Towards *In Vitro* Synthesis of Oxygen-tolerant CpI Hydrogenase Variants for Photosynthetic Hydrogen Production

by

Selin Turan

A Dissertation Submitted to the
Graduate School of Sciences and Engineering
in Partial Fulfillment of the Requirements for
the Degree of

Master of Science

in

Molecular Biology and Genetics

KOÇ
ÜNİVERSİTESİ

September 5, 2022

Towards *In Vitro* Synthesis of Oxygen-tolerant CpI Hydrogenase Variants for Photosynthetic Hydrogen Production

Koç University

Graduate School of Sciences and Engineering

This is to certify that I have examined this copy of a master's thesis by

Selin Turan

and have found that it is complete and satisfactory in all respects,
and that any and all revisions required by the final
examining committee have been made.

Committee Members:

_____

Assist. Prof. Cem Albayrak (Co-Advisor)

_____

Prof. İ Halil Kavaklı (Co-Advisor)

_____

Assist. Prof. Jamin Koo

_____

Prof. Batu Erman

_____

Assist. Prof. Caner Çağlar

Date: _____

ABSTRACT

Towards *In Vitro* Synthesis of Oxygen-tolerant CpI Hydrogenase Variants for Photosynthetic Hydrogen Production

Selin Turan
Master of Science in Molecular Biology and Genetics
September 5, 2022

Hydrogen is an attractive alternative to non-renewable energy sources like fossil fuels due to its carbon-free nature and high energy density. However, most of hydrogen production uses feedstock chemicals derived from fossil fuels, which leads to carbon dioxide emissions and thus contributes to the climate crisis. Photosynthetic hydrogen technology can provide an alternative, sustainable and entirely carbon-free energy cycle. However, oxygen sensitivity of the hydrogenase enzymes that catalyze hydrogen production is one of the most significant barriers against photosynthetic hydrogen production, as oxygen is a by-product of photosynthesis. Therefore, the hydrogenase enzyme to be used in the production of photosynthetic hydrogen must tolerate gaseous oxygen.

In this thesis, four locations known to be important for oxygen tolerance in the *Clostridium pasteurianum* (CpI) hydrogenase were replaced with the amber (TAG) stop codon for non-canonical amino acid (ncAA) incorporation. CpI was chosen as it is one of the most efficient and fast [Fe-Fe] hydrogenases characterized in terms of hydrogen production. Significant progress was made toward ncAA insertion using cell free protein synthesis (CFPS) to the selected spots on the enzyme.

The genes of six *M. jannaschii*-derived aminoacyl-tRNA synthetases (aaRSs) required for cysteine and tyrosine analog incorporation were first cloned into the high-copy pY71 vector. These aaRSs were then produced in *vivo* and purified with IMAC as they have genetically fused N-terminal 6xHis-tags. The orthogonal $tRNA_{CUA}$ template necessary for ncAA insertion was also synthesized with PCR and purified.

In the next part of the thesis, experiments were performed to synthesize *M. barkeri*-derived aminoacyl tRNA synthetase, HRS, necessary for histidine analog incorporation. Different conditions were tested *in vivo* and *in vitro* (CFPS) synthesis of the enzyme. However, HRS formed insoluble aggregates and could not be purified. Lastly, the *M. barkeri*-derived orthogonal tRNA template was produced by PCR assembly and purified.

In the last part, wild-type CpI was anaerobically produced and matured *in vivo*, followed by purification using StrepTactin affinity chromatography. The components required for the two biochemical assays to measure the hydrogen production activity of CpI, namely NADPH- and DTH-driven assays, were synthesized. With the activity assays, the anaerobically produced and purified CpI enzyme was shown to be in the active state. Attempts were made to produce CpI anaerobically in CFPS as well, however this was unsuccessful. But everything necessary for CFPS production of both wild-type and mutant CpI was prepared.

ÖZETÇE

Fotosentetik Hidrojen Üretimi için Oksijene Toleranslı CpI Hidrojenaz Varyantlarının *In Vitro* Üretim Çalışmaları

Selin Turan
Moleküler Biyoloji ve Genetik, Yüksek Lisans
5 Eylül 2022

Hidrojen, karbon içermeyen yapısı ve yüksek enerji yoğunluğu nedeniyle fosil yakıtlar gibi yenilenemeyen enerji kaynaklarına çekici bir alternatiftir. Bununla birlikte, hidrojen üretiminin çoğu, fosil yakıtlardan türetilen hammadde kimyasallarını kullanır, bu da karbondioksit emisyonuna yol açar ve iklim krizine katkıda bulunur. Fotosentetik hidrojen teknolojisi, alternatif, sürdürülebilir ve tamamen karbonsuz bir enerji döngüsü sağlayabilir. Bununla birlikte, oksijen fotosentezin bir yan ürünü olduğundan, hidrojen üretimini katalize eden hidrojenaz enzimlerinin oksijen duyarlılığı, fotosentetik hidrojen üretiminin önündeki en önemli engellerden biridir. Bu nedenle fotosentetik hidrojen üretiminde kullanılacak hidrojenaz enziminin gaz halindeki oksijeni tolere etmesi gerekir.

Bu tezde, *Clostridium pasteurianum* (CpI) hidrojenazında oksijen toleransı için önemli olduğu bilinen dört yer, kanonik olmayan amino asit (ncAA) eklenebilmesi için amber (TAG) durdurma kodonu ile değiştirildi. Hidrojen üretimi açısından karakterize edilen en verimli ve hızlı [Fe-Fe] hidrojenazlardan biri olduğu için CpI seçilmiştir. Enzim üzerinde seçilen noktalara hücresiz protein sentezi (CFPS) kullanılarak ncAA eklenmesi yönünde önemli ilerleme kaydedilmiştir.

Sistein ve tirozin analogunun dahil edilmesi için gerekli olan *M. Jannaschii* türevi altı aminoasil-tRNA sentetazının (aaRS'ler) genleri ilk olarak yüksek kopyalı pY71 vektörüne klonlandı. Bu aaRS'ler daha sonra *in vivo* olarak üretildi ve genetik olarak kaynaşmış N-terminal 6xHis-etiketlerine sahip oldukları için IMAC ile saflaştırıldı. ncAA yerleştirilmesi için gerekli olan ortogonal $tRNA_{CUA}$ şablonu da PCR ile sentezlendi ve saflaştırıldı.

Tezin bir sonraki bölümünde, histidin analoglarının katılması için gerekli olan *M. barkeri* türevi aminoasil tRNA sentetaz, HRS'nin sentezlenmesi için deneyler yapıldı. Enzimin *in vivo* ve *in vitro* (CFPS) sentezi için farklı koşullar test edildi. Ancak HRS, çözünmeyen agregatlar oluşturdu ve saflaştırılamadı. Son olarak, *M. barkeri*'den türetilen ortogonal tRNA şablonu, PCR montajı ile üretildi ve saflaştırıldı.

Son bölümde, doğal-tip CpI, anaerobik olarak *in vivo* olarak üretildi ve olgunlaştırıldı, ardından StrepTactin afinite kromatografisi kullanılarak saflaştırıldı. CpI'ın hidrojen üretim aktivitesini ölçmekte kullanılan iki biyokimyasal tahlil (NADPH ve DTH güdümlü) için gerekli olan bileşenler sentezlendi. Aktivite deneyleri ile anaerobik olarak üretilen ve saflaştırılan CpI enziminin aktif durumda olduğu gösterildi. CFPS'de de anaerobik olarak CpI üretme girişimleri yapıldı, ancak bu başarısız oldu. Ancak hem doğal-tip hem de mutant CpI'nin CFPS üretimi için gerekli olan her şey hazırlandı.

ACKNOWLEDGMENTS

First and foremost, I would like to express my gratitude to my co-advisor, Dr. Cem Albayrak. He had unending patience and understanding, and he always pointed me in the right direction while allowing me to make my own decisions.

I would like to thank Prof. Halil Kavaklı for being my co-advisor and allowing me to use their laboratory during my first year of studies.

I wish to thank my thesis committee members Dr. Jamin Koo, Prof. Batu Erman, and Dr. Caner Çağlar for their time and comments.

I would like to thank TÜBİTAK, for accepting me as one of their scholarship holders in the research projects with the number 118Z871 and title 'Hidrojenaz Enziminin Oksijen Toleransının Moleküler Seviyede İncelenmesi', and with the number 20AG044 and title 'İnflamazom Aracılı Otoenflamatuvar Hastalıkların Takip ve Tedavisine Yönelik Biyogösterge ve İleri Teknoloji Ürünü İlaçların Geliştirilmesi'.

I would also like to thank everyone in my lab group. Ehsan Sarayloo was my first mentor in the lab, and I will always be grateful for his patient, calm demeanor. From coworkers to my closest friends, Yağmur Ersoy and Beste Tunalı, thank you for showing me working is fun when you are with your best friends. I will forever miss the times we worked together. I want to thank Melike Uysal for her company in both the laboratory and the dorms, despite our brief time together.

I wish to thank Dr. Jamin Koo once more for being my mentor and for his support during my studies in his laboratory. I wish to thank him as well as his lab members Kim Seongwon and Cha Yeun for welcoming me to their lab at Hongik University. I also wish to thank Koç University and Hongik University for allowing me to participate in the global exchange program which allowed me to continue my research. I also want to thank all the friends I made during my exchange in South Korea, especially for Kim Subin and Lee Yunpyo; it was so much fun despite the hectic working schedule.

I want to thank my dearest family, Neslihan Turan, Rıdvan Turan, and Murat Turan for being the most encouraging people in my life.

I also want to thank Gizem Keskin and Selin Sevgi, my two best friends since undergrad for their unconditional support and love. I am also grateful to Ayça Öner for always being there for me since high school.

TABLE OF CONTENTS

List of Tables x

List of Figures xiii

Abbreviations xvii

Chapter 1: INTRODUCTION 1

Chapter 2: MATERIALS AND METHODS 16

2.1 Preparations for wild-type and mutant CpI Hydrogenase Production in CFPS 16

2.1.1 Production and Purification of o-tRNAs required for ncAA incorporation 16

2.1.2 Cloning of aminoacyl-tRNA synthetase (aaRS) genes into pY71 vector 19

2.1.3 CFPS to insert ncAAs to sfGFP 25

2.1.4 Site-directed mutation of CpI hydrogenase 26

2.2 Production of M. barkeri derived HRS 27

2.2.1 Cloning and SDM of HRS gene 27

2.2.2 Expression study of HRS with varying temperatures, IPTG concentrations, and additions 28

2.2.3 Production of HRS in CFPS 29

2.2.4 SDM to obtain an OmpT protease resistant HRS 33

2.3 Production, Purification and Analysis of CpI Holo-Hydrogenase 34

2.3.1 Anaerobic production and purification of CpI Hydrogenase 34

Production and purification of CpI Hydrogenase in vivo under the T7 promoter 35

2.3.2 Production and purification of Rice root FNR (RrFNR) 38

2.3.3 ΔiscR cell extract preparation 40

2.3.4 Maturase lysate preparation for in vitro CpI production 41

2.3.5 Cell-free production of wild-type CpI hydrogenase 42

2.3.6 Gas Chromatography to detect hydrogen production activity of CpI Hydrogenase 42

Chapter 3: RESULTS 46

3.1 Preparations for wild-type and mutant CpI Hydrogenase Production in CFPS 46

3.1.1 Production and Purification of o-tRNAs required for CFPS ncAA incorporation 46

3.1.2 Cloning of aminoacyl tRNA synthetase genes into pY71 vector 47

3.1.3 Production and purification of M. jannaschii derived aaRSs 49

3.1.4 CFPS to insert ncAAs to sfGFP 59

3.1.5 Site-directed mutation of CpI hydrogenase 61

3.2 M. barkeri derived aminoacyl tRNA synthetase HRS production 61

3.2.1 Cloning and SDM of HRS gene 61

3.2.2 Expression study of HRS with varying temperatures, IPTG concentrations, and additives 63

3.2.3 Production of HRS in CFPS 66

3.2.4 SDM to obtain an OmpT protease resistant HRS 74

3.2.5 Designing an HRS variant with a solubility tag 75

3.3 Production, Purification and Analysis of CpI Holo-Hydrogenase 76

3.3.1 E. coli BL21 (DE3) ΔiscR pACYC HydEFG competent cell preparation 76

3.3.2 Production and purification of CpI Hydrogenase in vivo under the T7 promoter 77

3.3.3 Production and purification of Rice root FNR 83

3.3.4 Maturase lysate preparation for in vitro CpI production 91

3.3.5 Gas Chromatography to detect hydrogen production activity of CpI Hydrogenase 94

3.3.6 Cell-free production of wild-type CpI hydrogenase 99

Chapter 4: DISCUSSION 104

Chapter 5: CONCLUSION 109

REFERENCES 111

LIST OF TABLES

Table 1.1 The aaRSs used in the thesis and their corresponding noncanonical amino acids. 2

Table 2.1 Primers used to amplify *M. jannaschii*-derived o-tRNA template from pY71 minHH-otDNAOpt plasmid and the final PCR product sequence. 16

Table 2.2 Large-scale PCR protocol to amplify *M. jannaschii* o-tRNA template. . 16

Table 2.3 Primers used for PCR assembly of *M. barkeri*-derived o-tRNAs. 17

Table 2.4 Amplification PCR primers and final products for *M. barkeri*-derived optimized o-tRNA..... 18

Table 2.5 Primers used for incorporation of NdeI and SalI cut sites along with C-terminal 6xHis-tag. 19

Table 2.6 PCR procedure to insert restriction sites and the 6xHis affinity tag. 20

Table 2.7 Restriction reaction reagents to cut the aaRS genes and pY71 vector. ... 20

Table 2.8 Ligation reaction setup. 21

Table 2.9 Primers used to remove the extra methionine via SDM. 21

Table 2.10 SDM procedure to remove the extra methionine in N-terminal of aaRS genes. 21

Table 2.11 Buffers used for purification of 6xHis-tagged proteins. 23

Table 2.12 FPLC procedure to purify 6xHis-tagged pPaFRS. 23

Table 2.13 Buffers used in the dialysis study for pAMFRS[45]. 24

Table 2.14 List of reagents in a CFPS reaction. 25

Table 2.15 Primers designed to incorporate TAG stop codon to CpI hydrogenase gene. 26

Table 2.16 SDM procedure to incorporate TAG stop codon to CpI hydrogenase gene. 26

Table 2.17 Primers used to remove the extra methionine via SDM. 27

Table 2.18 Required solutions for cell extract preparation..... 30

Table 2.19 Tested HRS production conditions in CFPS 31

Table 2.20 Buffers required in Western-Blot analysis of 6xHis-tagged proteins.... 33

Table 2.21 Primers used to change lysine at position 105 to alanine. 34

Table 2.22 Buffers used with StrepTactin Sepharose resin for CpI hydrogenase purification. 36

Table 2.23 Purification protocol of CpI Hydrogenase by StrepTactin affinity chromatography. 36
Table 2.24 Buffers used with Strep-Tactin®XT® 4Flow kit for CpI hydrogenase purification. 37
Table 2.25 Purification procedure for CpI hydrogenase with StrepTactin affinity chromatography. 37
Table 2.26 Buffers used with CM Sepharose resins for RrFNR purification. 38
Table 2.27 Purification procedure of RrFNR by cation exchange chromatography. 39
Table 2.28 Second-round purification procedure of RrFNR by cation exchange chromatography. 40
Table 2.29 H_2 dilutions to draw a calibration curve with gas chromatography. 43
Table 2.30 O_2 dilutions to draw a calibration curve with gas chromatography. 43
Table 2.31 NADPH activity assay reagents and required concentrations to measure hydrogen productivity of CpI Hydrogenase. 44
Table 2.32 DTH activity assay reagents and required concentrations to measure hydrogen productivity of CpI Hydrogenase. 45
Table 3.1 The promoter strength of $tRNA^{Pyl\text{-}opt}$. 46
Table 3.2 Protein concentration determination of aaRS elution fractions with DC assay. 59
Table 3.3 OmpT protease target sites and the resulting protein fraction sizes. 74
Table 3.4 Concentration determination of CpI Hydrogenase with Bradford assay. 80
Table 3.5 Concentration determination for EF2 of CpI Hydrogenase with Bradford assay. 82
Table 3.6 Concentration determination for EF3 of CpI Hydrogenase with Bradford assay. 82
Table 3.7 Calculated concentrations and molarities for pure EF2 and EF3 of CpI Hydrogenase. 82
Table 3.8 Concentration determination for RrFNR Lot1 with Bradford assay. 85
Table 3.9 Concentration determination for RrFNR Lot2 with Bradford assay. 86
Table 3.10 Calculated concentrations and molarities for RrFNR Lot 1 and Lot 2.. 86
Table 3.11 Concentration determination of RrFNR dilutions with Bradford assay.90
Table 3.12 Calculated concentration and molarity of RrFNR Lot3. 91

Table 3.13 H_2 mole calculations using ideal gas law and the measured areas of H_2 dilutions. 96
Table 3.14 O_2 mole calculations using ideal gas law and the measured areas of O_2 dilutions. 96
Table 3.15 The ratios of the lysates required to prepare the final maturase lysate. 102

LIST OF FIGURES

Figure 1.1 The designated noncanonical amino acids are planned to be inserted into CpI hydrogenase. 1

Figure 1.2 Activity levels of CpI enzyme[1]. 2

Figure 1.3 The suggested photosynthetic hydrogen production mechanism using CpI [Fe-Fe] hydrogenase[1]. 3

Figure 1.4 Structure of *Clostridium pasteurianum* (Cp) [Fe-Fe] hydrogenase. PDB: 3C8Y. 5

Figure 1.5 The hypothesized maturation process that leads to the active CpI holoenzyme. 6

Figure 1.6 The assay reactions to assess H_2 production rates of CpI in the presence of oxygen[8]. 8

Figure 1.7 The effect of SmbP on the solubility of MbSacRS[20]. 11

Figure 1.8 Comparison of $tRNA^{Pyl}$ and $tRNA^{Pyl\text{-}opt}$ [31]. 12

Figure 1.9 Schematic CFPS workflow from lysate preparation to *in vitro* protein production[33]. 13

Figure 1.10 Change in hydrogenase activity and ferredoxin-dependent hydrogen production *in vivo* with *isc* overexpression and *iscR* deletion[40]. 14

Figure 2.1 Planned PCR assembly design to create $tRNA^{Pyl\text{-}opt}$. 17

Figure 2.2 PCR assembly design to create $tRNA^{Pyl\text{-}opt\text{-}ribo}$. 17

Figure 3.1 Agarose gel analysis of o-tRNA templates produced by PCR, required for CFPS ncAA incorporation. 47

Figure 3.2 Agarose gel analysis of check digests from the cloned aaRS plasmids. 48

Figure 3.3 Sequence confirmation of the resulting plasmids after SDMs to remove second methionine at translation start. 49

Figure 3.4 Growth curve of *E. coli* BL21(DE3) pLysS pY71 pAFRS. 50

Figure 3.5 Purification of pAFRS. 51

Figure 3.6 Growth curve of *E. coli* BL21(DE3) pLysS pY71 CNFRS. 51

Figure 3.7 Purification of CNFRS. 52

Figure 3.8 Growth curve of *E. coli* BL21(DE3) pLysS pY71 OMeYRS. 53

Figure 3.9 Purification of OMeYRS. 53

Figure 3.10 Growth curve of *E. coli* BL21(DE3) pLysS pY71 pIFRS. 54

Figure 3.11 Purification of pIFRS. 55

Figure 3.12 Growth curve of *E. coli* BL21(DE3) pLysS pY71 pPaFRS..... 55
Figure 3.13 Purification of pPaFRS. 56
Figure 3.14 Growth curve of *E. coli* BL21(DE3) pLysS pY71 pAMFRS. 57
Figure 3.15 Purification of pAMFRS. 57
Figure 3.16 SDS-PAGE analysis of purified aaRS samples after dialysis. 58
Figure 3.17 Concentrations of soluble sfGFP variants produced in CFPS. 60
Figure 3.18 Concentrations of soluble sfGFP39TAG and sfGFP23TAG variants produced in CFPS. 60
Figure 3.19 Sequence alignments to confirm amber stop codon (TAG) insertions by SDM at L192, G194, T356, and S357 61
Figure 3.20 Agarose gel analysis of the check digest for the cloned HRS plasmid (1278 bp). 62
Figure 3.21 Sequence alignment to confirm the removal of the second methionine at translation initiation in in HRS by SDM. 62
Figure 3.22 Growth curve of *E. coli* BL21(DE3) pLysS pY71 HRS. 63
Figure 3.23 Purification of HRS. 64
Figure 3.24 SDS-PAGE analysis of HRS expression at varying temperatures and IPTG concentrations (HRS: 48.3 kDa). 65
Figure 3.25 SDS-PAGE analysis of HRS expression with ethanol addition at varying temperatures (HRS: 48.3 kDa)..... 66
Figure 3.26 Western blot analysis of cell-free HRS production with additions of Skp, GroEL-GroES chaperones, as well as oxidized (GSSG) and reduced (GSH) glutathione in varying concentrations (HRS: 48.3 kDa)..... 67
Figure 3.27 Western blot analysis of cell-free HRS production with additions of Skp chaperone, protease inhibitor and glutathione (HRS: 48.3 kDa). 68
Figure 3.28 Growth curve of (a) *E. coli* BL21 in LB medium. 69
Figure 3.29 Cell-free produced soluble sfGFP yields with different cell extracts. . 70
Figure 3.30 Western blot and SDS-PAGE analysis comparison for pPaFRS and HRS. 71
Figure 3.31 Densitometry analysis of the Western-blot image for pPaFRS dilutions. 72
Figure 3.32 BL21 cell extract magnesium concentration optimization for CFPS... 72
Figure 3.33 Comparison of opt. and BL21 cell extracts in terms of cell-free soluble sfGFP production. 73

Figure 3.34 Purification of HRS produced in preparative (3mL-scale) CFPS. 74
Figure 3.35 The designed HRS fusion protein with SmbP solubility tag. 76
Figure 3.36 Growth curve of *E. coli* BL21 (DE3) ΔiscR pACYC HydEFG for competent cell preparation. 77
Figure 3.37 (a) Growth curve of *E. coli* BL21 (DE3) ΔiscR pACYC HydEFG pET21b sCpI sIIC StrepII for CpI hydrogenase production. 78
Figure 3.38 Anaerobic purification of CpI hydrogenase by StrepTactin affinity chromatography. 78
Figure 3.39 SDS-PAGE analysis of CpI Hydrogenase purification by StrepTactin affinity chromatography. 79
Figure 3.40 Purification of CpI hydrogenase by StrepTactin affinity chromatography using IBA Strep-Tactin®XT® gravity flow column (CV = 0.2 mL). 80
Figure 3.41 SDS-PAGE analysis of CpI Hydrogenase purification by StrepTactin affinity chromatography using StrepTactin®XT® gravity flow column (IBA). 81
Figure 3.42 Growth curve of small-scale *E. coli* BL21 (DE3) pET21b RrFNR. 83
Figure 3.43 Purification of RrFNR by cation exchange chromatography. 84
Figure 3.44 SDS-PAGE analysis of RrFNR purification by cation exchange chromatography. 85
Figure 3.45 Growth curve of large-scale *E. coli* BL21 (DE3) pET21b RrFNR in TB medium. 87
Figure 3.46 Purification of RrFNR using cation chromatography. 88
Figure 3.47 SDS-PAGE analysis of RrFNR chromatography fractions. 89
Figure 3.48 Chromatography fractions of the large-scale RrFNR second purification. 90
Figure 3.49 Growth curve of small-scale *E. coli* BL21 (DE3) ΔiscR pACYC HydEFG. 91
Figure 3.50 SDS-PAGE analysis of maturase lysate preparation. 92
Figure 3.51 Growth curve of large-scale *E. coli* BL21 (DE3) ΔiscR pACYC HydEFG. 93
Figure 3.52 SDS-PAGE analysis of large-scale maturase lysate preparation. 93
Figure 3.53 SDS-PAGE analysis of different maturase lysate preparations. 94
Figure 3.54 Gas chromatography (GC) calibration curves. 97
Figure 3.55 Gas chromatography results from DTH and NADPH assays. 98
Figure 3.56 Gas chromatography NAPDH oxygen addition assay results. 99

Figure 3.57 SDS-PAGE analysis of CFPS sfGFP and CpI Hydrogenase production, comparing ΔiscR and maturase lysates. 100
Figure 3.58 Comparison of ΔiscR and maturase lysates in terms of soluble sfGFP production efficiency. 101
Figure 3.59 CFPS troubleshooting comparing old and new pY71 sfGFP plasmids and PEP stocks in terms of soluble sfGFP production efficiency. 102

ABBREVIATIONS

2YT	2x Yeast extract– tryptone broth
aaRS	Aminoacyl-tRNA synthetase
CFPS	Cell-free protein synthesis
CM	Carboxymethyl
CNF	Para-cyano phenylalanine
Cp	*Clostridium pasteurianum*
CV	Column volume
dH_2O	Distilled water
dNTP	Deoxynucleotide
DTH	Dithionite
E. coli	*Escherichia coli*
EDTA	Ethylenediaminetetraacetic acid
FPLC	Fast protein liquid chromatography
IMAC	Immobilized metal ion affinity chromatography
M. barkeri	*Methaosarcina barkeri*
M. jannaschii	*Methanocaldococcus jannaschii*
N. europaea	*Nitrosomonas europaea*
NADPH	Nicotinamide adenine dinucleotide phosphate
NaOAc	Sodium acetate
OMeY	Ortho-methyl tyrosine
pAF	Para-amino phenylalanine
pAMF	Para-azidomethyl-phenylalanine
PCR	Polymerase chain reaction
pIF	Para-iodo phenylalanine
pPaF	Para-propargyloxy-phenylalanine
RrFNR	Rice root ferredoxin $NADP^+$ reductase
RT	Room temperature
SDM	Site-directed mutagenesis
SDS-PAGE	Sodium dodecyl sulphate–polyacrylamide gel electrophoresis
sfGFP	Superfolder green fluorescent protein
SmbP	Small metal binding protein
SmbP	Small metal-binding protein

SynFd	*Synechocystis* sp. Ferredoxin
TCD	Thermal conductivity detector

Chapter 1:
INTRODUCTION

Thesis Overview

With technology covering all over daily life, there is an overwhelming need for an environmentally friendly, sustainable energy source to prevent a climate crisis. Hydrogen has the potential to be the main alternative carbon-free energy source to the current energy sources such as fossil fuels. However, the methods to produce H_2 are currently overpriced, require too much energy, and increase the carbon footprint. Unless it can be generated in a carbon-free manner efficiently, its use will be limited.

