysK PE=3 SV=1
A0A3E4V 8M1	Zn_Tnp_IS91 domain-containing protein OS=Ruminococcus gnavus OX=33038 GN=DXC31_06690 PE=4 SV=1
A0A2N5Q 0I3	50S ribosomal protein L24 OS=Ruminococcus gnavus OX=33038 GN=rplX PE=3 SV=1
A0A829N PB5	Uracil phosphoribosyltransferase OS=[Ruminococcus] gnavus CC55_001C OX=1073375 GN=upp PE=3 SV=1
A7B2K4	ATP synthase subunit b OS=Ruminococcus gnavus (strain ATCC 29149 / VPI C7-9) OX=411470 GN=atpF PE=3 SV=1
A7B1I8	AB hydrolase-1 domain-containing protein OS=Ruminococcus gnavus (strain ATCC 29149 / VPI C7-9) OX=411470 GN=RUMGNA_01414 PE=4 SV=1
A0A2N5N QY8	ABC_transp_aux domain-containing protein OS=Ruminococcus gnavus OX=33038 GN=CDL27_09670 PE=4 SV=1
A7B2C5	DUF5717_N domain-containing protein OS=Ruminococcus gnavus (strain ATCC 29149 / VPI C7-9) OX=411470 GN=RUMGNA_01703 PE=4 SV=1
A0A829N L85	Cytidylate kinase-like family protein OS=[Ruminococcus] gnavus CC55_001C OX=1073375 GN=HMPREF1201_00242 PE=4 SV=1
A0A412C4 59	Transketolase OS=Ruminococcus gnavus OX=33038 GN=DWY88_07680 PE=4 SV=1
A0A6N3F QB2	Coproporphyrinogen III oxidase OS=Ruminococcus gnavus OX=33038 GN=RGLFYP36_01962 PE=4 SV=1
A7AZ98	50S ribosomal protein L20 OS=Ruminococcus gnavus (strain ATCC 29149 / VPI C7-9) OX=411470 GN=rplT PE=3 SV=1