The aim of the thesis is to build towards the ultimate goal of developing hydrogenase variants with high oxygen resistance by inserting the non-canonical amino acids (ncAAs) given in Figure 1.1. Table 1.1 shows the aaRS enzymes that are produced in the project and the noncanonical amino acids they incorporate.

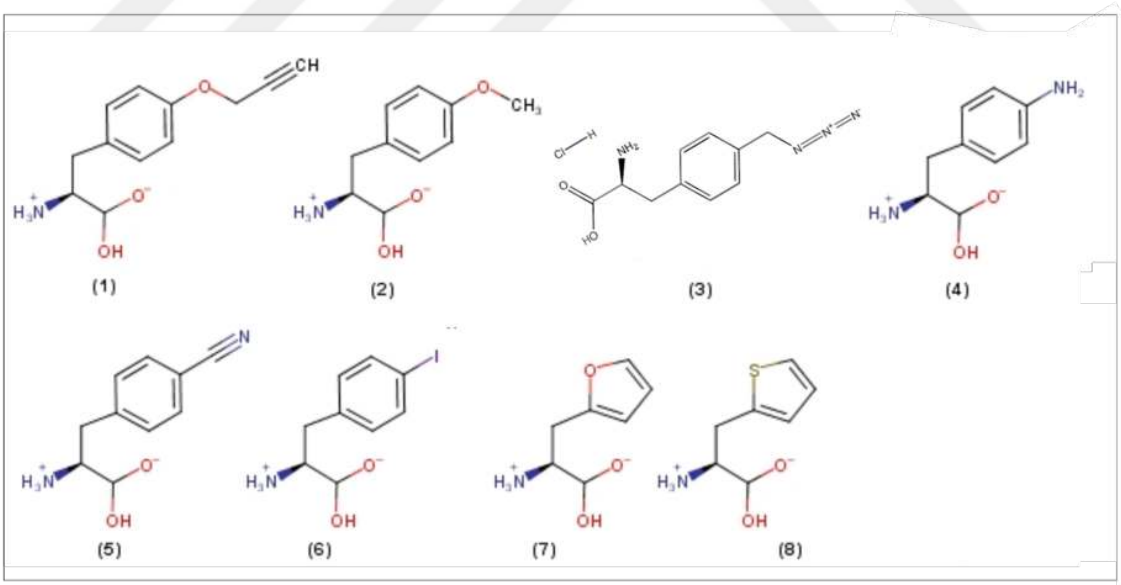


Figure 1.1 The designated noncanonical amino acids are planned to be inserted into CpI hydrogenase. (1) p-propargyloxy-l-phenylalanine. (2) o-methyl tyrosine. (3) p-azidomethyl-l-phenylalanine. (4) p-amino phenylalanine. (5) p-cyano phenylalanine. (6) p-iodo phenylalanine. (7) 2-furyl-alanine. (8) 2-thyenyl-alanine. Amino acids 1-3 are tyrosine analogs, 4-6 are phenyl analogs, and 7-8 are histidine analogs.

Table 1.1 The aaRSs used in the thesis and their corresponding noncanonical amino acids.

aaRS	Derived Organism	Noncanonical amino acid to insert
pPaFRS	*Methanocaldococcus jannaschii*	p-propargyloxy-l-phenylalanine
pAMFRS	*Methanocaldococcus jannaschii*	p-azidomethyl-l-phenylalanine
OMeYRS	*Methanocaldococcus jannaschii*	o-methyl-tyrosine
pIFRS	*Methanocaldococcus jannaschii*	p-iodo-l-phenylalanine
CNFRS	*Methanocaldococcus jannaschii*	p-cyano-l-phenylalanine
pAFRS	*Methanocaldococcus jannaschii*	p-amino-l-phenylalanine
HRS	*Methanosarcinae barkeri*	2-furyl-alanine, 2-thyenyl-alanine

In a previous study, to investigate the oxygen inactivation mechanism of CpI enzyme, locations close to the proximal catalytic center (PCC) and proximal delivery center (PDC) were analyzed. 4 hotspots, L192, G194, T356, and S357, on the CpI enzyme were observed to be the most important locations for oxygen tolerance of the enzyme (Figure 1.2)[1]. While L192 and T356 locations are determined to be related to the hydrogen production activity of the enzyme, G194 and S357 locations are related to the oxygen tolerance of CpI[1]. With Figure 1.2 being the starting point of this thesis, the 4 determined hotspots were chosen as the insertion sites of ncAAs.

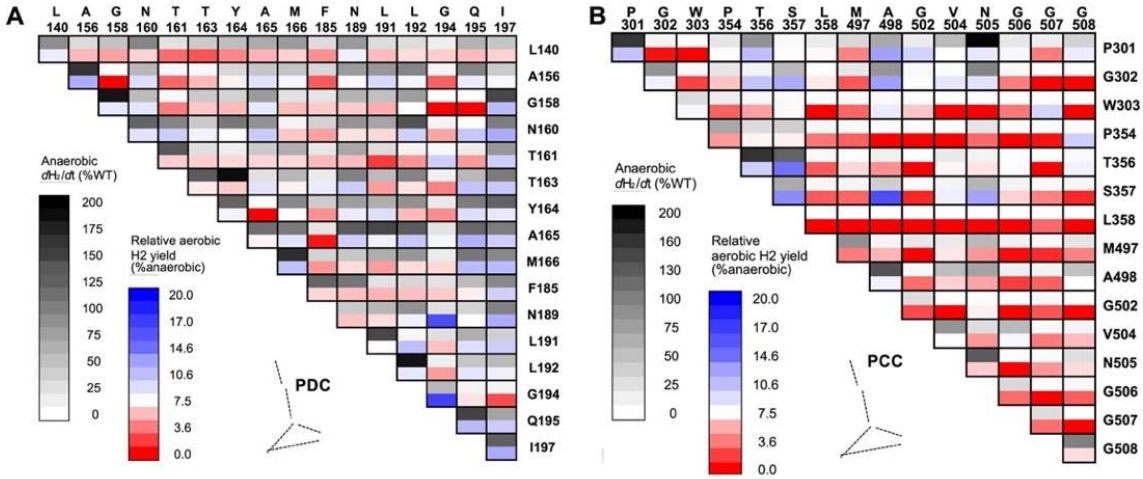


Figure 1.2 Activity levels of CpI enzyme[1]. (a) Anaerobic hydrogen production yield of the variants. (b) Relative aerobic hydrogen yields of the variants in the presence of 5% oxygen, compared to the anaerobic wild-type yield. Gray scale: anaerobic hydrogen production yield. Red-blue scale: Relative aerobic hydrogen yields.

After establishing oxygen-tolerant variants, these enzymes are planned to be inserted into a photosynthetic organism. Figure 1.3 shows the suggested photosynthetic hydrogen production mechanism. In the photosynthetic organism, photons from the sun activate the photosystems (PSII and PSI) and power hydrolysis of water. Electrons released by hydrolysis are then transferred from PSI to ferredoxin, and from there to iron-iron hydrogenase. The hydrogenase enzyme combines these electrons with protons to produce hydrogen gas. The critical point here is to ensure that the hydrogenase in the system is minimally affected by oxygen since oxygen is released as a byproduct of hydrolysis. Some of the released electrons also need to be transferred to NADPH via ferredoxin-NADPH reductase (FNR), as reduced NADPH is necessary for the cell to survive.

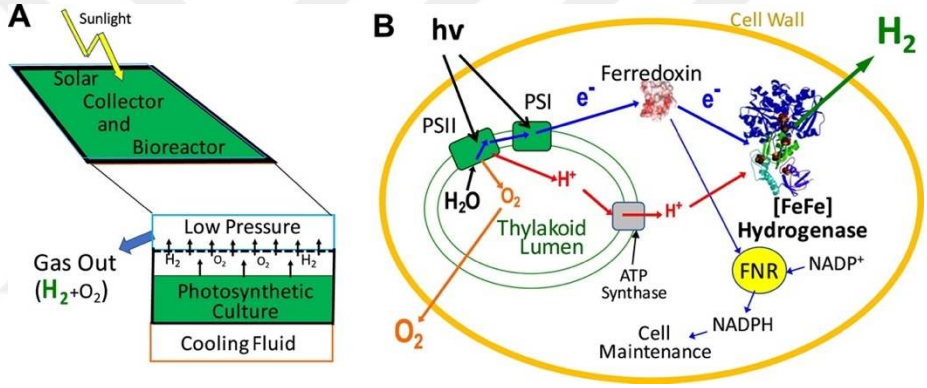


Figure 1.3 The suggested photosynthetic hydrogen production mechanism using CpI [Fe-Fe] hydrogenase[1].

Hydrogen Production Technologies

As hydrogen is plenty in nature, it can be obtained from sources like hydrocarbon, water, or biomass. There are three main categories of the extraction of hydrogen, being electrochemical, thermochemical, and biological methods. Electrochemical techniques include the application of an electric current followed by a chemical reaction. While electrolysis, breaking down water to hydrogen and oxygen, is one of the cleanest methods, it is highly costly. The photoelectrochemical method utilizes photons as the energy source. Thermochemical methods can be divided into oxidative and non-oxidative processes. An external or internal thermal energy input extracts hydrogen from the feedstock chemical(s) in this method. High temperature or another energy source such as

plasma or radiation is required. Hydrogen can also be produced with biological methods via photosynthetic and anaerobic organisms. These methods are promising in terms of clean and sustainable hydrogen production methods. Of the mentioned techniques, only oxidative thermochemical techniques are currently suitable for industrial processes. These include steam methane reformation (SMR), partial oxidation (PO_x), autothermal reforming (ATR), and steam-iron process (SIP)[2].

Biological Hydrogen Production

The traditional fermentative production of hydrogen from biomass gives low yields, making the method unfavorable. An alternative method being developed is the production of hydrogen from engineered cell extracts.

Hydrogen can also be produced in the photosynthetic pathway. In a photosynthetic organism, photosystems (PSI and PSII) are activated by photons and lead to hydrolysis of water. Electrons released by hydrolysis are then transferred from PSI to ferredoxin, and from there to the hydrogenase that would be inserted to the recombinant organism. The hydrogenase enzyme then produces hydrogen by combining electrons and protons.

Hydrogenases

Hydrogenases are bidirectional metalloproteins that both oxidize/consume or produce H_2. There are three main families of hydrogenases, based on the characteristics of their binuclear catalytic center, being either [Fe-Fe], [Ni-Fe], or [Fe] only. While [Ni-Fe] hydrogenases are more focused on hydrogen oxidation, [Fe-Fe] hydrogenases mostly focus on the production of hydrogen by lowering the number of protons with the help of electrons generated during fermentation[3][4]. Compared to the [Fe-Fe] hydrogenases, [Ni-Fe] hydrogenases show more tolerance towards oxygen as well as carbon monoxide[5]. However, their activity levels are 10 to 100 times less, and they exhibit higher affinity toward hydrogen[3].

Clostridium pasteurianum, a Gram-positive anaerobic N_2-fixing bacterium, converts hexose sugars into butyrate, carbon dioxide, and acetate during fermentation, and reduced ferredoxin, an electron carrier, accumulates in the cell. To sustain the fermentation cycle, the reduced ferredoxin is oxidized by a [Fe-Fe] hydrogenase enzyme,

which combines electrons and protons to form H_2[6]. This enzyme is the first hydrogenase identified in *C.pasteurianum*, therefore called CpI. It is one of the fastest known hydrogenases in terms of H_2 production. The other one is called CpII, and during nitrogen fixation, it is responsible for the oxidation of H_2, which is a byproduct of the reduction reaction of nitrogen to ammonia by the endogenous nitrogenase enzyme. Even though both hydrogenases are bidirectional, CpI displays an H_2 production rate that is 550 times faster compared to CpII meanwhile, the H_2 consumption rate by CpI is 30% slower[7].

CpI is a 63.8 kDa protein[8], and has 4 structural domains, with the largest one containing the active site of the enzyme, called the H-cluster (HC, Figure 1.4). The rest of the domains contain the [Fe-S] clusters, which are named FS4A, FS4B, FS4C, and FS2[9]. Inside the H-cluster, there is a $[4Fe\text{-}4S]_H$ cluster connected to a $[2Fe]_H$ cluster by a cysteinyl sulfur ligand. The $[2Fe]_H$ cluster organizes three CO and two CN^- ligands as well as a dithiol bridge[10].

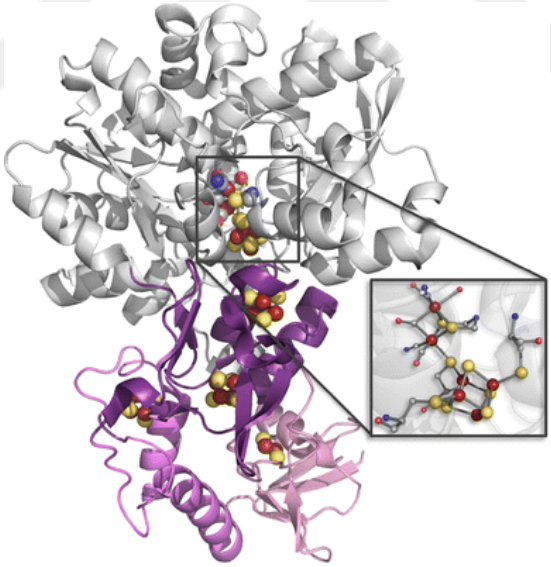


Figure 1.4 Structure of *Clostridium pasteurianum* (Cp) [Fe-Fe] hydrogenase. PDB: 3C8Y. Domain colors: Gray: largest domain of the enzyme, contains the H-cluster (active site). Dark purple: domain with 2 [4Fe-4S] clusters (FS4A and FS4B). Light purple: domain with one [4Fe-4S] cluster (FS4C). Pink: domain with a [2Fe-2S] cluster. Atom colors: Red: oxygen. Dark red: iron. Yellow: sulfur. Blue: nitrogen. Gray: carbon. Reprinted by permission from Joan B. Broderick et al, 2014, Springer Nature [JBIC Journal of Biological Inorganic Chemistry], License No: 5373670044687.

There are three maturation enzymes that take part in the formation of active CpI holoenzyme. Figure 1.4 shows the maturation pathway of a [Fe-Fe] hydrogenase, where

the $[4Fe\text{-}4S]_H$ is first formed, followed by the transfer of $[2Fe]_H$ cluster from HydF to hydrogenase, resulting in the activated enzyme. A radical S-adenosyl methionine (SAM) enzyme, HydG, first converts free tyrosine into CO and CN^- ligands and possibly to the dithiol bridge (DTMX). These ligands are combined with a HydG-bound [4Fe-4S] cluster to form HydG-co, a [2Fe] subcluster precursor. This precursor is then transferred to HydF GTPase, where GTP is hydrolyzed for a possible conformational change. The precursor is then assembled with the hydrogenase apoenzyme, forming the active protein. It is believed that HydE, also a radical SAM enzyme, helps transport the [2Fe] precursor[10].

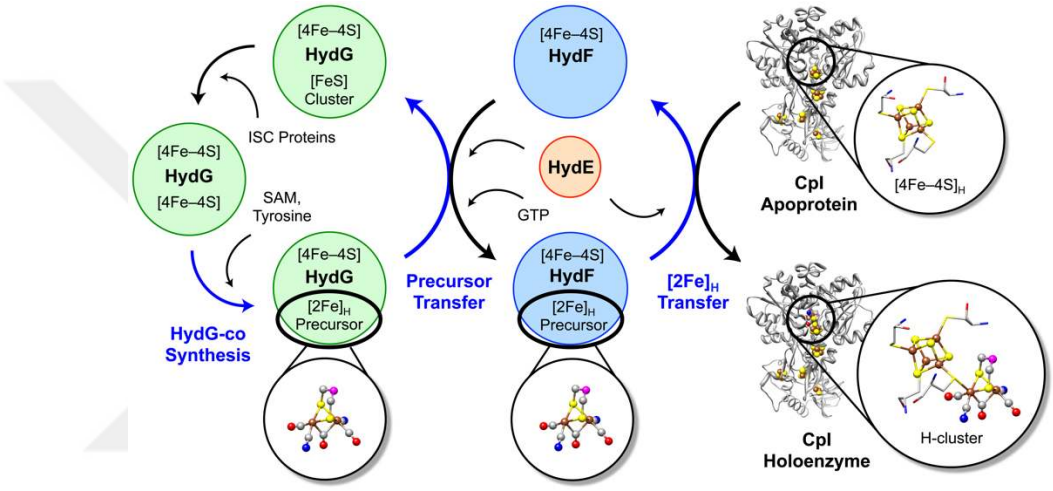


Figure 1.5 The hypothesized maturation process that leads to the active CpI holoenzyme. Blue arrow: formation and direction of $[2Fe]_H$ cluster. Atom colors: Brown: iron. Yellow: sulfur. Oxygen: red. Blue: nitrogen. Gray: carbon. Magenta: unknown atom (X)[10].

While CpI gets inactivated in aerobic conditions in a couple of minutes, the exact mechanism of how oxygen irreversibly inactivates the enzyme is still not clear. One hypothesis is that O_2 binds to one of the iron atoms in the $[Fe\text{-}Fe]_H$ subcluster, causing the formation of reactive oxygen species (ROS). As a result, the [4Fe-4S] cluster inside the active site (H-cluster) disintegrates[11]. The proposed mechanism of inactivation consists of two rate-limiting steps, the first is the diffusion of oxygen into the H-cluster and the second is the irreversible activity loss[12]. In another study, it was hypothesized that the $[Fe\text{-}Fe]_H$ subcluster detaches from the H-cluster due to inactivation caused by oxygen[13].

[Ni-Fe] hydrogenases are more oxygen tolerant than the [Fe-Fe] ones, so the underlying mechanisms of this tolerance were studied in a previous study. Two additional cysteines occupy the [Fe-S] cluster next to the catalytic center of a [Ni-Fe] hydrogenase

from *R. eutropha*. To make the enzyme tolerant to oxygen, these cysteines act as an extra electron supply in the presence of oxygen, reducing oxygen to water and preventing damage to the active site[14]. Therefore, to mimic this [Ni-Fe] hydrogenase for its oxygen tolerance, locations close to the proximal catalytic center (PCC) and proximal delivery center (PDC) of CpI were analyzed in another study[1]. 4 hotspots on the CpI enzyme are determined to be the most crucial points for oxygen tolerance of the enzyme. L192 and T356 locations, when mutated into cysteine and produced with CFPS, were shown to increase the hydrogen production activity of the enzyme. G194 and S357 (Figure 1.1), when mutated into cysteine, increased the oxygen tolerance of CpI compared to the wild-type enzyme[1].

Assessment of hydrogen production activity of hydrogenases

There are several methods to assess the hydrogen production efficiency and oxygen sensitivity of hydrogenases. One of these methods is protein film voltammetry (PFV). In this technique, the protein of interest is attached to the surface of an electrode, and the activity measurement is done by analyzing the currents generated by the enzyme when a positive or negative potential is applied. The issue with this technique is that the system is being charged with unnatural electrons and there is a possibility that different hydrogenases will attach to the film with varying angles, leading to different electron transfer distances. Therefore, it may not represent an accurate and global prediction of O_2 sensitivity[8][15].

A different method was suggested, which aimed to supply electrons to the enzyme in a natural manner[8]. Two parameters that are important in oxygen sensitivity analysis can be measured with the established assays, the first one being the rate constant of inactivation and the second one being the residual hydrogenase activity following oxygen exposure. One of the two assay pathways (Figure 1.5a), inspired by the photosynthetic pathway, utilizes nicotinamide adenine dinucleotide (NADPH), ferredoxin (Fd), and ferredoxin-$NADP^+$ reductase (FNR) to transfer electrons to CpI. As the ultimate aim is to use oxygen-tolerant CpI variants for photosynthetic hydrogen production, Fd from *Synechocystis sp.,* a photosynthetic cyanobacterium was chosen. FNR from rice root

(*Oryza sativa*) was picked as it was previously shown to give faster results in terms of hydrogen productivity in the NADPH-driven pathway.

The next assay uses dithionite (DTH) as the electron source, and the electrons are transferred to CpI by Fd (Figure 1.5b). As DTH has high reducing power, it can consume the remaining oxygen in the assay vial after the oxygen exposure of CpI. This allows the analysis of the residual activity of CpI after exposure[8].

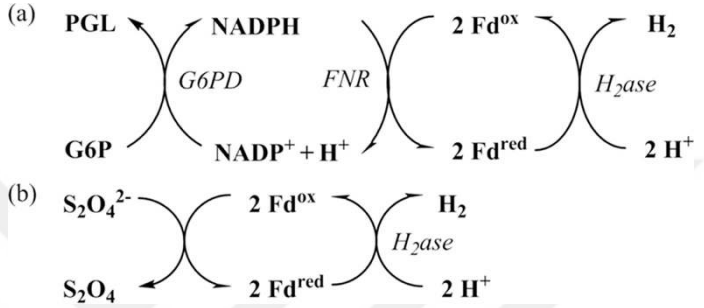


Figure 1.6 The assay reactions to assess H_2 production rates of CpI in the presence of oxygen[8]. (a) NADPH-driven assay. (b) DTH-driven assay.

Non-canonical amino acids & genetic code expansion

Proteins have a broad spectrum of functions, from acting as messengers to initiating biochemical reactions and taking part in transcription. Despite the wide functional range of proteins, only 20 amino acids exist in nature as their building blocks. Therefore, it is inevitable that proteins require additions and modifications to increase functional diversity. These changes include the addition of cofactors, metal ions, or post-translational modifications such as acetylation and methylation[16]. Some archaea and eubacteria encode non-canonical amino acids (UAAs/ncAAs) like selenocysteine and pyrrolysine[17], regardless of the conserved spectrum of amino acids. Therefore, it is not unreasonable to expect that by carefully expanding the genetic code, new functional skills caused by personal, chemical, or biological properties may be picked up by proteins[16]. However, the challenge is to incorporate the ncAAs into a polypeptide chain with the efficiency and fidelity as those of natural amino acids.

Insertion of non-canonical amino acids (ncAAs), whether global or site-specific, has a vast application field. It gives control over the physical and chemical characteristics

of proteins, granting an evolutionary advantage. ncAAs can be used in therapeutical areas like incorporating immunogenic amino acids to break immunogenic tolerance and generate vaccines against self-proteins to battle cancer or inflammation. Moreover, proteins with an attached toxin, radioisotope, or another protein can be created[16]. ncAA insertion can also be used for protein labeling (i.e. BONCAT/FUNCAT) to visualize, track and tag proteins[18].

For expanding the genetic code using ncAAs, three components are necessary. The first one is a stable ncAA that can pass through the cell membrane or be synthesized inside the cell. The key here is that the ncAA should not be a substrate of endogenous aminoacyl-tRNA synthetases. Secondly, an operative orthogonal tRNA and aminoacyl-tRNA synthetase pair is required. The pair should bind the amino acid in question but not react with endogenous tRNA's and aminoacyl-tRNA synthetases. Lastly, a codon specific to the orthogonal tRNA's anticodon but not recognized by endogenous tRNAs is necessary. The unique codon can be easily recognized by choosing a blank or nonsense codon and having a corresponding tRNA. However, aminoacylation accuracy and orthogonality in aminoacyl-tRNA synthetase/tRNA pair are more challenging. As the tRNA recognition may be species or domain-specific, the challenges can be overcome by importing a heterologous aminoacyl-tRNA synthetase and tRNA pair from a different domain of life[16]. Mutagenesis and selection can be further performed if there is a need for higher specificity in orthogonality. Most popular orthogonal pairs are $tRNA_{CUA}$/tyrosyl-tRNA synthetase derived from *Methanococcus jannaschii* and $tRNA_{CUA}$/pyrrolysyl-tRNA synthetase derived from *Methanosarcinae* species.

The incorporation of ncAAs can be global or site-specific. In global replacement, the target natural amino acid is replaced by its substitute ncAA in the entire proteome. The concentrations of ncAAs and the selectivity of synthetases play an important role in reactions since two amino acids compete for one position. As changes in the physical and chemical composition of the resulting protein are often observed, the advantage of global reassignment is in materials design and proteomic analysis[19].

If one needs to introduce point mutations to proteins while preserving the structural integrity, site-specific incorporation is the optimal method. Stop codons (mostly amber codon) are used in ncAA insertion[19]. However, the target protein sequence must

be known beforehand. Moreover, the incorporation of more than one ncAA is challenging with this method, and therefore it is unlikely to make any global changes to the protein[18].

Methanosarcina barkeri-derived orthogonal tRNA-aaRS systems

M. barkeri (Mb)-derived aaRS that naturally encodes pyrrolysine, PylRS, is a popular choice to insert ncAAs due to its orthogonality with both bacterial and eukaryotic organisms. PylRS-$tRNA^{Pyl}$ pairs are engineered so that they can incorporate over 100 ncAAs, including pyrrolysine, phenyl, histidine, or tryptophan analogs[20]. While the enzyme has no problem with solubility in vivo, previous studies reported that it has low solubility *in vitro*[21]. The reason behind this is reported to be the hydrophobic N-terminal domain that plays a crucial role in the recognition, binding, and charging of $tRNA^{Pyl}$ *in vivo*[22]. Mutations to the N-terminus or exchanging the N-terminal domain of *Mb*PylRS with *Mm*PylRS N-terminal domain, which is known to be more soluble, led to improvements in protein solubility[20].

To overcome the solubility issue of a protein, an easy approach would be to decrease the IPTG concentration or temperature of the protein. The addition of up to 3% (v/v) ethanol[23] and glycerol[24] has also been shown to help form the correct shape of the proteins by slowing protein production. Chaperones and chaperonins also help the proteins to fold into their natural 3D shapes. For example, Skp is a 17 kDa periplasmic chaperone; the hydrophobic cage formed by trimeric Skp traps and prevents OMPs (outer membrane protein) from aggregation and misfolding while they pass from the inner and outer membranes in Gram-negative bacteria[25]. Chaperonins are large double-ringed barrel structures that assist denatured proteins fold by ATP hydrolysis. GroEL, a homo-oligomer, is the most well-known chaperonin in *E.coli.* When its cochaperonin GroES binds to GroEL, GroEL undergoes a conformational change, increasing the volume of the central cavity where the substrate is bound. This creates a polar environment, which aids in the formation of the substrate's native state[26].

Changing the redox state of the environment can also prevent misfolding of proteins with disulfide bonds. Glutathione, a tripeptide that is derived from glutamate, cysteine,

and glycine, along with its oxidized state (GSSG) aids to change the redox potential of its surrounding area for optimal disulfide bond formation[27].

Another strategy to increase protein solubility is by genetically fusing a solubility tag to either N- or C- terminus of the target gene[20]. SmbP (small metal-binding protein) is a 9.9 kDa protein that is found in the periplasm of *Nitrosomonas europaea.* The monomer protein consists of only 93 amino acids and has 10 back-to-back repeats of 7 amino acid motif. The motif has a conserved histidine residue in the 4th location[28]. When SmbP is fused to the N-terminal of MbSacRS (*M. barkeri* derived aaRS that encodes Sac ncAA), the protein could be produced in the soluble form while the untagged version was mostly insoluble (Figure 1.7)[20]. When the aaRS is co-expressed with its target tagged protein under the lpp promoter, an overproduction of SmbP-MbSacRS is observed on the SDS-PAGE gel at the expected size (57 kDa). When both tagged, and untagged proteins were expressed with the T7 promoter, the soluble SmbP-MbSacRS can clearly be seen on the SDS-PAGE gel. However, there is also an overwhelming amount of insoluble protein. The untagged protein only shows up as an insoluble aggregate.

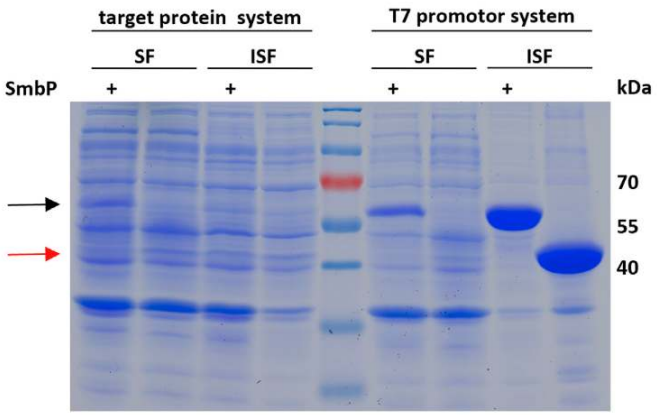


Figure 1.7 The effect of SmbP on the solubility of MbSacRS[20]. Left: aaRS co-expression driven by lpp promoter, used in ncAA incorporation. Right: aaRS overproduction under T7 promoter. Black arrow: SmbP-MbSacRS (57.5 kDa) Red arrow: MbSacRS (47.6 kDa) SF: Soluble fraction. ISF: Insoluble fraction.

PylRS-derived HRS

PylHRS, an aaRS enzyme that incorporates histidine analogs, was obtained by mutating 5 locations on the *M. barkeri* PylRS enzyme. L270, Y271, L274, and C313 locations on the enzyme are related to the recognition of pyrrolysine and thus were mutated to the best fit result of random mutagenesis. L270I, Y271F, L274G, and C313F mutations were implemented. Y349F mutation was also implemented as a previous study showed that this mutation increases aminoacylation efficiency. With the novel aaRS, 3-methyl-histidine, 3-pyridyl-alanine, 2-furyl-alanine, 3-(2-thienyl)-alanine, and 2-(5-bromo-thienyl)-alanine were inserted into proteins both in *E.coli* and eukaryotic cells[29].

The o-tRNA$^{Pyl\text{-}opt}$ produced in the thesis is a version of the tRNA^{Pyl} with alterations in the acceptor and T stems (Figure 1.8a) that interact with the elongation factor-Tu (EF-Tu)[30]. As tRNA^{Pyl} is derived from archaea, it has been optimized for interaction with *E.coli* EF-Tu. The resulting o-tRNA$^{Pyl\text{-}opt}$ is approximately three times more efficient in the insertion of ncAAs into sfGFP (super-folder green fluorescent protein) variants (Figure 1.8b)[31].

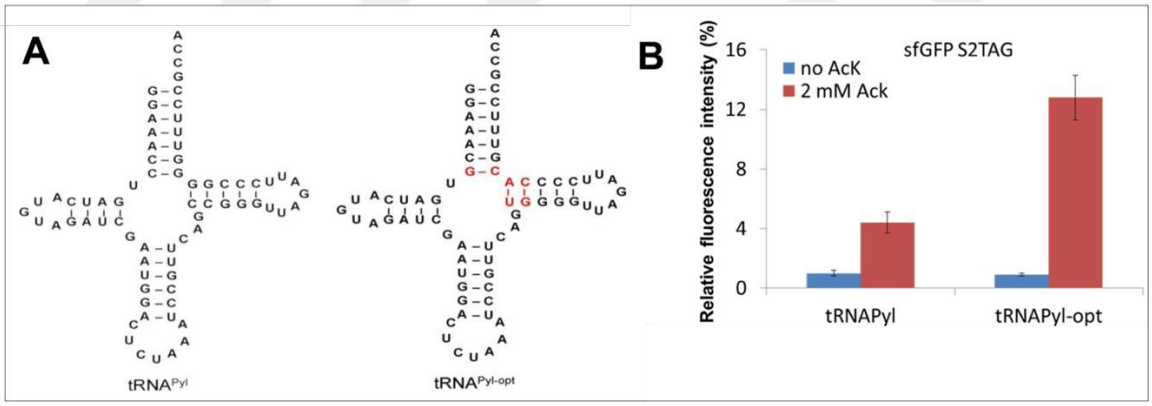


Figure 1.8 Comparison of tRNA^{Pyl} and tRNA$^{Pyl\text{-}opt}$ [31]. (a) Sequence comparison of the two tRNA molecules. (b) ncAA insertion efficiency of the tRNAs.

Cell-free protein synthesis

An effective technique that offers an alternative to traditional *in vivo* protein production is cell-free protein synthesis. PANOx-SP (PEP, amino acids, nicotinamide adenine dinucleotide (NAD), oxalic acid, spermidine, putrescine) system was used in the CFPS reactions[32]. The first stage of CFPS is fermenting the genetically modified *E. coli*

strains for optimum protein synthesis conditions. These alterations are mostly related to the inactivation of enzymes that break down amino acids, such as serine deaminases and arginine decarboxylase. Following the fermentation, the cells are either homogenized or subjected to chemical lysis to lyse them. To the resulting lysate, the following additions are made: amino acids and nucleotides, which are the building blocks of macromolecules, a carbon source that will provide energy to the system (e.g., phosphoenolpyruvate (PEP), glucose), glutamate salts that give additional power by supplying the central metabolism, coenzymes (NAD, coenzyme A), polyamines that maintain the stability of nucleic acids (putrescine, spermidine), purified T7 RNA polymerase that produces the mRNA, and the gene of the protein of interest. As the lysate is purified from genomic DNA, only the protein of interest is synthesized in the reaction mixture (Figure 1.9)[33].

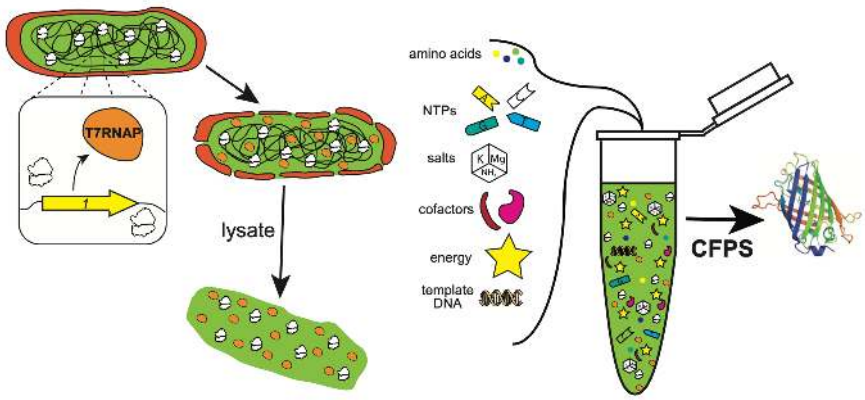


Figure 1.9 Schematic CFPS workflow from lysate preparation to *in vitro* protein production[33]. Reprinted by permission from Benjamin J. Des Soye et al, 2019,Elsevier [Cell Chemical Biology] License No: 5382560063598.

Since there is no live cell to consume energy, all the chemical energy in the solution can be used for protein production in CFPS. Furthermore, the lack of the cellular membrane makes manipulation of the reaction conditions and environment easier, such as radiolabeling, the addition of chaperones, ncAAs, and tRNA/tRNA synthetase pairs for ncAA insertion. With the CFPS systems, one can synthesize toxic proteins, membrane proteins, proteins with ncAAs, and other proteins that are difficult to express in vivo.

Cell extracts used in CFPS systems

ΔiscR strain

In bacteria, enzymes encoded by three operons are responsible for iron-sulfur cluster formation and incorporation. These are nitrogen fixation (NIF), sulfur utilization factor (SUF), and iron-sulfur cluster (ISC) operons[34]. In *E. coli, isc* operon encoded enzymes oversee the proper Fe-S cofactor assembly[35]. Previous studies show that if *isc* is overexpressed, it could increase ferredoxin, a protein that contains iron-sulfur clusters, production significantly. Similarly, work on other proteins that harbor Fe-S clusters proved that the formation and attachment of these clusters are limiting factors for most of the recombinant Fe-S proteins in *E.coli*[36][37][38]. By increasing the dosage of the *isc* gene, this problem could be overcome. Another strategy that has been proven to increase *isc* expression is the deletion of *iscR*, the transcriptional negative regulator of *isc*. Even though the deletion also affects the expression of 35 other genes, most of their roles and importance in *E.coli* is unknown[39].

Figure 1.10 shows that hydrogenase activity (black bars) is highest in the *ΔiscR* strain whether the culture is grown in the presence of 5% (v/v) oxygen or in anaerobic conditions. On the other hand, the amount of accumulated hydrogen is higher for the *ΔiscR* strain grown in aerobic conditions than the one grown in anoxic conditions. This is simply due to the final cell density of the aerobic culture being 5 times higher than the anaerobic one[40].

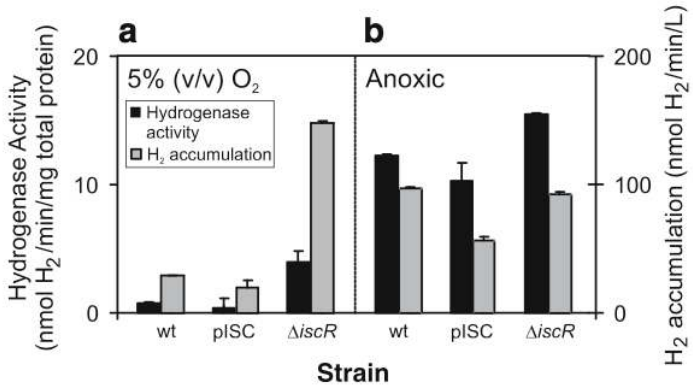


Figure 1.10 Change in hydrogenase activity and ferredoxin-dependent hydrogen production *in vivo* with *isc* overexpression and *iscR* deletion[40]. Each strain contains the plasmids that encode hydrogenase and maturation enzymes, and ferredoxin. The error bars represent standard error (n=2). (a) The cultures grown in 5% (v/v) oxygen. (b) The cultures grown anaerobically. Reprinted by permission from M. Kalim Akhtar et al, 2008,

759.T7.opt strain

To incorporate ncAAs into proteins, the amber suppression method is used. However, this suppression gets interrupted due to the competition with release factor 1 (RF1), a factor that terminates translation when the ribosome encounters a UAG stop codon[40]. If RF1 outcompetes the o-tRNA with an attached ncAA, a truncated protein is produced[41]. Thus, to increase the CFPS yields, an *E. coli* strain with an altered genome was designed. All native amber codons that the *E. coli* genome harbors were swapped with the ochre stop codon (UAA), which enabled the deletion of RF1 from the altered genome. Moreover, T7 RNA polymerase gene was inserted into the genome of the strain. Since the polymerase is subject to proteolysis catalyzed by OmpT protease, the gene sequence of the T7 polymerase was altered in a way to eliminate the cut sites, K183/R184 and K190/K191. For this purpose, K183G and K190L mutations were implemented. The resulting strain, 759.T7.Opt (called opt. in the thesis), yields significantly higher modified protein production and suppression efficiency[33].

BL21 & ARG1 strains

E.coli BL21 strain lacks the Lon and OmpT proteases, while the KC6-derived strain *E.coli* ARG1 strain lacks the OmpT protease[43]. Thus, these two strains would be ideal to be used as cell extracts where these proteases target the protein to be synthesized.

Chapter 2:
MATERIALS AND METHODS

2.1 Preparations for wild-type and mutant CpI Hydrogenase Production in CFPS

2.1.1 Production and Purification of o-tRNAs required for ncAA incorporation

pY71 minHH-otDNAOpt (#30) from lab stocks contains the optimized o-tRNA template used to incorporate ncAAs using *M. jannaschii*-derived aminoacyl-tRNA synthetases. This o-tRNA sequence was improved for better interaction with the endogenous *E.* coli release factor 1 (RF1)[44]. To create a linear tRNA template, a "large-scale" (0.5 ml) PCR reaction was performed with the primers in Table 2.1 and following the protocol given in Table 2.2.

Table 2.1 Primers used to amplify *M. jannaschii*-derived o-tRNA template from pY71 minHH-otDNAOpt plasmid and the final PCR product sequence.

Primer Name	Sequence	Purpose
pY71.seq.F	CCACCTCTGACTTGAGCGTCG	To amplify o-tRNA
tRNA.R	TGGTCCGGCGGAGGGGATTTG	To amplify o-tRNA
Final product	CCACCTCTGACTTGAGCGTCGATTTTTGTGATGCTCGTCAG GGGGGCGGAGCCTATGGAAACGAATTCAGATCTTAATAC GACTCACTATAGGGAGACCGGCTGATGAGTCCGTGAGGA CGAAACGGTACCCGGTACCGTCCCGGCGGTAGTTCAGCA GGGCAGAACGGCGGACTCTAAATCCGCATGGCAGGGGTT CAAATCCCCTCCGCCGGACCA	

Table 2.2 Large-scale PCR protocol to amplify *M. jannaschii* o-tRNA template.

Template	1 µL	98°C	5 min	
Primers	25 µL each (from 10mM stock solution)	98°C 55°C	30 sec 30 sec	x20
dNTP	1.5 µL	72°C	2 min	
Phusion Polymerase	10 µL	72°C	7 min	
Phusion Buffer	100 µL	12°C	∞	
dH_2O	Up to 500 µL			
The volume was divided into 10 PCR tubes, each containing 50 µL of the reaction.				

Following the reaction, ethanol precipitation was followed to purify the PCR products. The reaction mixture was collected in a 2 mL Eppendorf tube. 3M sodium acetate (NaOAc) was added to the PCR mixture with a final concentration of 0.3M. Ethanol was also added in 3X volume, and the mixture was incubated at -20ºC for 30 minutes. The tube was then centrifuged at 18,000g for 30 minutes at 4ºC. After removing the supernatant, the pellet was washed with 500 µL ethanol, and it was recentrifuged at the previous conditions. The supernatant was carefully discarded, and the pellet was dried at 40ºC for 30 minutes, followed by resuspension in 250 µL elution buffer of Nucleospin Gel and PCR Clean-up kit (Macherey-Nagel™ 740609.50).

For the $tRNA^{Pyl}$-derived orthogonal tRNA compatible with *M. barkeri*-derived aminoacyl-tRNA synthetases, 2 versions of the tRNA were prepared, one with a hammerhead ribozyme sequence ($tRNA^{Pyl\text{-}opt\text{-}ribo}$) (Figure 2.2) and one without ($tRNA^{Pyl\text{-}opt}$) (Figure 2.1). The primer design was done with Benchling software. First, a PCR assembly was performed to obtain the desired o-tRNA sequence with the primers in Table 2.3. T_m was adjusted to 50ºC for both cases, and $tRNA^{Pyl\text{-}opt}$ assembly reaction had only 3 cycles while $tRNA^{Pyl\text{-}opt\text{-}ribo}$ assembly had 20 cycles. The products were then amplified in a large-scale PCR reaction using the primers in Table 2.4 and T_m of 63ºC.

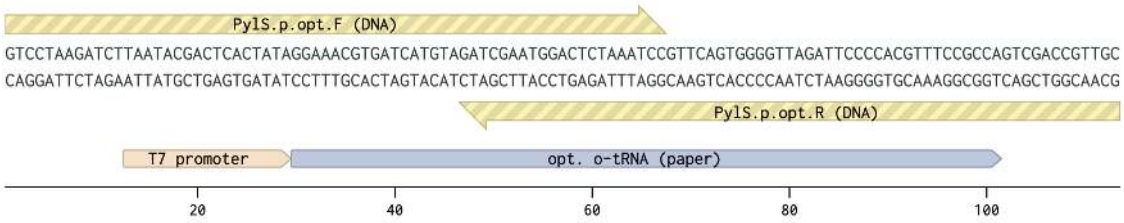


Figure 2.1 Planned PCR assembly design to create $tRNA^{Pyl\text{-}opt}$.

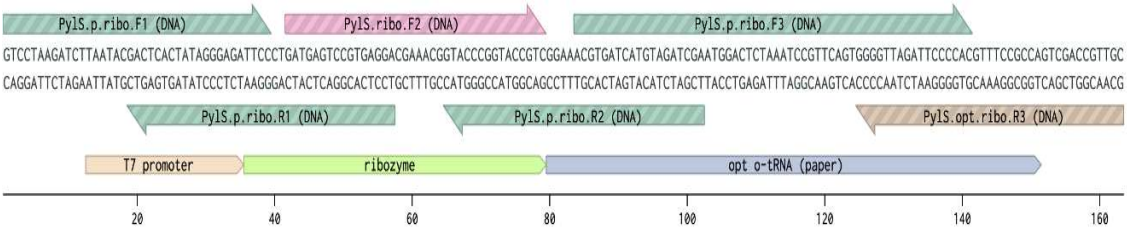


Figure 2.2 PCR assembly design to create $tRNA^{Pyl\text{-}opt\text{-}ribo}$.

Table 2.3 Primers used for PCR assembly of *M. barkeri*-derived o-tRNAs.

Primer Name	Sequence	Purpose
PylS.p.opt.F	GTCCTAAGATCTTAATACGACTC ACTATAGGAAACGTGATCATGT AGATCGAATGGACTCTAAATCC	PCR assembly of $tRNA^{Pyl\text{-}opt}$
PylS.p.opt.F	GCAACGGTCGACTGGCGGAAAC GTGGGGAATCTAACCCCACTGA ACGGATTTAGAGTCCATTCGATC	PCR assembly of $tRNA^{Pyl\text{-}opt}$
PylS.p.ribo.F1	GTCCTAAGATCTTAATACGACTC ACTATAGGGAGATTCC	PCR assembly of $tRNA^{Pyl\text{-}opt}$
PylS.p.ribo.R1	TCCTCACGGACTCATCAGGGAA TCTCCCTATAGTGAGTC	PCR assembly of $tRNA^{Pyl\text{-}opt\text{-}ribo}$
PylS.ribo.F2	GATGAGTCCGTGAGGACGAAAC GGTACCCGGTACCGTC	PCR assembly of $tRNA^{Pyl\text{-}opt\text{-}ribo}$
PylS.p.ribo.R2	TCGATCTACATGATCACGTTTCC GACGGTACCGGGTAC	PCR assembly of $tRNA^{Pyl\text{-}opt\text{-}ribo}$
PylS.p.ribo.F3	ACGTGATCATGTAGATCGAATG GACTCTAAATCCGTTCAGTGGG GTTAGATTCCCCAC	PCR assembly of $tRNA^{Pyl\text{-}opt\text{-}ribo}$
PylS.opt.ribo.R3	GCAACGGTCGACTGGCGGAAAC GTGGGGAATCTAACCCC	PCR assembly of $tRNA^{Pyl\text{-}opt\text{-}ribo}$
$tRNA^{Pyl\text{-}opt}$ product	GTCCTAAGATCTTAATACGACTCACTATAGGAAACG TGATCATGTAGATCGAATGGACTCTAAATCCGTTCA GTGGGGTTAGATTCCCCACGTTTCCGCCAGTCGACCG TTGC	
$tRNA^{Pyl\text{-}opt\text{-}ribo}$ product	GTCCTAAGATCTTAATACGACTCACTATAGGGAGAT TCCCTGATGAGTCCGTGAGGACGAAACGGTACCCGG TACCGTCGGAAACGTGATCATGTAGATCGAATGGAC TCTAAATCCGTTCAGTGGGGTTAGATTCCCCACGTTT CCGCCAGTCGACCGTTGC	

Table 2.4 Amplification PCR primers and final products for *M. barkeri*-derived optimized o-tRNA.

Primer Name	Sequence	Purpose
PylS.f.p.F	AGATCTTAATACGACTCACTATAGGAA ACGTGATCATG	Amplification of $tRNA^{Pyl\text{-}opt}$
PylS.f.opt.R	TGGCGGAAACGTGGGGAATCTAAC	Amplification of $tRNA^{Pyl\text{-}opt}$ or $tRNA^{Pyl\text{-}opt\text{-}ribo}$
PylS.p.ribo.F1	GTCCTAAGATCTTAATACGACTCACTA TAGGGAGATTCC	Amplification of $tRNA^{Pyl\text{-}opt\text{-}ribo}$

Final product ($tRNA^{Pyl\text{-}opt}$)	AGATCTTAATACGACTCACTATAGGAAACGTGATCATGTA GATCGAATGGACTCTAAATCCGTTCAGTGGGGTTAGATTC CCCACGTTTCCGCCA
Final product ($tRNA^{Pyl\text{-}opt\text{-}ribo}$)	GTCCTAAGATCTTAATACGACTCACTATAGGGAGATTCCC TGATGAGTCCGTGAGGACGAAACGGTACCCGGTACCGTC GGAAACGTGATCATGTAGATCGAATGGACTCTAAATCCGT TCAGTGGGGTTAGATTCCCCACGTTTCCGCCA

2.1.2 Cloning of aminoacyl-tRNA synthetase (aaRS) genes into pY71 vector

4 aaRS genes (CNFRS, pAFRS, pIFRS, OMeYRS) were planned to be cloned into the pY71 vector by restriction-ligation method. pEVOL OMeYRS, pDule pAFRS, and pEVOL pIFRS plasmids that served as gene sources were provided by Prof. Peter Schultz at The Scripps Research Institute (La Jolla, CA). pDule CNFRS and pULTRA CNFRS plasmids were obtained from Addgene Plasmid Bank, and the HRS gene was synthesized by Twist Biosciences and cloned into one of their proprietary vectors. All DNA and protein sequence analyses, as well as primer design were done on Benchling software. Initially, primers were designed to include 5'-NdeI and 3'-SalI restriction sites along with a C-terminal 6xHis affinity marker to the genes of interest, except for HRS as it already contained the necessary restriction sites and affinity marker. Table 2.5 shows the primers used for the additions.

Table 2.5 Primers used for incorporation of NdeI and SalI cut sites along with C-terminal 6xHis-tag.

Primer Name	Sequence	Purpose
CNFpAFpIFOMeY.F	AAGGAGATATACATATGATGGACGAATTTGAAATGATAAAGA	Incorporation of 5'-NdeI cut site
CNFpAFpIFOMeY.R	GGCCGGTCGAGTCGACTTAGTGATGGTGATGGTGATGTAATCTCTTTCTAATTGGCTCTAAAATC	Incorporation of C-terminal 6xHis-tag and 3'-SalI cut site

With the designed end primers, aaRS genes were amplified by PCR (Table 2.6), and the PCR products were purified using Nucleospin Gel and PCR Clean-up kit (Macherey-Nagel™ 740609.50).

Table 2.6 PCR procedure to insert restriction sites and the 6xHis affinity tag.

Template	1.2 µL	98°C	5 min	
Primer	3 µL each (10mM stock solution)	98°C	30 sec	
dNTP	1.5 µL	55°C	30 sec	x30
Phusion Polymerase	1.5 µL	72°C	1 min	
Phusion Buffer	12 µL	72°C	7 min	
dH_2O	Up to 60 µL	12°C	∞	

The purified PCR products and a pY71 plasmid containing the anakinra gene from the lab stocks were double digested with NdeI and SalI restriction enzymes overnight at 37ºC (Table 2.7). Then the genes were ligated to pY71 vectors via overnight incubation. The cloning efficiency was checked with agarose gel electrophoresis. 5 µL from each reaction mixture was run on an agarose gel to confirm the digestion. Then the mixture was purified by Nucleospin Gel and PCR Clean-up kit (Macherey-Nagel™ 740609.50) as gel purification was observed to lower the yield of purification. Afterwards, the insert and vectors were ligated together in an overnight reaction at RT (Table 2.8). The next day, 5 µL of the reaction mixture was transformed into chemically competent *E. coli* DH5α cells. 3 colonies were picked from each plate and made into a liquid culture. The cultures were then miniprepped using Monarch® Plasmid Miniprep Kit (#T1010S) and a check restriction digest was performed by double digesting the purified plasmids and running the reaction mixture on an agarose gel. After verification by restriction digestion, the samples containing the genes of interest and the pY71 vector were sent for sequencing.

Table 2.7 Restriction reaction reagents to cut the aaRS genes and pY71 vector.

NdeI	1.8 µL	37ºC overnight
SalI	1.8 µL	65 ºC 10 min
DNA	1200 ng	
3:1 buffer	6 µL	
dH_2O	Up to 60 µL	

Table 2.8 Ligation reaction setup.

Vector	10 ng	RT overnight 65 ºC 10 min
Insert	(1:5 vector:insert ratio)	
T4 Buffer	2 µL (NEB #B0202)	
T4 DNA Ligase	1 µL (NEB #M0202L)	
dH_2O	Up to 20 µL	

Once the sequence was confirmed, the plan was to commence protein production. However, it was noticed that there was an extra methionine at the beginning of the aaRS sequences due to a mistake in the initial primer design. These methionines were removed by SDM using the primers in Table 2.9 and protocol in Table 2.10.

Table 2.9 Primers used to remove the extra methionine via SDM.

Primer Name	Sequence	Purpose
Mj.SDM.F	GAAGGAGATATACATATGGACGAATTT GAAATGATAAAG	Site-directed mutagenesis
Mj.SDM.R	CTTTATCATTTCAAATTCGTCCATATGTA TATCTCCTTC	Site-directed mutagenesis

Table 2.10 SDM procedure to remove the extra methionine in N-terminal of aaRS genes.

Template	0.5 µL	98°C	5 min	
Primers	1 µL each (from 10mM stock)	98°C	30 sec	x30
dNTP	0.5 µL	55°C	30 sec	
Phusion Polymerase	0.5 µL	72°C	2 min	
Phusion Buffer	4 µL	72°C	7 min	
dH_2O	Up to 20 µL	12°C	∞	
DpnI digestion for 2 hours at 37°C				

After the PCR reaction, DpnI (NEB #R0176) digestion was performed at 37°C for 2 hours to remove the original methylated DNA put into the PCR reaction. Then, PCR purification was performed with Nucleospin Gel and PCR Clean-up kit (Macherey-Nagel™ 740609.50), and 5 µL of the reaction mixture was transformed into chemically competent *E. coli* DH5α cells. 2 colonies from each plate were grown in liquid culture

and, after miniprepping with Monarch Plasmid Miniprep Kit (NEB #T1010S), sent for sequence confirmation. When confirmed, to the cells were taken to protein production step.

2.1.3 Aminoacyl-tRNA synthetase production & purification

Production and purification of pPaFRS

The gene for pPaFRS was received from Prof. James Swartz at Stanford University with a 6xHis-tag in pY71 vector. The plasmid was transformed into *E. coli* BL21(DE3) pLysS cells. A starter culture was grown overnight at 37ºC in 30 mL selective 2YT medium containing 25 ug/mL chloramphenicol and 50 ug/mL kanamycin. The following morning, 1L fresh 2YT medium was inoculated using the overnight culture. When 0.6 OD_{600} was reached, enzyme expression was induced with 1 mM IPTG. After 3 hours of induction, the culture was harvested by centrifugation at 3,000 g for 15 minutes at 4ºC. The supernatant was decanted, and the cell pellet was washed with 100 mL cold S30 buffer. One more round of centrifugation was done, and the cell pellet was stored at -80ºC.

To purify the protein containing a C-terminal 6xHis-tag, IMAC method was selected. For this purpose, HisTrap FF 1mL columns and ÄKTA Pure fast protein liquid chromatography (FPLC) system were used. Initially, the cell pellet was suspended in binding buffer containing EDTA-free protease inhibitor (Roche cOmplete Tablets, Mini EDTA-free, *EASY* pack, Cat. No: 4693159001). 1 tablet was dissolved in 2mL dH_2O. Per g of cell pellet, 0.2 mL protease inhibitor solution and 5 mL binding buffer were used. Later, the cells were lysed by passing the cell suspension once through the Emulsiflex C-5 homogenizer between 17,500 and 25,000 psi. To remove insoluble proteins and cell debris, the mixture was centrifuged at 12,000 g and 4ºC for 30 minutes. In the meantime, 1 ml HisTrap FF affinity column (GE Healthcare Cat. No: 17-5319-01) was equilibrated by washing with 10 CV binding buffer. The supernatant was filtered and loaded onto the column using the Superloop™ apparatus for purification. Imidazole (500mM) was used with gradient elution as the competitive reagent. IMAC buffers and procedure can be seen in Tables 2.11 and 2.12, respectively.

Table 2.11 Buffers used for purification of 6xHis-tagged proteins.

IMAC Buffers	Ingredients
Binding Buffer	20mM sodium phosphate 0.5M NaCl 30mM imidazole
Elution Buffer	20mM sodium phosphate 0.5M NaCl 500mM imidazole

Table 2.12 FPLC procedure to purify 6xHis-tagged pPaFRS.

Step	Buffer	Volume
Equilibration	Binding buffer	10 CV
Sample Application		
Column wash	Binding buffer	10 CV
Elution	Linear gradient: Binding + Elution buffer	15 CV
Column Wash	0.5M NaOH	10 CV
Column Wash	dH_2O	15 CV
Column Wash	20% Ethanol	10 CV

The elution fractions containing the peak corresponding to the protein of interest were collected and separately dialyzed against 10mM potassium phosphate buffer (pH 7.4) containing 0.05% Tween20 for 1 hr. After one more hour of dialysis with fresh buffer, the fractions were left overnight with potassium phosphate buffer containing 20% sucrose (pH 7.4). The following day, the dialysis fractions were collected and centrifuged at 12,000 g for 15 minutes at 4ºC to remove protein aggregates. Afterwards, the purified soluble protein was stored in -80ºC. SDS-PAGE and DC assay were performed to determine purity and concentration of the protein, respectively.

Production and purification of the other M. jannaschii-derived aaRSs

The plasmids containing the genes aaRss (CNFRS, pAFRS, OMeYRS, pIFRS) were transformed into *E.coli* BL21(DE3) pLysS cells. The enzymes were prepared as explained in the previous section.

Production and purification of pAMFRS

The pAMFRS-His6x gene was previously constructed in the lab by Yağmur Ersoy. The production and purification of the enzyme was done using the methods described in the '*Production and purification of pPaFRS*' section above, apart from the dialysis buffers used. It was previously observed in the lab that pAMFRS enzyme forms aggregates during dialysis when a buffer with a pH 7.4 is used. Therefore, a study was performed together with Yağmur Ersoy by varying dialysis buffers used. For this, a literature search was done to find buffers that increase stability and solubility of the proteins. First, a fresh batch of the enzyme was produced and purified. All the buffers mentioned in Table 2.13 except PBS were prepared inside a 10mM potassium phosphate buffer with pH of 8.4. The elution fractions were first dialyzed for 1 hour at room temperature in the buffers mentioned at Table 2.13 and then with a fresh buffer they were left for overnight dialysis at 4ºC. Then DC assay was performed for each fraction to determine protein concentration.

Table 2.13 Buffers used in the dialysis study for pAMFRS[45].

Buffer name	**Ingredient**
Sodium citrate	0.5M Sodium citrate in 10mM potassium phosphate
NaCl	1M NaCl in 10mM potassium phosphate
PEG	1% w/v PEG800 in 10mM potassium phosphate
SDS	0.1% w/v SDS in 10mM potassium phosphate
PBS	0.137M NaCl 0.0027M KCl 0.01M Na_2HPO_4 0.0018M KH_2PO_4
Imidazole	50mM imidazole in 10mM potassium phosphate
Arginine	0.1M arginine in 10mM potassium phosphate
Potassium phosphate	10mM potassium phosphate

2.1.3 CFPS to insert ncAAs to sfGFP

PANOx SP system was used in CFPS reactions, and a reaction would consist of 3 replicates, each being 20 µL in volume in 2 mL Eppendorf tubes, unless stated otherwise. The reactions were incubated at 30 ºC for 16 hours. pY71 sfGFP plasmid was used to produce wild-type sfGFP as a positive control. To produce sfGFP variants that contain ncAAs, pY71 sfGFP23TAG, pY71 sfGFP39TAG or pY71 sfGFP151TAG plasmids were used. In reactions where ncAA incorporation is aimed, 2 mM ncAA of interest, 382 nM o-tRNA template, 0.5 mg/mL aaRS of interest were added to the reactions aside from the reagents in Table 2.14. The cell extract changed depending on the experiment.

Table 2.14 List of reagents in a CFPS reaction.

Magnesium glutamate	8 mM
Ammonium glutamate	10 mM
Potassium glutamate	175 mM
ATP	1.2 mM
GMP	0.85 mM
UMP	0.85 mM
CMP	0.85 mM
Folinic acid	34 µg/ml
20 amino acids	2 mM each
Phosphoenolpyruvate (PEP)	33.3 mM
Nicotinamide adenine dinucleotide (NAD)	0.33 mM
Coenzyme A (CoA)	0.26 mM
Oxalic acid (OA)	2.7 mM
Spermidine	1.5 mM
Putrescine	1 mM
T7 RNA polymerase	0.1 mg/mL
Plasmid	6 nM
Cell extract	25% v/v

2.1.4 Site-directed mutation of CpI hydrogenase

The designed primers to mutate 4 predetermined locations on CpI hydrogenase enzyme to TAG amber stop codon with SDM can be seen in the Table 2.15 and the SDM procedure in Table in Table 2.16.

Table 2.15 Primers designed to incorporate TAG stop codon to CpI hydrogenase gene.

Primer Name	Sequence	Purpose
H192.F	CGATGACACCAATTGTCTGTAGTGTGGTCA ATGTATCATCG	L192TAG mutation
H192.R	CGATGATACATTGACCACACTACAGACAA TTGGTGTCATCG	L192TAG mutation
H194.F	GACACCAATTGTCTGCTGTGTTAGCAATGT ATCATCGCCTGTCCA	G194TAG mutation
H194.R	TGGACAGGCGATGATACATTGCTAACACA GCAGACAATTGGTGTC	G194TAG mutation
H356.F	CAAAGAATGTATTTACTGTAACCGTTATGC CGTGTTAGTCCAAAAAATTTGAAGCAGA	T356TAG mutation
H356.R	TCTGCTTCAAATTTTTTGGACTAACACGGC ATAACGGTTACAGTAAATACATTCTTTG	T356TAG mutation
H357.F	TGTAACCGTTATGCCGTGTACTTAGAAAA AATTTGAAGCAGATCGTC	S357TAG mutation
H357.R	GACGATCTGCTTCAAATTTTTTCTAAGTAC ACGGCATAACGGTTACA	S357TAG mutation

Table 2.16 SDM procedure to incorporate TAG stop codon to CpI hydrogenase gene.

Template	0.5 µL	98°C 5 min
Primers	1 µL each (from 10mM stock solution)	98°C 30 sec 55°C 30 sec 72°C 4 min x20
dNTP	0.5 µL	
Phusion Polymerase	0.5 µL	72°C 7 min
Phusion Buffer	4 µL	12°C ∞
dH_2O	Up to 20 µL	
DpnI (0.5 µL) digestion for 2 hours at 37°C.		

After the SDM reaction, DpnI (NEB #R0176) digestion was done at 37°C for 2 hours to remove the methylated original DNA put into the PCR reaction. Then, it was followed by PCR purification with Nucleospin Gel and PCR Clean-up kit (Macherey-Nagel™ 740609.50), and 5 µL of the reaction mix was transformed into chemically competent *E. coli* DH5α cells. From each plate, 3 colonies were picked and after miniprepping, they were sent for sequence confirmation.

2.2 Production of M. barkeri derived HRS

2.2.1 Cloning and SDM of HRS gene

The gene for HRS, an M. barkeri derived aminoacyl tRNA synthetase, containing 5'-NdeI, 3'-SalI and C-terminal 6xHis affinity tag was ordered from Twist in a clonal vector. The gene was cloned into the pY71 vector upon its arrival. Before moving to the production of the enzyme, it was noticed that the gene had an extra methionine sequence at the N-terminal. Therefore, standard SDM procedure was performed to remove it. The designed primers can be seen in the Table 2.17 below. After the sequence confirmation by Macrogen, the plasmid was transformed into *E. coli* BL21 (DE3) pLysS cell line. The production and purification protocol described for *M.jannaschii* derived aaRSs was then followed.

Table 2.17 Primers used to remove the extra methionine via SDM.

Primer Name	Sequence	Purpose
HRS.SDM.F	TTAAGAAGGAGATATACATATGGACA AAAAACCGCTG	Site-directed mutagenesis
HRS.SDM.R	CAGCGGTTTTTTGTCCATATGTATATCT CCTTCTTAAAG	Site-directed mutagenesis

2.2.2 Expression study of HRS with varying temperatures, IPTG concentrations, and additions

Expression at 25ºC and 30ºC and 0.25 mM & 1 mM IPTG concentration

Alternative to the classic expression method, expression study with 0.25 mM and 1 mM IPTG concentrations and varying temperatures was performed. For this purpose, a 120 mL culture was grown at 37ºC in 2YT media supplemented with kanamycin and chloramphenicol until OD_{600} reached 0.6, then it was split into 4 falcons. The different reaction condition combinations can be seen below.

- 30°C – 0.25 mM IPTG
- 30°C – 1 mM IPTG
- 25°C – 0.25 mM IPTG
- 25°C – 1 mM IPTG

After induction with IPTG, protein expression was carried overnight. The following morning, the cultures were harvested by centrifugation at 3,000 g for 15 minutes. The pellets were dissolved in binding buffer with 1:5 ratio and lysed by sonication for 9 cycles. Then the lysates were centrifuged at 12,000 g for 30 minutes. The resulting pellets and supernatants were analyzed with SDS-PAGE.

Expression study at 4ºC

BL21 (DE3) pLysS pY71 HRS cell line was grown at 37ºC in 30 mL 2YT media supplemented with kanamycin and chloramphenicol in a 50 mL falcon tube until OD_{600} reached 0.6. Induction was done with 0.5 mM IPTG, and the temperature of the incubator shaker was switched to 4ºC. The protein expression was continued for 4 days (104.7 hours). After preparing the lysate as mentioned in the previous paragraph, it was centrifuged at 12,000 for 30 minutes. The resulting pellets and supernatants were analyzed with SDS-PAGE.

Expression study with glycerol and ethanol additives at different temperatures

4 falcon tubes filled with 30mL 2YT media supplemented with kanamycin and chloramphenicol were prepared. The effect of ethanol[23] and glycerol[24] addition to the

media on protein solubility was tested. To 2 of the tubes 2% ethanol was added and to the other 2, 15% glycerol and 2% ethanol were added. Each tube was inoculated with an overnight starter culture of BL21 (DE3) pLysS pY71 HRS cell line and the cultures were grown at 37ºC until OD_{600} reached 0.6. Protein expression under T7 promoter was induced with 1 mM IPTG. One tube from each condition was placed in an incubator shaker set to 27ºC and the other two were left in 37ºC shaker. All expressions were continued overnight. After preparing the lysates as mentioned before, they were centrifuged at 12,000 g for 30 minutes. The resulting pellets and supernatants were analyzed with SDS-PAGE.

2.2.3 Production of HRS in CFPS

HRS was also produced in CFPS. The effect of addition of different concentrations of Skp, GroEL, and GroES chaperones, varying ratios of reduced and oxidized glutathione, protease inhibitor concentration, cell extract type, reaction temperature and length on protein solubility was tested. All the tested conditions can be seen in Table 2.30.

Cell extract preparation from E. coli BL21 and E. coli ARG1 cell lines

30 mL starter cultures were grown overnight for both *E. coli* BL21 and *E. coli* ARG1 cell lines. The following morning, the necessary media were prepared, as indicated in Table 2.18. 1 L LB medium was inoculated with *E. coli* BL21 overnight culture and 1 L M9 medium was inoculated with *E. coli* ARG1 overnight culture. The cultures were grown until hr^{-1} value fell below 0.6 and then harvested by centrifugation at 3,000g for 15 minutes at 4ºC. The pellets were stored at -80ºC. Alternatively, ARG1 cell line was also grown in LB medium with the described conditions. To prepare the cell extracts, S30 extract preparation protocol was followed. The day of extract preparation, 1 mL S30 buffer per gram wet cell and the pellets were incubated on ice for 1.5 hours to thaw them. Later, the cell suspensions were lysed using the Emulsiflex C-5 homogenizer between 17,500 and 25,000 psi. The lysates were collected in separate 30 mL centrifuge tubes, and they were cleared at 30,000g for 30 minutes at 4ºC. The resulting supernatant were centrifuged once more at the same conditions. The final supernatants were taken to clean tubes, wrapped in aluminum foil, and incubated on a rotary shaker at 37ºC and 120 rpm

for 80 minutes. The cell extracts were then dialyzed against 1 L S30 buffer for 30 minutes 4 times using a 7K MWCO dialysis membrane. The extracts were then centrifuged at 10,000 g for 20 minutes at 4ºC and the final supernatants were aliquoted. Then they were incubated at -150ºC for 5 minutes and stored at -80ºC.

Table 2.18 Required solutions for cell extract preparation.

5X M9 Salts	21.4g $Na_2HPO_4.2H_2O$ 7.5 g KH_2PO_4 1.25 g NaCl 6.6g $(NH_4)_2SO_4$ Up to 500 mL dH_2O
M9 Media (pH 7.0 with HCl)	200 mL 5X M9 Salts 10 mL filter-sterilized 20% Glucose 1 mL autoclaved 2M $MgSO_4$ 100 µL autoclaved $CaCl_2$ Up to 1L autoclaved dH_2O
S30	10 mL 1M Tris-acetate pH 8.2 10 mL 6M Potassium acetate 10 mL 1.4M Magnesium acetate

CFPS reactions

The protocol in Section 2.1.3 and conditions in the Table 2.14 were followed. After the reactions, they were incubated at -20ºC for 5 minutes. The Eppendorf tubes were briefly vortexed and 5 µL from each reaction replica was taken and pooled in separate PCR tubes. The Eppendorf tubes were then centrifuged at 20,000g for 15 minutes and again 5 µL from each reaction replica was taken. The samples were prepared to load on SDS-PAGE gels along with a pure pPaFRS protein previously purified by Yağmur Ersoy to serve as positive control. 8 µL from each sample was loaded to 2 separate gel and run at 100V. After the run, one of the gels were taken for Western-blot analysis.

Table 2.19 Tested HRS production conditions in CFPS

Addition	Cell extract	Temperature	Reaction length
–	759.T7 opt.	30ºC	16 hours
1 µM Skp	759.T7 opt.	30ºC	16 hours
4 µM Skp	759.T7 opt.	30ºC	16 hours
4 µM GroEL, 1µM GroES	759.T7 opt.	30ºC	16 hours
4mM oxidized glutathione (GSSG), 1mM reduced glutathione (GSH)	759.T7 opt.	30ºC	16 hours
3mM oxidized glutathione (GSSG), 2mM reduced glutathione (GSH)	759.T7 opt.	30ºC	16 hours
2.5mM oxidized glutathione (GSSG), 2.5mM reduced glutathione (GSH)	759.T7 opt.	30ºC	16 hours
1 µM GroEL, 1 µM GroES	759.T7 opt.	30ºC	16 hours
4 µM GroEL, 4 µM GroES	759.T7 opt.	30ºC	16 hours
4 µM Skp, 1X Protease inhibitor	759.T7 opt.	30ºC	16 hours
4 µM Skp, 5X Protease inhibitor	759.T7 opt.	30ºC	16 hours
4mM oxidized glutathione (GSSG), 1mM reduced glutathione (GSH), 1X protease inhibitor	759.T7 opt.	30ºC	16 hours
4mM oxidized glutathione (GSSG), 1mM reduced glutathione (GSH), 5X protease inhibitor	759.T7 opt.	30ºC	16 hours
4 µM Skp, 4mM oxidized glutathione (GSSG), 1mM reduced glutathione (GSH)	759.T7 opt.	30ºC	16 hours

4 µM Skp, 4mM oxidized glutathione (GSSG), 1mM reduced glutathione (GSH), 1X protease inhibitor	759.T7 opt.	30ºC	16 hours
4 µM Skp, 4mM oxidized glutathione (GSSG), 1mM reduced glutathione (GSH), 1X protease inhibitor	759.T7 opt.	30ºC	16 hours
4 µM Skp	ARG1	30ºC	16 hours
4 µM Skp, 1X protease inhibitor	ARG1	30ºC	16 hours
4 µM Skp	BL21	30ºC	16 hours

A large-scale (3 mL) CFPS reaction was also performed with 4 µM Skp and 1X protease inhibitor additions and BL21 cell extract at 30ºC for 16 hours in 6-well plates. The reaction was then incubated at -20ºC for 5 minutes. A sample was taken for SDS-PAGE and Western-blot analysis and then the reaction was centrifuged at 20,000 g for 15 minutes. After taking a sample from the supernatant for analysis, it was filtered and loaded onto FPLC. The purification and dialysis were proceeded as described in Section 2.1.3.

Western-blot analysis of CFPS produced C-terminal 6xHis-tagged HRS

Wet transfer method was followed to transfer the proteins on the SDS-PAGE gel to the membrane with the indicated buffers in Table 2.20, First the blotting filter papers, and the SDS-PAGE gel were washed with the transfer buffer. Then, the foam pads were placed to the either side of the gel holder cassette. A membrane at the size of the gel was cut and wet in methanol for 30 seconds followed by wash in transfer buffer. To the black side of the cassette, the SDS-PAGE gel was placed and on top of it the membrane was placed carefully with a tweezer. Then the cassette was closed and placed into the electrode assembly and the assembly was put into the tank with a cooling unit. The transfer was done at 100V for 1 hour. Alternatively, the transfer was also done at 25V overnight at 4ºC.

Table 2.20 Buffers required in Western-Blot analysis of 6xHis-tagged proteins.

Transfer Buffer	2.9 g Tris 14.4 g Glycine 200 mL Methanol Up to 1 L dH_2O
TBST	10 mL 1 M Tris-HCl pH 7.4 9 g NaCl 2 mL Tween-20 Up to 1 L dH_2O
Blocking Buffer	5% w/v skim milk powder in TBST
TBS	1 mL 1 M Tris-HCl pH 7.4 0.9 g NaCl Up to 100 mL dH_2O

After the transfer, the membrane was put inside a 50 mL falcon tube with a tweezer and washed with 10 mL TBST for 5 minutes. In the meantime, 15 mL blocking buffer was prepared, and the blocking was done by gently rotating the membrane on a rotator for 1.5 hours at RT. Afterwards, the membrane was washed 5 times with 10 mL TBST. His-tag (27E8) Mouse mAB (HRP conjugate) (Cell Signaling #9991) was diluted in 10 mL blocking buffer with 1:2000 ratio. Antibody conjugation was done overnight at 4ºC. The next day the membrane was washed with TBST 4 times and 1 time with TBS at RT for 5 minutes. Signal detection mixture was prepared by mixing 200 µL of each detection chemical and 200 µL dH_2O, making a total of 600 µL. The membrane was then placed on the detection tray and the mixture was pipetted on top of the membrane. The tray was placed inside the machine and from the software 'Chemiluminescence' imaging was picked. After adjusting the focus of the camera, signal detection was done.

2.2.4 SDM to obtain an OmpT protease resistant HRS

K104/K105 on HRS was identified to be the main target of OmpT protease. To remove this cut site, K105 was mutated to alanine by SDM with the protocol described in Section 2.1.1 and primers in Table 2.21, with the only change being the T_m in PCR

reaction was set to 60ºC. The resulting plasmids were sent to Macrogen for sequence confirmation.

Table 2.21 Primers used to change lysine at position 105 to alanine.

Primer Name	Sequence	Purpose
HRS.K105A.F	CACCGAAGGTTAAAGCAGCAATGCCGAAAAG	K105A mutation
HRS.K105A.R	CTTTTCGGCATTGCTGCTTTAACCTTCGGTG	K105A mutation

2.3 Production, Purification and Analysis of CpI Holo-Hydrogenase

2.3.1 Anaerobic production and purification of CpI Hydrogenase

E. coli BL21 (DE3) ΔiscR pACYC HydEFG competent cell preparation

Mix & Go! E. coli Transformation Kit (Zymo Research #T3001) was used to make competent cells from *E. coli* BL21 (DE3) ΔiscR pACYC HydEFG cell line. A 5 mL starter culture was grown overnight at 37ºC. The following morning, a 50 mL LB medium supplemented with 50 μg/mL kanamycin and 125 μg/mL chloramphenicol was inoculated with 1 mL overnight culture. The cells were grown at 37ºC until OD_{600} 0.4-0.6 was reached. The culture was incubated on ice for 10 minutes. In the meantime, 2X Wash and Competent Buffers provided with the kit were diluted in Dilution Buffer which was also in the kit with 1:1 ratio. The culture was then harvested by centrifugation at 2,500g for 10 minutes at 4ºC and the supernatant was discarded. The pellet was resuspended in 5 mL 1X Wash Buffer, and the suspension was centrifuged once more at previous conditions. The resulting pellet was resuspended in 5 mL 1X cold Competent Buffer. The ready-to-use competent cells were aliquoted and stored at -80 ºC.

Production and purification of CpI Hydrogenase in vivo under the T7 promoter

pET21b sCpI sIIC StrepII plasmid carrying a C-terminal StrepII-tagged CpI hydrogenase enzyme was transformed into *E. coli* BL21 (DE3) ΔiscR pACYC HydEFG competent cells simply by mixing 1 µL plasmid with 50 µL competent cell and spreading the mixture onto an LB-agar plate containing 50 µg/mL kanamycin, 125 µg/mL chloramphenicol, and 100 µg/mL ampicillin. The resulting cell line was used in the *in vivo* production of CpI.

A 5 mL starter culture was grown overnight at 37ºC in LB medium supplied with 50 µg/mL kanamycin, 125 µg/mL chloramphenicol, and 100 µg/mL ampicillin. The following day, a 50mL LB medium supplemented with the appropriate antibiotics and additions (100mm MOPS/KOH pH 7.8, 0.5% w/v glucose, and 2mM ferric ammonium citrate) was inoculated with the overnight culture. After growing the cells at 37ºC until OD600 ~0.5-0.6, the culture was transferred to a 1L of LB media with the mentioned supplementations. The cells were grown at 37ºC until OD600 ~0.5-0.7, then the culture was transferred into an anaerobic glove box and a stir bar was added. While mixing on a stir plate, 10mM sodium fumarate and 5mM L-cysteine were added to the media. After 15 minutes of stirring, protein production was induced with 0.25 mM IPTG. The culture was left for overnight induction. The next morning, the cells were harvested by centrifugation at 6,000 rpm for 20 minutes. The pellet was suspended in a lysis buffer and mixed for 30 minutes in the rotator at RT. The lysate was then centrifuged for 16,000 rpm for 10 minutes. The supernatant was taken for purification in an anaerobic glovebox.

StrepTactin affinity chromatography was performed to purify CpI Hydrogenase using StrepTactin Sepharose High Performance resin (Cytiva #28935599). Prior to the purification, 5 mL of the resin was poured into three 10 mL Pierce™ Disposable Columns (ThermoFisher Cat. No: # 29924), and it was allowed to settle by gravity. The buffers and purification procedure can be found in Tables 2.22 and 2.23, respectively. SDS-PAGE and Bradford assay were performed to determine purity and concentration of the protein. After, the protein was concentrated 2-fold using a 0.5mL Pierce Protein concentrator (PES, 10K MWCO, Thermo Scientific Cat. No: 88513).

Table 2.22 Buffers used with StrepTactin Sepharose resin for CpI hydrogenase purification.

Lysis Buffer	1X Bugbuster 100mM HEPES/KOH pH 8.4 25U/mL benzonase 1KU/mL lysozyme 1 EDTA-free protease inhibitor tablet (Roche) 0.1% Tween-20 1mM fresh dithionite (added directly as powder)
Wash Buffer	50mM HEPES/KOH pH 8 150mM KCl 0.1% Tween-20 0.25mM fresh dithionite (added directly as powder)
Elution Buffer	50mM HEPES/KOH pH 8 150mM KCl 0.1% Tween-20 0.25mM fresh dithionite (added directly as powder) 2.5mM desthiobiotin
Regeneration Buffer	0.5M NaOH

Table 2.23 Purification protocol of CpI Hydrogenase by StrepTactin affinity chromatography.

Step	Buffer	Volume
Equilibration	Wash Buffer	5 CV
Sample Application		
Column wash	Wash Buffer	1 CV
Elution	Elution Buffer	2 CV
Regeneration	0.5M NaOH	5 CV
Column Wash	dH_2O	5 CV
Column Wash	20% Ethanol	3 CV

Alternatively, the protein was also purified anaerobically using Strep-Tactin®XT® 4Flow trial kit (IBA Lifesciences, Cat. No: 2-5999-999) with the protocol in Table 2.25. The kit was composed of a 0.2 mL Strep-Tactin®XT® 4Flow pre-packed gravity column, Buffer W, and Buffer BXT (ingredients given in Table 2.24). 1 mM $MgCl_2$ was added both to Buffer W and BXT to chelate EDTA in the buffers before purification (1:1 stoichiometry) as EDTA inactivates metalloproteins[46]. The sample was filtered through a 0.45 μm syringe filter before loading to prevent clogging of the column. After purification was finished, the column was regenerated by running 6 CV of 10 mM NaOH. The column was then washed with Buffer W 2x4CV and stored at 4ºC. The purified protein was run on an SDS-PAGE gel, and Bradford assay was performed to determine protein concentration. The protein was subsequently concentrated using a 0.5 mL Pierce Protein concentrator (PES, 10K MWCO, Thermo Scientific Cat. No: 88513).

Table 2.24 Buffers used with Strep-Tactin®XT® 4Flow kit for CpI hydrogenase purification.

Buffer W	100mM Tris/HCl pH 8.0 150mM NaCl 1mM EDTA
Elution Buffer	10mM Tris/HCl pH 8.0 150mM NaCl 1mM EDTA 50mM biotin

Table 2.25 Purification procedure for CpI hydrogenase with StrepTactin affinity chromatography.

Step	Buffer	Volume
Equilibration	Buffer W	2 x 1 CV
Sample Application		
Column wash	Buffer W	5 x 1 CV
Elution	Buffer BXT	0.6 CV (E1) 1.6 CV (E2) 0.8 CV (E3)
Regeneration	10mM NaOH	6 CV

Column Wash	Buffer W	2 x 4 CV
Column Wash	20% Ethanol	3 CV

2.3.2 Production and purification of Rice root FNR (RrFNR)

E. coli BL21 (DE3) pET21B RrFNR cell line was used for production of FNR from *Oryza sativa* (rice root). A 5 mL starter culture supplemented with ampicillin was grown overnight. The next morning, it was transferred to a 50 mL LB medium, and the culture was grown at 30°C and 280 rpm. When 0.6 OD_{600} was reached, protein expression was induced with 0.5 mM IPTG. After overnight expression, the cells were harvested by centrifugation at 3,000 g for 15 minutes at 4ºC.

The cell pellet was suspended in a lysis buffer (Table 2.26) with 1:4 ratio and was incubated on a rotator for 30 minutes in RT. Later, the lysate was cleared by centrifugation at 20,000 g for 30 minutes, and the supernatant was diluted in wash buffer with 1:5 ratio. The pH of the lysate was then adjusted to pH 5.5 with 1M HCl. The resulting cloudy suspension was cleared by centrifugation at 6,500 rpm 30 minutes. The purification of RrFNR was conducted by cation exchange chromatography, using gradient elution method. A 10 mL Pierce™ disposable column (ThermoFisher Cat. No: 29924) was packed with 5 mL CM Sepharose® Fast Flow resin (Sigma-Aldrich CCF100-50ML), and the column was equilibrated by running down 4 CV of wash buffer. The full purification protocol can be seen in the Table 2.27.

Table 2.26 Buffers used with CM Sepharose resins for RrFNR purification.

Lysis Buffer	1X Bugbuster 1 Protease Inhibitor Tablet 100 mM HEPES/KOH pH 8.1 (pH 8.0), 25 U/mL Benzonase 1 KU/mL lysozyme 0.1% Tween 20
Wash Buffer	25mM Bis-Tris pH 5.5
Elution Buffer	EB1: 150 mM NaCl in 25 mM Bis-Tris pH 5.5 EB2: 250 mM NaCl in 25 mM Bis-Tris pH 5.5

	EB3: 350 mM NaCl in 25 mM Bis-Tris pH 5.5
Exchange Buffer	25 mM Tris, pH 8.0

Table 2.27 Purification procedure of RrFNR by cation exchange chromatography.

Step	Buffer	Volume
Equilibration	Wash Buffer	4 CV
Sample Application		
Column wash	Wash Buffer	5 x 1 CV
Elution	EB1	1 CV
Elution	EB2	1 CV
Elution	EB3	1.6 CV
Regeneration	1 M NaCl	3 CV
Column wash	Wash Buffer	3 CV

The elution fractions resulting from EB2 and EB3 runs (EF2 & EF3 respectively) were pooled together and immediately loaded to the Amicon® Stirred Cell concentrator (Merck UFSC05001) containing an Ultracel® 5 kDa Ultrafiltration Disc (MERCK Millipore Cat. No: PLGC04310) The concentrator was connected to a N_2 tube to apply pressure. The volume was first reduced to one-third, and two 5X buffer exchanges were done with the Exchange Buffer (Table 2.26). Using 0.5 mL Pierce Protein concentrator (PES, 10K MWCO, Thermo Scientific Cat. No: 88513), the volume was reduced to one-fifth. 5% glycerol was added to the final sample, and it was stored at -80ºC. An SDS-PAGE analysis and Bradford assay were conducted later.

Alternatively, the protein was also produced in large-scale with 300mL TB medium. The obtained lysate was diluted in wash buffer 1:1.5 ratio and 10 mL CM Sepharose® Fast Flow resin (Sigma-Aldrich CCF100-50ML) was packed into a 20 mL disposable column. After pH adjustment with 1 M HCl, the lysate was centrifuged 2 times, once at 6,500 rpm for 30 minutes followed by 20,000 g for 15 minutes. The procedure in Table 2.22 was followed for purification, with the modification of EB2 volume being 1.5 CV instead of 1 CV. The purified protein was not concentrated with Pierce Protein concentrator, instead it was stored at -80ºC after addition of 5% glycerol.

After an SDS-PAGE analysis of the purification process, the pure sample was purified once more to get rid of the impurities. Same as before, 10 mL CM Sepharose® Fast Flow resin (Sigma-Aldrich CCF100-50ML) was used in purification. The sample was diluted in 1:5 ratio with the wash buffer, pH was adjusted to 5.5 with 1 M HCl, and centrifuged at 6,500 rpm 30 minutes. 5 elution buffers were prepared with NaCl concentration varying from 150 mM to 350 mM, with 50 mM increments. The purification protocol can be seen in Table 2.28.

Table 2.28 Second-round purification procedure of RrFNR by cation exchange chromatography.

Step	Buffer	Volume
Equilibration	Wash Buffer	4 CV
Sample Application		
Column wash	Wash Buffer	5 x 1 CV
Elution	EB1	1 CV
Elution	200mM NaCl in 25mM Bis-Tris pH 5.5	1 CV
Elution	EB2	1 CV
Elution	300mM NaCl in 25mM Bis-Tris pH 5.5	1 CV
Elution	EB3	1 CV
Regeneration	1M NaCl	3 CV
Column wash	Wash Buffer	3 CV

2.3.3 ΔiscR cell extract preparation

BL21(DE3) ΔiscR cells were grown in 1 L LB medium supplemented with 100 mM MOPS, 1 mM ferric ammonium citrate, and 40 μg/ml kanamycin at 37ºC. When OD_{600} reached 0.6, 1 mM cysteine was added to the media. After 3 hours, the culture was harvested. After washing the pellet with S30 buffer once, the suspension was centrifuged at 6,000g and 4ºC for 15 minutes. The resulting pellet was dissolved in S30 with 1mL S30 per gram of cell ratio. The cell suspension was passed once through the Emulsiflex C-5 homogenizer between 17,500 and 25,000 psi. The lysate was centrifuged for 30 mins at 30,000 g and 4ºC, and the centrifugation was repeated for the supernatant. The cleared

lysate was then wrapped in aluminum foil and incubated 80 minutes at 37ºC with gentle shaking at 100 rpm. After one more centrifugation at the previous conditions, the supernatant was aliquoted, flash frozen in liquid nitrogen and stored at -80ºC.

2.3.4 Maturase lysate preparation for in vitro CpI production

E. coli BL21 (DE3) ΔiscR pACYC HydE HydF HydG strain was used to prepare cell extracts that will be used in CFPS CpI hydrogenase production. Initially, a small-scale production was done to see if the maturases were present after the induction. A 5 mL starter culture in LB medium supplemented with kanamycin and chloramphenicol was grown overnight at 37ºC. The next day, the culture was transferred to a 50 mL LB medium supplemented with the mentioned antibiotics, 100 mM MOPS/KOH pH 7.8, 0.5% w/v glucose, and 2 mM ferric ammonium citrate. After growing the cells at 28ºC until OD_{600} ~0.5-0.6, the falcon tube that the culture was growing in was placed on a rotator inside an anaerobic glovebox. 10 mM sodium fumarate and 2 mM L-cysteine were added to the culture. After 15 minutes of stirring, enzyme production was induced with 0.5 mM IPTG. The cells were left for overnight induction. The following morning, the culture was harvested by centrifugation at 6,000 rpm for 30 minutes. The pellet was resuspended in BugBuster master mix with a 1:3 ratio that contains 1X BugBuster Mix, 12.5 U/ml benzonase, 0.25 mg/ml chicken egg lysozyme and, 100 mM HEPES pH 8.0. After 30 minutes of mixing on the rotator at RT anaerobically, the lysate was centrifuged at 18,000 g for 30 minutes. The supernatant was aliquoted and stored at -80ºC to be used in CFPS experiments. An SDS-PAGE analysis was performed.

The strain was also grown in 1 L of LB media. For this, everything was done exactly as described above until OD_{600} of the 50 mL culture reached 0.5-0.6. Then 1 L of LB media with the supplementations was inoculated. The cells were grown at 28ºC until OD_{600} ~0.5-0.7, then the culture was transferred into an anaerobic glove box and a stir bar was added to the culture flask. While mixing on a stir plate, 10 mM sodium fumarate and 2 mM L-cysteine were added to the culture. After 15 minutes of stirring, enzyme production was induced with 0.5 mM IPTG and left for overnight protein production. The following day, the culture was centrifuged at 6,000 rpm for 30 minutes. The pellet was transferred into a 50 mL falcon tube in the anaerobic glovebox and lysis was proceeded as explained above.

Alternatively, the maturases were also grown separately. For this purpose, *E. coli* BL21 (DE3) ΔiscR pACYC HydF, *E. coli* BL21 (DE3) ΔiscR pACYC HydG and *E. coli* BL21 (DE3) ΔiscR pACYC HydE cell lines were grown in separate 5 mL overnight starter cultures. The following day, 50 mL LB media containing kanamycin, chloramphenicol, 100 mM MOPS/KOH pH 7.8, 0.5% w/v glucose, and 2 mM ferric ammonium citrate were inoculated with the overnight cultures. They were grown at 28ºC until OD_{600} reached ~0.55 and then cycled inside the anaerobic glovebox. 10 mM sodium fumarate and 2 mM L-cysteine were added to the cultures and after 15 minutes on rotator, protein production was induced with 0.5 mM IPTG. The maturase lysates were prepared with the same procedure mentioned above.

2.3.5 Cell-free production of wild-type CpI hydrogenase

To produce wild type CpI hydrogenase *in vitro*, the procedure in Section 2.1.3 was followed. Aside from the reagents listed in Table 2.14, 2 mM S-adenosyl methionine (SAM), 0.8 mM ammonium iron (II) sulfate, 0.1 mg/mL 31R1 (an anti-foaming agent), and 0.8 mM sodium sulfide were added to the reactions. The maturase lysates mentioned in the previous section were used as the cell extract. The reaction mixtures were incubated in a nitrogen-glovebox at room temperature for 16 hours. Wild-type sfGFP was produced to provide a positive control to the CFPS trials.

2.3.6 Gas Chromatography to detect hydrogen production activity of CpI Hydrogenase

Drawing calibration curves for hydrogen and oxygen

Before plotting calibration curves, YL6500 Gas Chromatography settings were adjusted. To have the peaks of hydrogen and oxygen completely separated, column temperature, oven temperature and carrier gas flow rate were changed. The oven temperature was decreased to 25ºC, the column temperature to 70ºC and flow rate to 3 mL/min.

Varying percentages of H_2 samples were prepared inside 2 mL glass vials. The gas samples were taken from 2 different anaerobic glovebox, one filled with N_2 gas only and one with a gas mix of 98.1% N_2 and 1.9% H_2. The details about the dilution of H_2 can be seen in the Table 2.29. Similarly, an oxygen calibration curve was drawn with

samples from with N_2 gas only anaerobic glovebox and air directly, with the assumption that it consists of 79% N_2 and 21% O_2 (Table 2.30). The measurements for the oxygen calibration curve were made by Kim Seongwon in Dr. Jamin Koo Lab, Hongik University.

Table 2.29 H_2 dilutions to draw a calibration curve with gas chromatography.

Vial	N2 gas volume (mL)	Gas mix volume (mL)	Total volume (mL)	H_2 percentage
1	0	2	2	1.9
2	0.2	2	2.2	1.73
3	0.5	2	2.5	1.52
4	2	0.5	2.5	0.38
5	2	0.2	2.2	0.17
6	2	0	2	0

Table 2.30 O_2 dilutions to draw a calibration curve with gas chromatography.

Vial	N2 gas volume (mL)	Air volume (mL)	Total volume (mL)	O2 percentage
1	0	2	2	21
2	0.5	2	2.5	16.8
3	2	0.47	2.47	4
4	2	0.21	2.21	2
5	2	0.1	2.1	1
6	2	0.05	2.05	0.5

YL6500 Gas Chromatography was turned on and YL-Clarity software from the connected computer were opened. From the ‘GC’ tab at method editor the oven temperature was set to 25ºC. The packed column was used in the GC trials for hydrogen and oxygen presence determination and the temperature of the column was adjusted to be 70ºC. The carrier gas was set to be N_2 with the flow rate of 3 mL/min. TCD was used as the detector, and the temperature was set to 120ºC. From the ‘Measurement’ tab the run time of the reading was set. Once the system reached the set temperatures and flow rate, filament of TCD was turned on. 50 µL samples from each vial was taken with a GC syringe and injected to the GC system through the septa. The measurements were taken 3 times to minimize human error. From the automatically measured areas of the observed

peaks, a calibration curve that shows the change in the area (mV*s) with varying moles of the gas of interest was drawn.

NADPH and DTH assays to measure hydrogen production activity of in vivo produced CpI hydrogenase

For the NADPH-driven assay, reagents with concentrations given in Table 2.31 were added inside a 2 mL GC vial in an anaerobic glovebox. For DTH assay, the reagents given in Table 2.32 were added to the GC vial. To another vial 200 µL dH_2O was added to serve as the blank sample. A magnetic stir-bar was added to each reaction vial inside the N_2 only glovebox and the caps of the vials were crimped, after that they were taken outside for GC measurements. The vials were placed on top of a magnetic stirrer and the speed was set to 240 rpm. A 50 µL sample was taken every 30 minutes until 1 hour is completed and injected to GC separately. The resulting peak areas and retention times were noted. From the calibration curved previously drawn, the produced hydrogen amount and leaked oxygen amount were determined.s

Table 2.31 NADPH activity assay reagents and required concentrations to measure hydrogen productivity of CpI Hydrogenase.

Chemical	Stock Concentration	Final Concentration
Tris-HCl, pH 7	100 mM	50 mM
NADPH	100 mM	10 mM
SynFd	597.8 µM	5 µM
RrFNR	84.04 µM	5 µM
CpI	37.76 nM	100 nM

For the DTH assay, 50 mM Tris-HCl (pH 7.0), 5 µM SynFd, 100 nM CpI hydrogenase, and a magnetic stir-bar were added to the GC vial. To another vial 200µL dH_2O was added to serve as the blank sample. 10 mM DTH was added to the assay vial just before taking it out of the N_2 only glovebox and the vials were crimped immediately. After taking the vials out of the glovebox, they were placed on top of a magnetic stirrer and the speed was set to 240 rpm. 50 µL gas was sampled every 30 minutes until 1 hour is completed. The hydrogen and oxygen amounts in the vials were calculated from the calibration curve.

Table 2.32 DTH activity assay reagents and required concentrations to measure hydrogen productivity of CpI Hydrogenase.

Chemical	Stock Concentration	Final Concentration
Tris-HCl, pH 7	100 mM	50 mM
DTH	500 mM	10 mM
SynFd	597.8 μM	5 μM
CpI	37.76 nM	100 nM

Chapter 3:
RESULTS

3.1 Preparations for wild-type and mutant CpI Hydrogenase Production in CFPS

3.1.1 Production and Purification of o-tRNAs required for CFPS ncAA incorporation

To incorporate ncAAs into a protein of interest carrying amber stop codon, two things are required. First is an orthogonal tRNA and the second is an aminoacyl tRNA-synthetase. To obtain the o-tRNA that will be used with M. jannaschii derived synthetases, the plasmid containing the tRNA sequence (pY71-minHH opt. o-tRNA[46] (#30) from lab stocks) was amplified using the primers in Table 2.1. The o-tRNA that is inside the vector contains a T7 promoter, a ribozyme, and the tRNA sequence. Following the large-scale PCR reaction given in Table 2.2, ethanol precipitation was done to purify the PCR product. Figure 3.1 proves that the *M. jannaschii* o-tRNA was indeed produced and purified as the band observed in the agarose gel is at the correct place (218 bp).

M. barkeri derived Pyl o-tRNA was produced in a different manner, as there was no plasmid in the laboratory stocks containing the sequences of interests. A modified tRNA sequence was found in a literature search and the experiments were continued with that sequence[48]. First, the promoter strength of the o-tRNA sequence was calculated based on the first 6 nucleotides[49].

Table 3.1 The promoter strength of $tRNA^{Pyl\text{-}opt}$.

Position	The natural base	The bases in $tRNA^{Pyl\text{-}opt}$	Relative promoter strength
+1	G	G	1
+2	G	G	1
+3	G	A	0.88
+4	A	A	1
+5	G	A	0.86
+6	A	C	0.97
		Calculated promoter strength: = 1 x 1 x 0.88 x 1 x 0.86 x 0.97 = 0.73	

Even though the promoter strength is relatively high, two o-tRNAs, one with a hammerhead ribozyme sequence and one without it were designed. The ribozyme helps the gene of interest to be produced effectively, where the promoter strength is low[47]. Two separate PCR assembly reactions were done to obtain the two designs. Then a large-scale amplification PCR was done with the end primers. The ethanol precipitated PCR products are given in Figure 3.1. $tRNA^{Pyl\text{-}opt}$ is 95 bp and $tRNA^{Pyl\text{-}opt\text{-}ribo}$ is 151 bp long. The agarose gel image shows that the end products are at the right sizes.

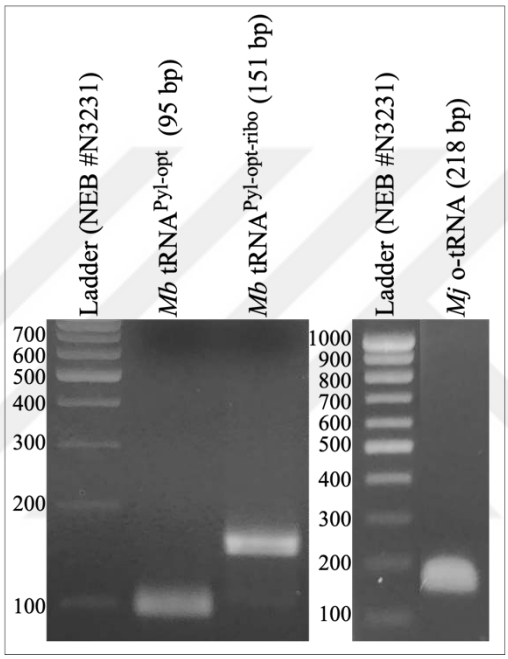


Figure 3.1 Agarose gel analysis of o-tRNA templates produced by PCR, required for CFPS ncAA incorporation.

3.1.2 Cloning of aminoacyl tRNA synthetase genes into pY71 vector

The genes for the aminoacyl tRNA synthetases were planned to be cloned into the pY71 vector as it is a high-copy number plasmid, small in size, and it is designed for CFPS specifically. Before starting the cloning experiments, primers given in Table 2.5 were used to insert NdeI and SalI restriction sites to the N-terminal and C-terminal and of each gene. Also, to the C-terminal of the genes, a 6xHis-tag affinity marker sequence was added. After a successful PCR, the purified PCR product and a pY71 sfGFP plasmid (that contains NdeI and SalI restriction sites) were double digested using NdeI and SalI enzymes initially for 2 hours. However, it was noticed that a large amount of the pY71 vector stayed intact after this step. Therefore, the restriction time was extended to

overnight, while keeping the temperature at 37ºC. Gel purification was used to extract the genes with the right sizes directly from the gels. Unfortunately, this purification method led to very low purification yields that are insufficient to provide 1:5 vector:insert ratio. Moreover, the UV exposure during cutting the DNA fragment of interest could damage the DNA, thus this method was abandoned. Instead, PCR purification was used to purify the DNA fragments.

Ligation reaction of the pY71 vector and gene of interests were run for 2 hours initially. However, this resulted to no colonies on the transformation plate. Therefore, the run time for the reactions were extended to overnight, while keeping the temperature unchanged at RT. It was proven previously that SOC medium increases transformation efficiency as it helps the recovery of the bacteria[50]. Therefore, this media was used in the transformation of the ligation products. Still, after the check-digest, still sfGFP gene was observed instead of the aaRS genes. This time the source plasmid was changed from pY71 sfGFP to a pY71Anakinra vector found in the lab stocks. After following the procedure with the mentioned optimizations, the genes were successfully cloned into the pY71 vector, as can be seen in Figure 3.2. As all the aaRS genes were derived from the same organism, *M. jannaschii*, they had the same size and very similar sequences.

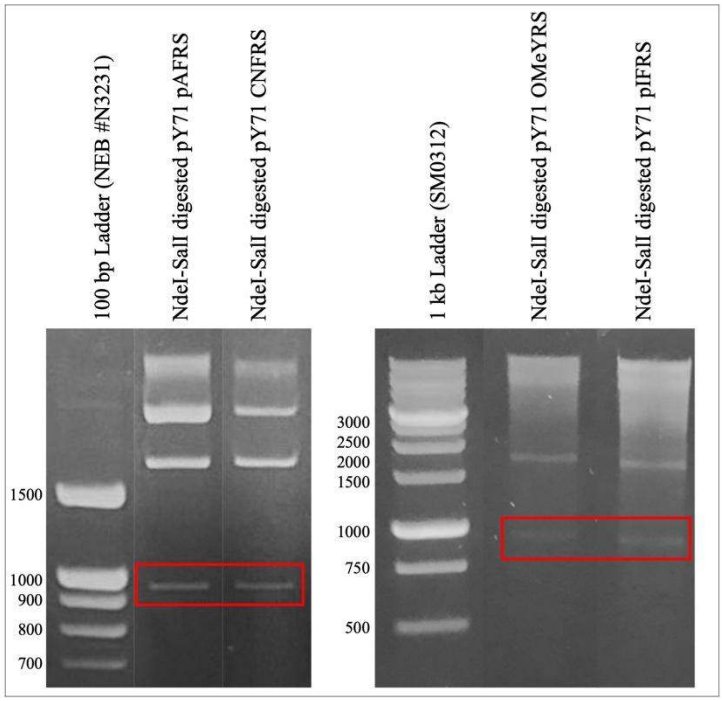


Figure 3.2 Agarose gel analysis of check digests from the cloned aaRS plasmids. pAFRS, CNFRS, OMeYRS, pIFRS are each 939bp long.

After sequence confirmation of the restriction-ligation process, it was noticed that the aaRS genes contained two methionine codons back-to-back at the beginning of the sequence due to a design mistake with the primers. Thus, an SDM was performed to remove these extra 3 bases with the primers in Table 2.9 and protocol in Table 2.10. Figure 3.3 shows the sequence confirmation of the SDMs, as there is only one methionine sequence at the beginning.

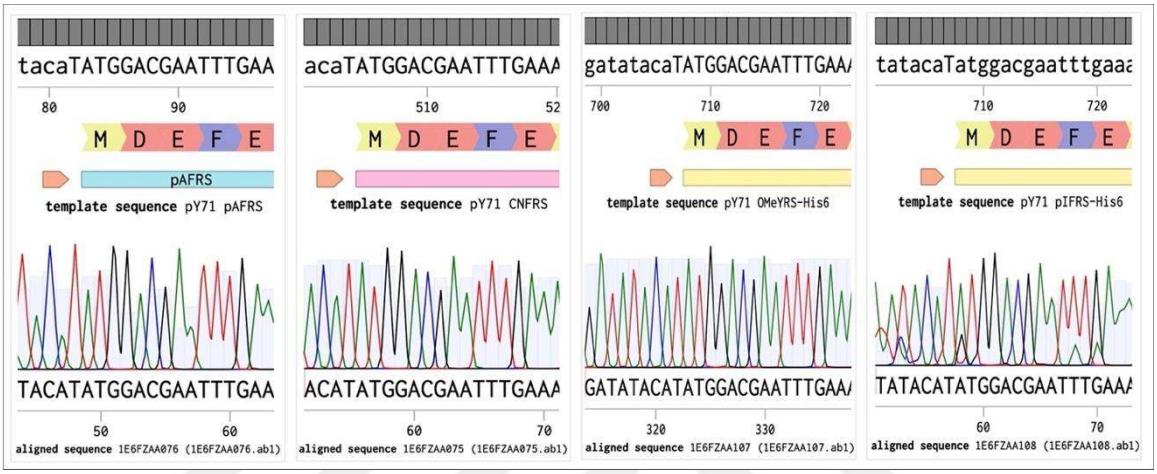


Figure 3.3 Sequence confirmation of the resulting plasmids after SDMs to remove second methionine at translation start.

3.1.3 Production and purification of M. jannaschii derived aaRSs

After successfully cloning the aaRS genes to the pY71 vector, it was proceeded with the production of these proteins *in vivo*. *M. jannaschii* derived synthetases, except for pAMFRS, were produced and purified with the exact same protocol. First the plasmids containing the gene of interest was transformed to *E.coli* BL21(DE3) pLysS cell line, which is optimized for protein production. The cultures were grown in 1L 2YT medium and induced around 0.6 OD_{600} with 1 mM IPTG. After 3 hours of induction, the cultures were harvested.

To produce pAFRS synthetase, *E. coli* BL21(DE3) pLysS pY71 pAFRS was induced with 1 mM IPTG at OD_{600} 0.78, when the culture was in its log phase (Figure 3.4). The cells were harvested after 3 hours at OD_{600} 2.284, and a total of 5.31 g cell pellet was obtained from 1 L culture.

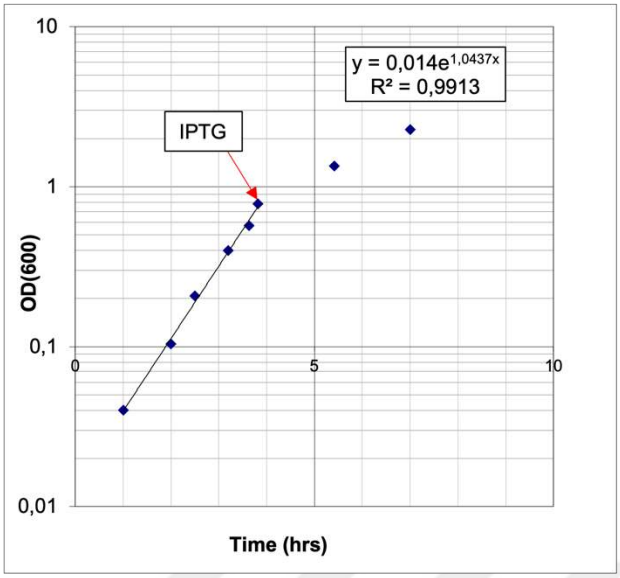


Figure 3.4 Growth curve of *E. coli* BL21(DE3) pLysS pY71 pAFRS. Specific growth rate (μ): 1.04 hr^{-1}. Red arrow: Induction point with IPTG.

All aaRS genes derived from *M. jannaschii* contain a 6xHis-tag sequence at the C-terminus. Therefore, they were purified using a HisTrap column with fast protein liquid chromatography (FPLC). *E. coli* BL21(DE3) pLysS pY71 pAFRS cell pellet was first resuspended with the binding buffer (Table 3.7) containing a protease inhibitor tablet with a 5 mL buffer per gram of cell ratio. The suspension was lysed using Emulsiflex C-5 homogenizer. After centrifuging the lysate to clear it from cell debris and insoluble protein aggregates, the supernatant was filtered. The FPLC chromatogram of pAFRS purification, showing a large peak at the elution can be seen in Figure 3.5A. After the purification, an SDS-PAGE gel was run to analyze the process. The gel image (Figure 3.5b) shows that there was a small amount of protein accumulated in the pellet, however, most of the pAFRS remained soluble after the centrifugation. Most of the protein bound to the column after the sample application and stayed bound even after the column wash. The elution fractions display large bands containing the protein, however there are impurities as well. Both elution fractions were dialyzed against a 10 mM potassium phosphate buffer (pH 7.4).

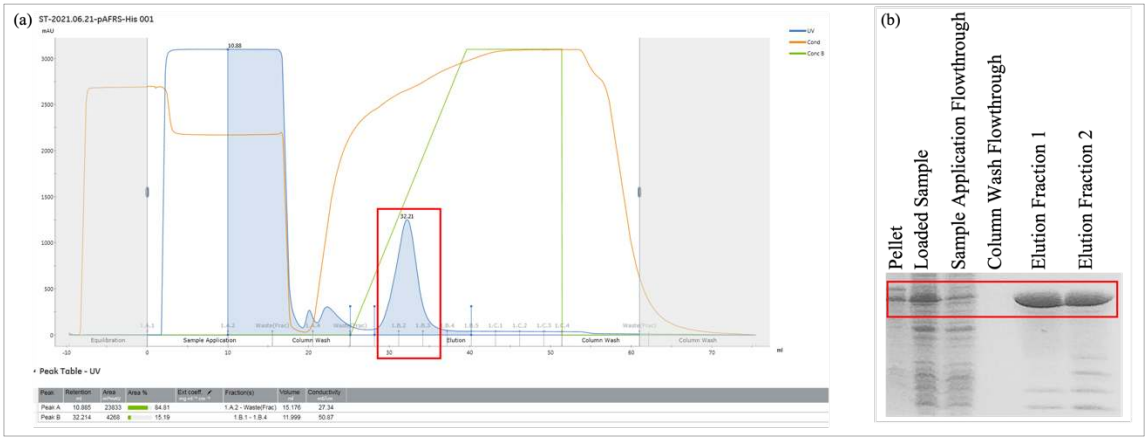


Figure 3.5 Purification of pAFRS. (a) FPLC Chromatogram obtained after His-tag affinity chromatography. Red box: Elution peak of pAFRS. Blue line: UV adsorption at 280 nm for protein concentration determination. Orange line: Conductivity for salt concentration measurement. Green line: Concentration of the elution buffer. (b) SDS-PAGE analysis of the purification (pAFRS: 35.7 kDa). Red box: expected elution area on the gel.

The growth curve for CNFRS production in 2YT medium can be seen in Figure 3.6. the culture was induced at OD_{600} 0.604 with 1 mM IPTG. After 3 hours of protein production the cells were harvested at OD_{600} 2.288, and 5.37 g cell pellet was obtained.

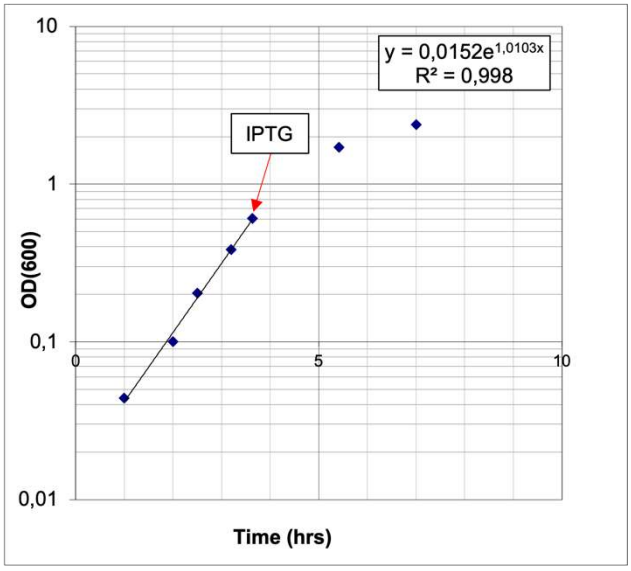


Figure 3.6 Growth curve of *E. coli* BL21(DE3) pLysS pY71 CNFRS. Specific growth rate (μ): 1.01 hr^{-1}. Red arrow: Induction point with IPTG.

Figure 3.7a shows the FPLC chromatogram of CNFRS purification. There were 2 large peaks observed during elution. When an SDS-PAGE gel was run (Figure 3.7b), it

was seen that the elution fractions from both the peaks contain pure CNFRS protein. All elution fractions were dialyzed separately. The pellet contained the protein of interest however, compared to the load, it can be concluded that most of the protein stayed soluble. There was unbound CNFRS after loading the sample to the column.

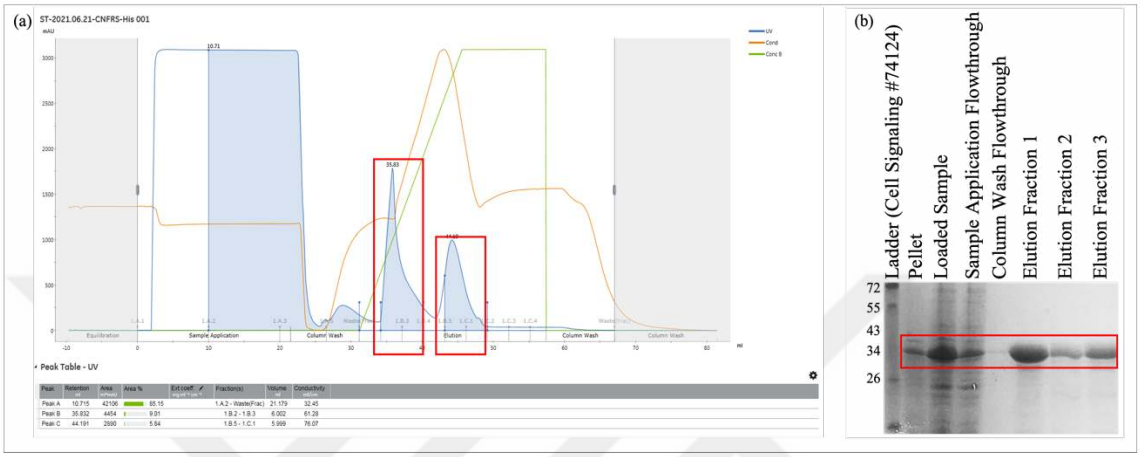


Figure 3.7 Purification of CNFRS. (a) FPLC Chromatogram obtained after His-tag affinity chromatography. Red box: Elution peak of CNFRS. Blue line: UV adsorption at 280 nm for protein concentration determination. Orange line: Conductivity for salt concentration measurement. Green line: Concentration of the elution buffer. (b) SDS-PAGE analysis of the purification (CNFRS: 35.8 kDa). Red box: expected elution area on the gel.

E. coli BL21(DE3) pLysS pY71 OMeYRS culture was grown in 2YT medium until OD_{600} 0.708 and then induced with 1 mM IPTG. After 3 hours induction, the cells were harvested at OD_{600} 3.116, and the cell pellet weighted 6.65 g. Figure 3.8 shows the growth curve of the culture.

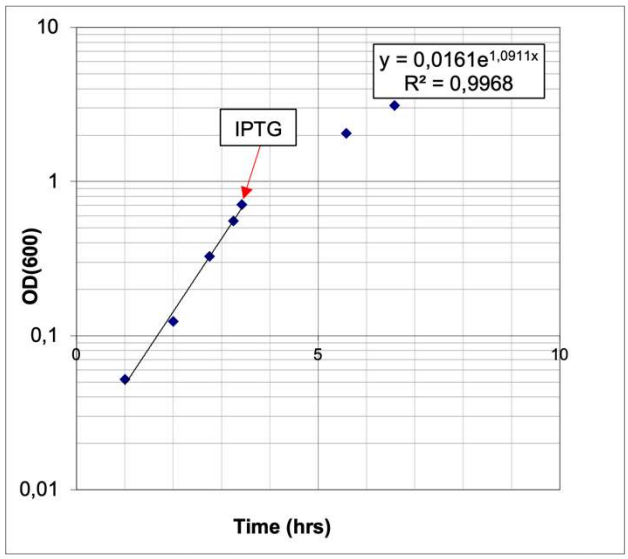


Figure 3.8 Growth curve of *E. coli* BL21(DE3) pLysS pY71 OMeYRS. Specific growth rate (μ): 1.09 hr^{-1}. Red arrow: Induction point with IPTG.

The FPLC chromatogram of OMeYRS shows a high peak (Figure 3.9a). The fractions in the peak were run on an SDS-PAGE gel (Figure 3.9b) and OMeYRS band which is around 35 kDa, was observed in each peak. Thus, they were dialyzed separately. Sample application and wash flowthroughs contain a small amount of protein as well but the major part is in elution.

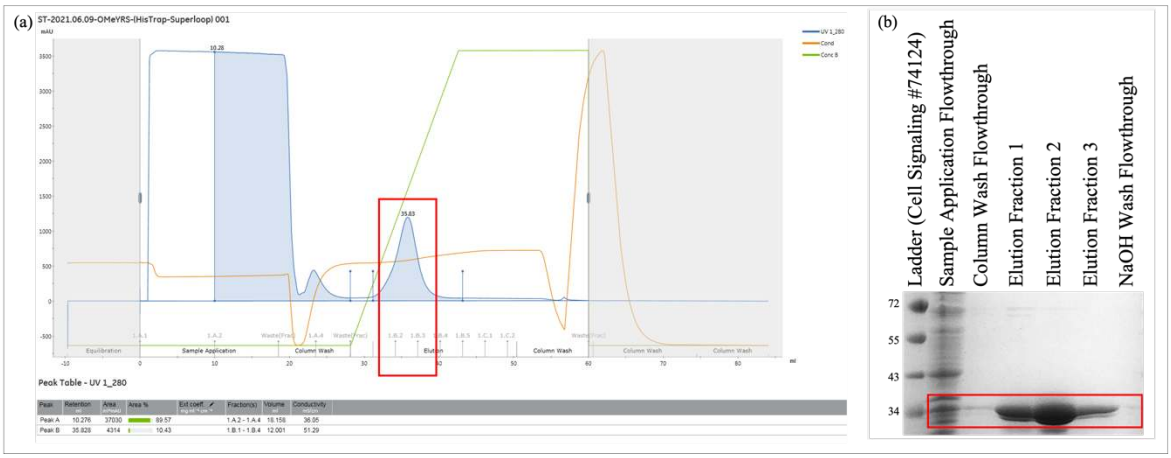


Figure 3.9 Purification of OMeYRS. (a) FPLC Chromatogram obtained after His-tag affinity chromatography. Red box: Elution peak of OMeYRS. Blue line: UV adsorption at 280 nm for protein concentration determination. Orange line: Conductivity for salt concentration measurement. Green line: Concentration of the elution buffer. (b) SDS-PAGE analysis of the purification (OMeYRS: 35.7 kDa). Red box: expected elution area on the gel.

Figure 3.10 shows the growth curve for pIFRS production. The protein production was induced at OD_{600} 0.6132 with 1 mM IPTG. The culture was then harvested after 3 hours at OD_{600} 2.896 and resulted to 5.86 g cell pellet.

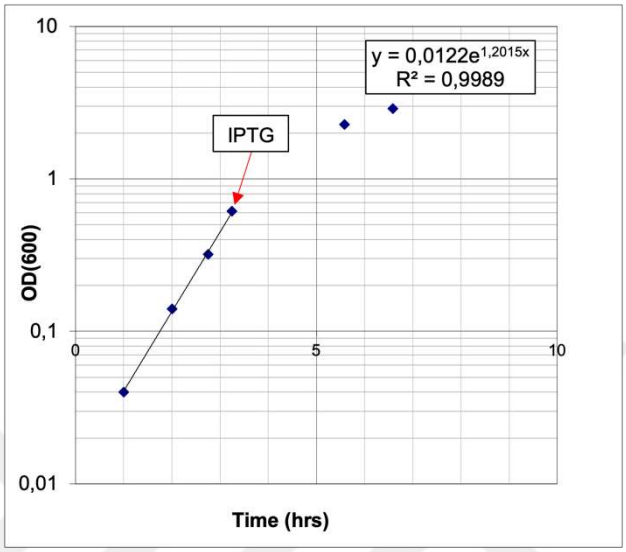


Figure 3.10 Growth curve of *E. coli* BL21(DE3) pLysS pY71 pIFRS. Specific growth rate (μ): 1.20 hr^{-1}. Red arrow: Induction point with IPTG.

On the FPLC chromatogram given below for pIFRS (Figure 3.11a), there are two peaks in a merged state at elution. When an SDS-PAGE gel was run, shown in Figure 3.11b, the protein of interest is observed in both sample application and column wash flowthroughs. The first two elution fractions are from the peak 1 and the last two are from peak 2. pIFRS is present in each elution fractions, but it is low in amount in fraction 1. Therefore, all elution fractions except EF1 were dialyzed separately to later test which fraction works best in incorporating pIF to sfGFP.

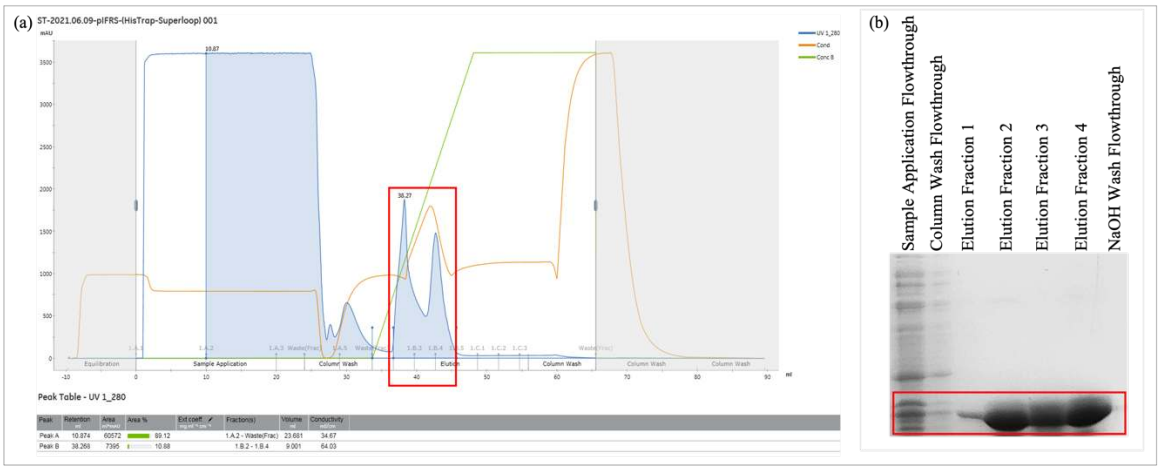


Figure 3.11 Purification of pIFRS. (a) FPLC Chromatogram obtained after His-tag affinity chromatography. Red box: Elution peak of pIFRS. Blue line: UV adsorption at 280 nm for protein concentration determination. Orange line: Conductivity for salt concentration measurement. Green line: Concentration of the elution buffer. (b) SDS-PAGE analysis of the purification (pIFRS: 35.8 kDa). Red box: expected elution area on the gel.

The plasmid carrying pPaFRS gene, pY71 pPaFRS, was transformed to *E. coli* BL21(DE3) pLysS cell line. Even though the culture was grown in 2YT medium at 37°C, like the rest of the *M. jannaschii* synthetases, the observed growth rate (h^{-1}) was much lower (Figure 3.12). Still, the protein production was induced when OD_{600} 0.592 was reached. The culture was harvested after 3 hours, when OD_{600} was 1.036.

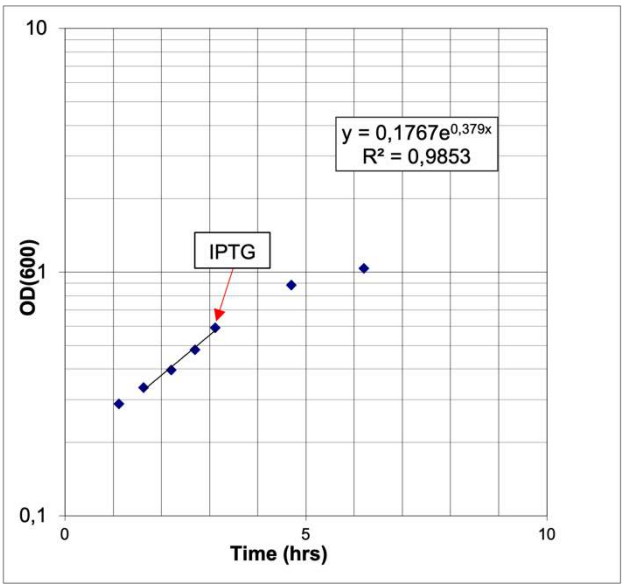


Figure 3.12 Growth curve of *E. coli* BL21(DE3) pLysS pY71 pPaFRS. Specific growth rate (μ): 0.38 hr^{-1}. Red arrow: Induction point with IPTG.

The FPLC chromatogram of pPaFRS is given in Figure 3.13a. There is a peak in the elution section, indicating pure pPaFRS. The SDS-PAGE (Figure 3.13b) for the purification suggests that almost all the loaded protein bound to the column and eluted with increasing imidazole concentration. Elution fraction 2 has more protein than from fraction 1. Still, both fractions were carried on to dialysis.

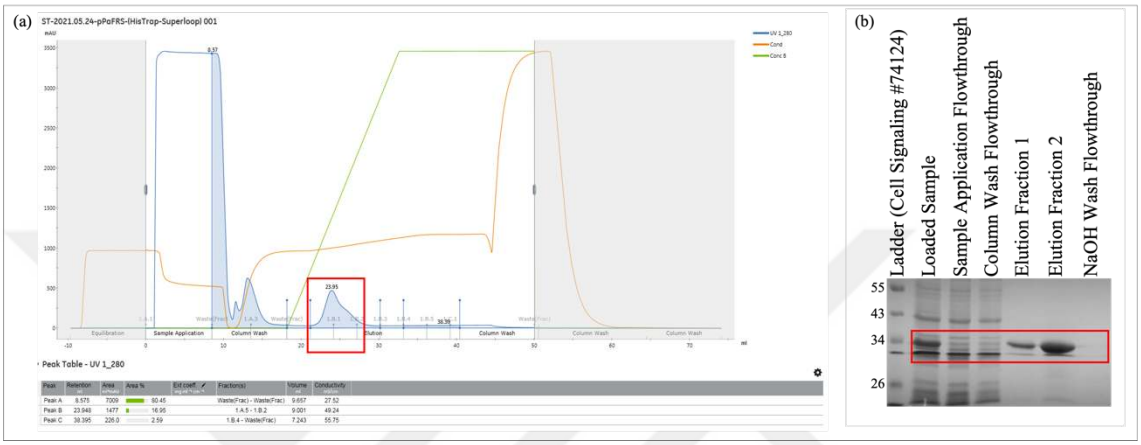


Figure 3.13 Purification of pPaFRS. (a) FPLC Chromatogram obtained after His-tag affinity chromatography. Red box: Elution peak of pPaFRS. Blue line: UV adsorption at 280 nm for protein concentration determination. Orange line: Conductivity for salt concentration measurement. Green line: Concentration of the elution buffer. (b) SDS-PAGE analysis of the purification (pPaFRS: 35.7 kDa). Red box: expected elution area on the gel.

To produce pAMFRS protein, pY71 pAMFRS was transformed to *E. coli* BL21(DE3) pLysS competent cells. The culture was grown in a 2YT medium at 37°C and protein production was induced at 0.61 OD_{600}. After 3 hours of induction, the culture was harvested at 1.68 OD_{600} (Figure 3.14). The cell pellet weighed 3.45 g.

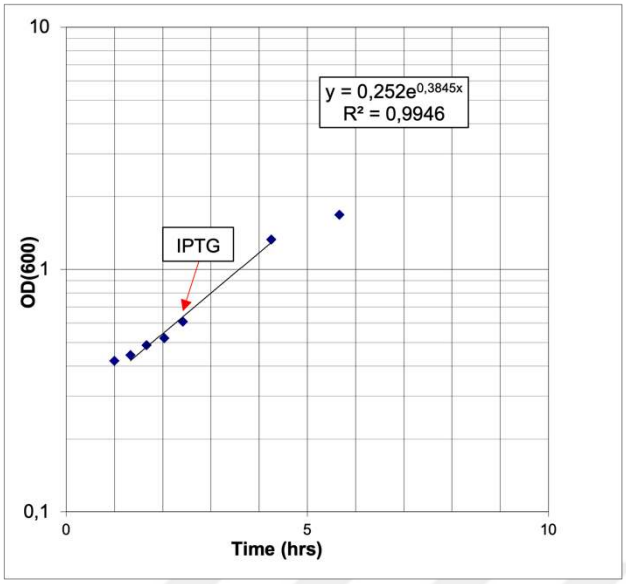


Figure 3.14 Growth curve of *E. coli* BL21(DE3) pLysS pY71 pAMFRS. Specific growth rate (μ): 0.38 hr^{-1}. Red arrow: Induction point with IPTG.

The chromatogram and SDS-PAGE image of pAMFRS purification are given in Figure 3.15. There were 2 peaks in the elution on the chromatogram, and when the elution fractions were run on an SDS-PAGE gel, they all had the protein of interest in the correct size.

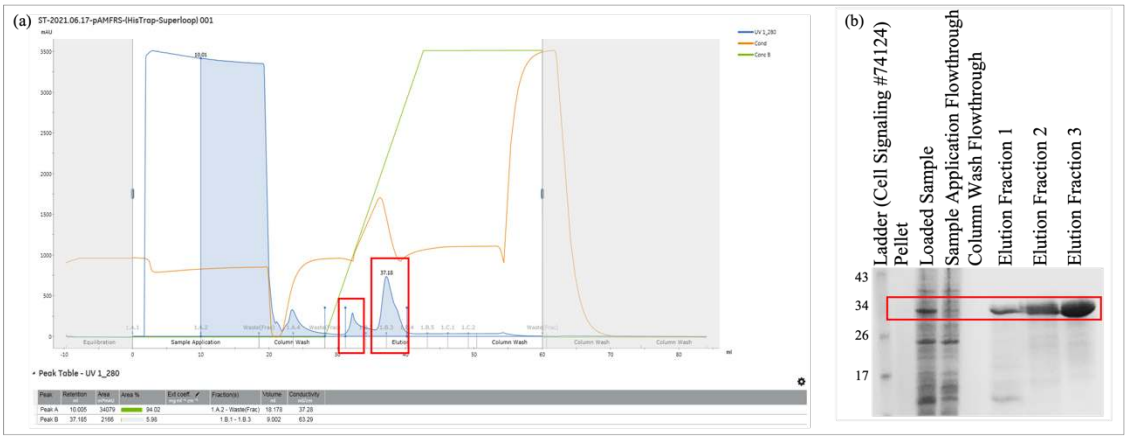


Figure 3.15 Purification of pAMFRS. (a) FPLC Chromatogram obtained after His-tag affinity chromatography. Red box: Elution peak of pAMFRS. Blue line: UV adsorption at 280 nm for protein concentration determination. Orange line: Conductivity for salt concentration measurement. Green line: Concentration of the elution buffer. (b) SDS-PAGE analysis of the purification (pAMFRS: 35.7 kDa). Red box: expected elution area on the gel.

After an overnight dialysis in a potassium phosphate buffer containing 20% sucrose (pH 7.4), the elution fractions were analyzed. For pAMFRS, PBS buffer was used

instead (pH 7.1) However, dialysis fragments had aggregation. There was no aggregation in CNFRS and pAFRS samples except for CNFRS EF3 and a slight aggregation was present in OMeYRS EF3. The most precipitated sample was pIFRS EF3. In the pPaFRS EF1 sample, there was a little aggregate formation as well.

To make sure all fractions were cleared from these aggregates, all dialysis fractions were centrifuged for 5 minutes at 12.000g to remove the aggregates. Then each fraction was run on a gel with several dilutions (Figure 3.16) to check for impurities and lysed protein. Compared to CNFRS EF1 and EF2, EF3 does not contain much protein. EF1 and EF2 fractions had lysed protein, as there were multiple bands under the main 35 kDa bands. EF1 has the most protein out of the 3 fractions. Both pAFRS EF1 and EF2 contain a large amount of lysed protein. But they also have the full-size protein with a high concentration. pAMRS EF1 and EF3 has more protein than EF2. All 3 OMeYRS elution fractions have the intact protein of interest. While pIFRS fractions also have full-size protein, EF2 displays a lysed protein fraction. But it is also more abundant in protein, compared to EF1 and EF3. Lastly, pPaFRS elution fractions show the protein at the right place, and EF2 has a higher concentration of both intact and lysed protein.

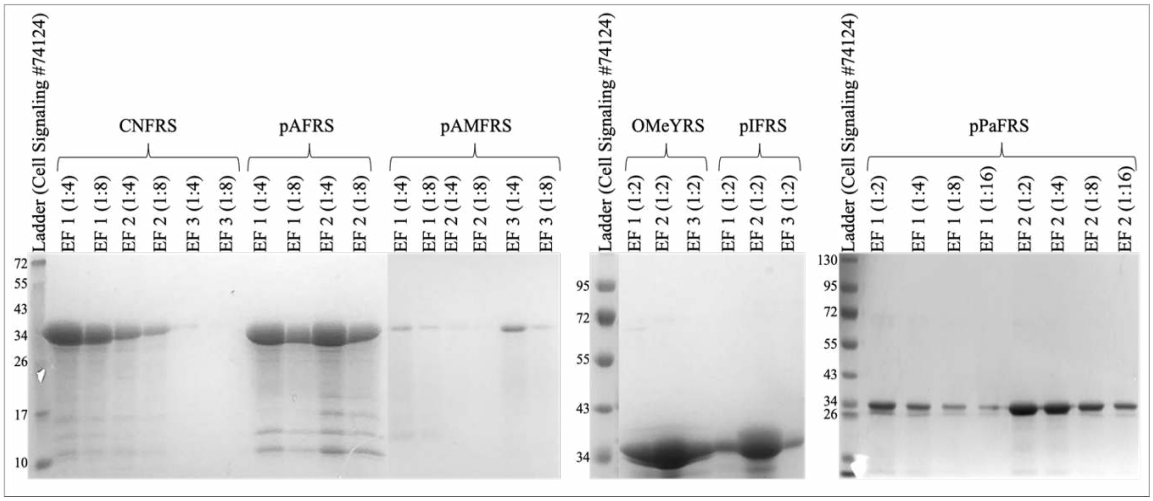


Figure 3.16 SDS-PAGE analysis of purified aaRS samples after dialysis.

DC assay was used to calculate the concentrations of purified and dialyzed aaRS samples (Table 3.2). The concentrations are required as each fraction would be tested in CFPS for their efficiency in ncAA incorporation.

Table 3.2 Protein concentration determination of aaRS elution fractions with DC assay.

Sample	Concentration (mg/mL)	Standard deviation
CNFRS EF1	8.58	0.27
CNFRS EF2	2.00	0.36
pAFRS EF1	1.50	0.18
pAFRS EF2	0.89	0.79
OMeYRS EF1	2.31	0.24
OMeYRS EF2	5.51	0.75
OMeYRS EF3	1.50	0.18
pIFRS EF1	0.89	0.07
pIFRS EF2	4.38	0.72
pIFRS EF3	0.77	0.05
pPaFRS EF1	0.78	0.04
pPaFRS EF2	1.82	0.10
pAMFRS EF1	8.29	0.05
pAMFRS EF2	10.61	0.02

3.1.4 CFPS to insert ncAAs to sfGFP

Mj o-tRNA and the purified synthetases were tested for their efficiency in ncAA incorporation to sfGFP. CNFRS EF3 was not tested in cell-free reactions there was slight to no band in the post-dialysis gel in Figure 3.14. All the synthetases except for CNFRS showed activity as they successfully incorporated their corresponding ncAA to sfGFP151TAG (Figure 3.17). The best working elution fractions were picked to continue the CFPS experiments. Therefore, pPaFRS EF2, OMeYRS EF1, pIFRS EF1 and were chosen for the future ncAA incorporation trials.

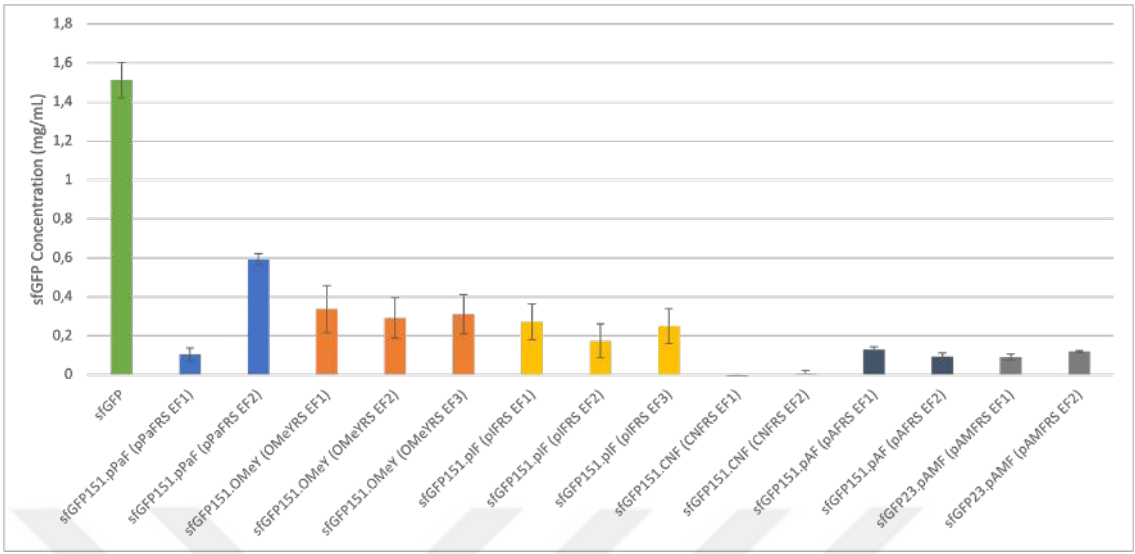


Figure 3.17 Concentrations of soluble sfGFP variants produced in CFPS. Error bars represent the difference in soluble protein concentration between reaction replicates (n=9 for sfGFP, n=3 for variants). EF: Elution Fraction. sfGFPX.Y: **Y** ncAA is inserted at the **X** position of sfGFP.

Since sfGFP151.CNFRS variants did not show any fluorescence activity and sfGFP151.pAF variants had very little, whether these synthetases were inactive or there was another factor for the low fluorescent values were investigated. ncAA incorporation with sfGFP23TAG and sfGFP39TAG plasmids were tested for this purpose. Figure 3.18 shows that CNFRS EF1 showed more efficiency in incorporating into sfGFP23TAG than to sfGFP39TAG and the situation was the opposite for EF2. sfGFP39pAF had a higher concentration compared to sfGFP23TAG. Either way, each synthetase was active.

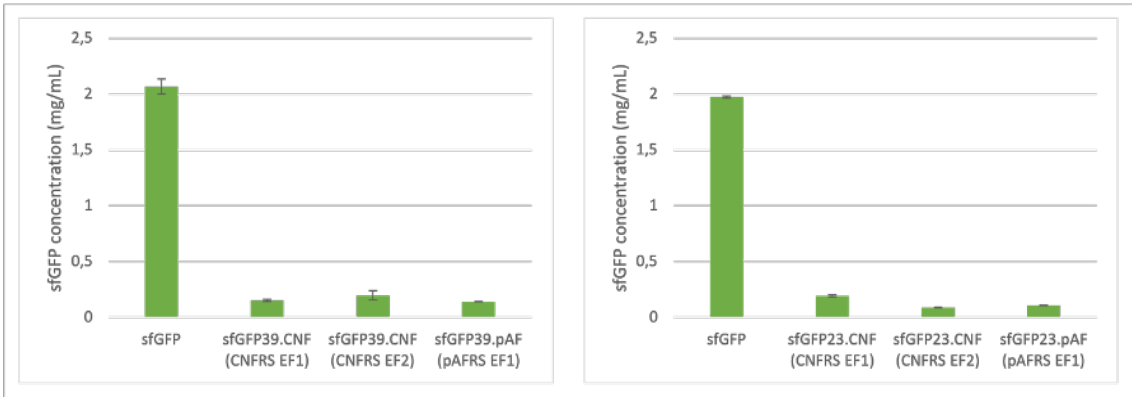


Figure 3.18 Concentrations of soluble sfGFP39TAG and sfGFP23TAG variants produced in CFPS. Error bars represent the difference in soluble protein concentration between reaction replicates (n=X). EF: Elution Fraction. sfGFPX.Y: **Y** ncAA is inserted at the **X** position of sfGFP.

3.1.5 Site-directed mutation of CpI hydrogenase

To incorporate ncAAs into CpI hydrogenase, 4 points on the gene of the protein were chosen. An SDM reaction was performed to mutate these selected sited to amber stop codon (TAG). Figure 3.19 shows the sequence confirmation of the SDM process. The mutation was successful.

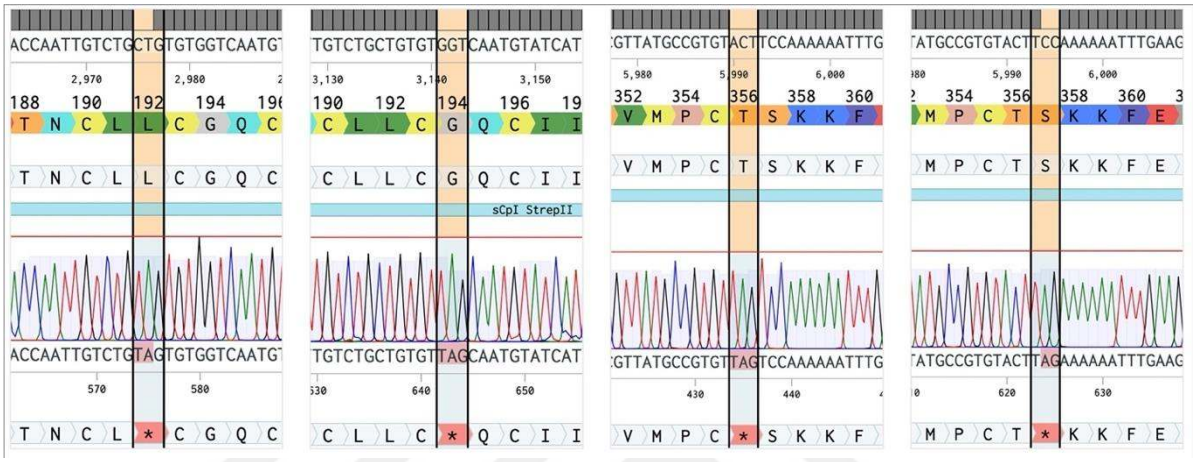


Figure 3.19 Sequence alignments to confirm amber stop codon (TAG) insertions by SDM at L192, G194, T356, and S357 .

3.2 M. barkeri derived aminoacyl tRNA synthetase HRS production

3.2.1 Cloning and SDM of HRS gene

A clonal vector containing HRS gene with 5'-NdeI, 3'-SalI and C-terminal 6xHis-tag was ordered from Twist. The gene of interest was then cloned into pY71 vector with restriction-ligation method. The check-digest result of the cloned plasmid is given in Figure 3.20. The agarose gel image displays two DNA fragments, the upper one around 1700 kDa being pY71 vector and the lower one being HRS with 1278 bp. After this confirmation, a sequence verification was also done.

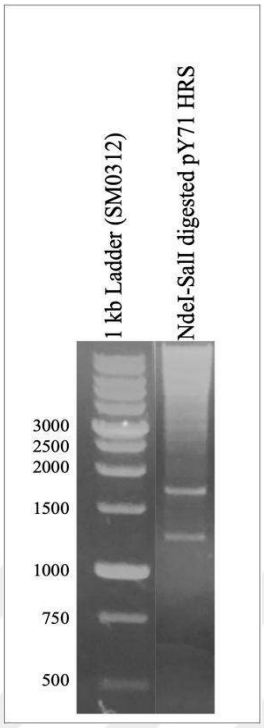


Figure 3.20 Agarose gel analysis of the check digest for the cloned HRS plasmid (1278 bp).

Just like *M. jannaschii* derived synthetases, HRS also had an extra methionine at translation initiation by a design mistake. Therefore, SDM was performed to delete this extra methionine. Figure 3.21 shows the successful sequence confirmation of the process.

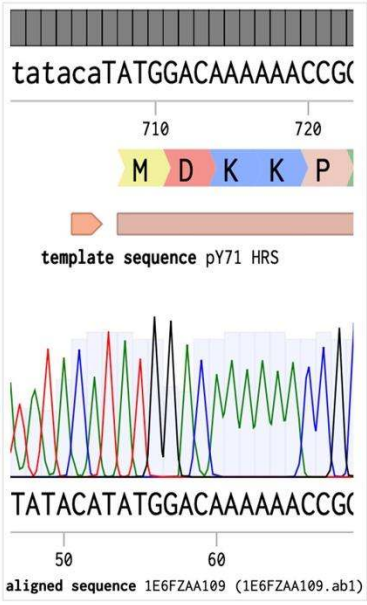


Figure 3.21 Sequence alignment to confirm the removal of the second methionine at translation initiation in in HRS by SDM.

3.2.2 Expression study of HRS with varying temperatures, IPTG concentrations, and additives

Traditional expression and purification HRS

After the final version of the gene inside the pY71 vector was obtained, the plasmid was transformed into *E. coli* BL21 (DE3) pLysS competent cells. The cell line was then grown at 37°C in 1 L 2YT medium until OD_{600} 0.684 was reached. Then it was induced with 1 mM IPTG, and protein production continued for 3 hours. Then the culture was harvested with a final OD_{600} of 3.484 and 4.69 g cell pellet was obtained. After confirming induction with an SDS-PAGE gel, the protein was purified with a His-tag affinity column. However, there was no band at the right place for HRS in the elution fractions. Initially, it was suspected that either HRS was incapable of binding to the column, or the protein was forming aggregates in the pellet. To eliminate any human error in the experiments, the production and purification experiments were repeated. *E. coli* BL21 (DE3) pLysS pY71 HRS culture was grown in 1 L 2YT medium until OD_{600} 0.604 (Figure 3.22) was reached. Protein production was induced with 1 mM IPTG and after 3 hours of induction time, the culture was harvested with a final OD_{600} of 3.452 and 4.47g cell pellet was obtained.

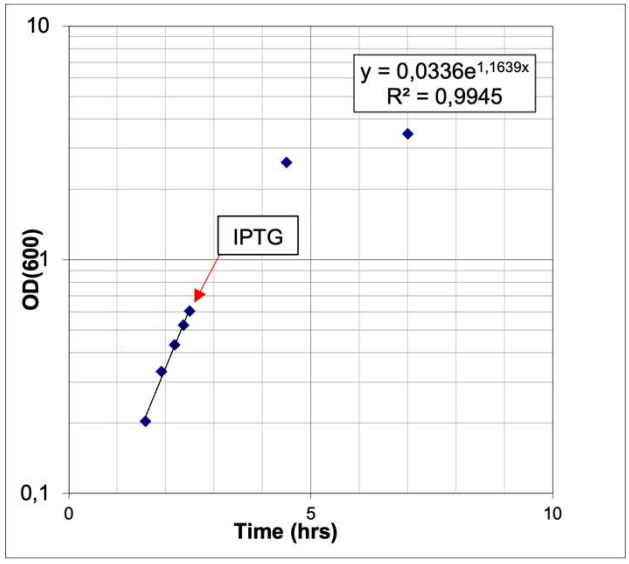


Figure 3.22 Growth curve of *E. coli* BL21(DE3) pLysS pY71 HRS. Specific growth rate (μ): 1.16 hr^{-1}. Red arrow: Induction point with IPTG.

After lysing the pelleted cells and centrifuging the suspension, the supernatant was filtered and loaded to HisTrap FF affinity column. The FPLC chromatogram given in Figure 3.23A shows that there was no elution peak for HRS. The SDS-PAGE gel image of the purification process (Figure 3.23B) proves that the problem with purification was indeed the accumulation of HRS in the pellet, rather than it not binding to the column. Compared to the pellet, there is no distinct band for HRS around 48 kDa in the loaded sample. Sample application and column wash flowthroughs also does not harbor the protein of interest. Since there is no band in the elution fractions either, the purification of HRS with the conventional aaRS purification method established previously was failed.

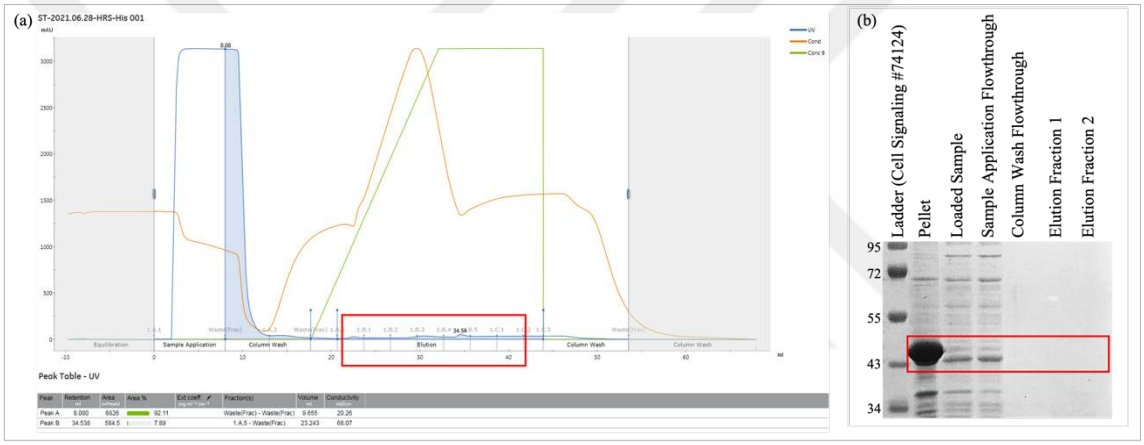


Figure 3.23 Purification of HRS. (a) FPLC Chromatogram obtained after His-tag affinity chromatography. Red box: Elution peak of HRS. Blue line: UV adsorption at 280 nm for protein concentration determination. Orange line: Conductivity for salt concentration measurement. Green line: Concentration of the elution buffer. Red box: expected elution area. (b) SDS-PAGE analysis of the purification (HRS: 48.3 kDa). Red box: expected elution area on the gel.

Expression at 25ºC, 30ºC, 4ºC and 0.25 mM, 0.5 mM and 1 mM IPTG concentration

In Figure 3.23b, it was proved that HRS was accumulating in the pellet after centrifugation, hinting possible inclusion body formation. Major strategies against inclusion body formation are decreasing the expression temperature and using a lower final IPTG concentration. Thus, expression and solubility of HRS was tested at two different temperatures (25ºC and 30ºC), with two different IPTG concentration (0.25 mM and 1 mM). The induction was carried overnight due to temperature decrease. Figure 3.24 shows the SDS-PAGE analysis of the expression study for the pellet and supernatant

comparison. 1 mM IPTG and 30ºC combination has the strongest expression in the 4 tested conditions. The other 3 conditions show faint bands for the expression. Unfortunately, in all 4 conditions HRS accumulates in the pellet after being centrifuged before protein purification, with 1 mM IPTG and 30ºC having the highest amount of protein. An expression study at 4ºC for 5 days was also performed with 0.5 mM IPTG. However, same as the other conditions, there was no band in the SDS-PAGE gel for the supernatant.

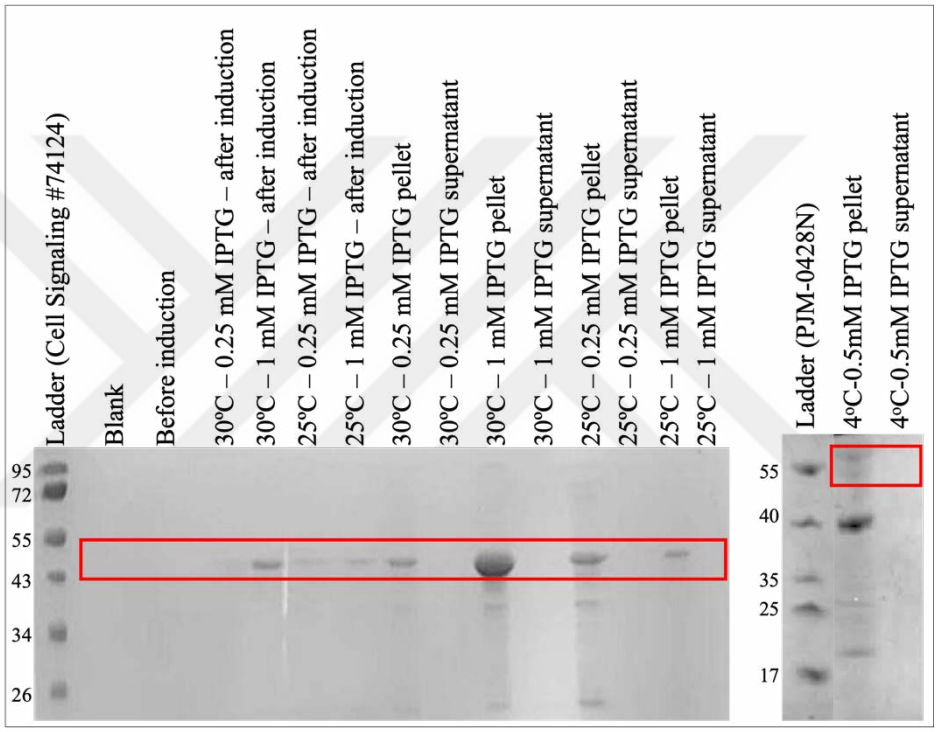


Figure 3.24 SDS-PAGE analysis of HRS expression at varying temperatures and IPTG concentrations (HRS: 48.3 kDa).

Expression study with glycerol and ethanol additives at different temperatures

In the literature, it was seen that the additives such as ethanol and glycerol could increase the solubility, thus decrease inclusion body formation. 2% of ethanol, 15% glycerol and addition of both together were tested. In all 3 conditions, the additives were added to the medium before inoculation of the 30 mL cultures. However, only the culture with 2% ethanol addition reached to OD_{600} 0.6. The cultures with 15% glycerol did not grow past OD_{600} 0.3. Therefore, only the 2% ethanol cultures were investigated for HRS

solubility. Figure 3.25 shows the SDS-PAGE analysis of the expressions. In 27ºC, the expression is stronger than 37ºC. In both temperatures, HRS accumulated in the pellet.

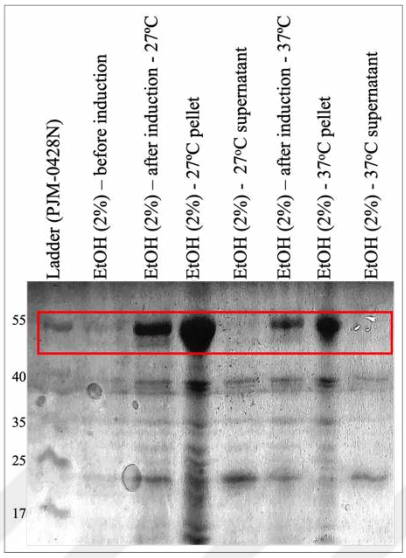


Figure 3.25 SDS-PAGE analysis of HRS expression with ethanol addition at varying temperatures (HRS: 48.3 kDa).

3.2.3 Production of HRS in CFPS

After the unsuccessful *in vivo* production and purification trials, it was decided to test HRS production in CFPS. As the CFPS SDS-PAGE gels are too crowded to see HRS production, western-blot analysis was performed for CFPS samples with 6xHis-tags. Initially, the effect of chaperones and redox potential of the environment on the solubility of HRS was investigated. For this purpose, Skp, GroES and GroEL, as well as varying concentrations of reduced and oxidized glutathione were added to the cell-free reactions. Figure 3.26 shows the western-blot image of the overnight CFPS reaction performed at 30ºC. The antibody dilution rate was 1:3000 and the images were obtained after 30 second of signal detection. A pPaFRS sample (35 kDa) produced and purified by Yağmur Ersoy was used as the western-blot positive control. The pure protein had 3.29 mg/mL concentration and it was diluted 1:10 ratio before uploading to the SDS-PAGE gel.

In the western-blot, the membrane was incubated with the primary antibody (His-tag mouse mAB) overnight at room temperature instead of at 4ºC by mistake, therefore the signals are not very strong due to possible antibody degradation. Still, the Figure 3.26 shows that when HRS was tried to be produced directly in CFPS, there is no band at the

expected area. Instead, without any chaperone or redox agent addition, there is a faint band around 35 kDa in the total protein sample. This result suggests that even if HRS is being produced, it is either produced in the incorrect size or it is being lysed by a protease in the CFPS mixture. Still, the change in the solubility in these lysed proteins was observed. The reaction with 4 µM Skp addition displays a strong band in the soluble protein sample, suggesting Skp is indeed a chaperone that helps the protein to fold correctly and 4 µM concentration gave better results. 4mM oxidized glutathione (GSSG) and 1mM reduced glutathione (GSH), addition also increased soluble protein amount. Bu with the other glutathione concentrations, there was no HRS production. In the reaction with 4 µM GroEL and 4 µM GroES addition, the strongest signal was observed for total protein, however there was almost no protein in the soluble protein sample.

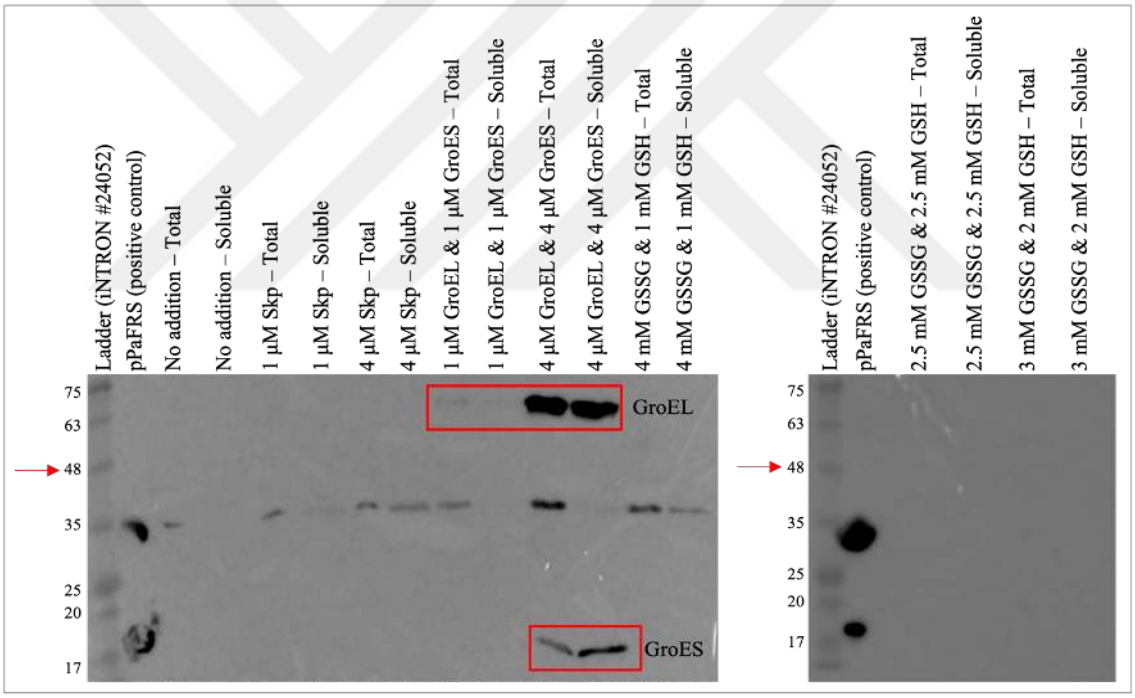


Figure 3.26 Western blot analysis of cell-free HRS production with additions of Skp, GroEL-GroES chaperones, as well as oxidized (GSSG) and reduced (GSH) glutathione in varying concentrations (HRS: 48.3 kDa). Red arrow: Expected area of full-size HRS.

As it was suspected that there was protease degradation in the protein of interest, a protease inhibitor cocktail was added to the cell-free reaction with varying concentrations. The CFPS reactions ran overnight at 30ºC. In the western blot, a different antibody was used (His-tag (27E8) Mouse mAB (HRP Conjugate)) with 1:2000 ratio. This time the antibody incubation was carried overnight at 4ºC as it should have been.

Signal was obtained after 5 minutes of detection. In the western-blot image in Figure 3.27 the positive control (3.29 mg/mL pPaFRS 1:10 diluted) proves that the blotting worked fine. Possibly due to fresh antibody usage, the intact protein band around 48 kDa could be observed on the membrane unlike previous cell-free reaction western image (Figure 3.26). Without any chaperone or protease inhibitor addition, HRS is produced in the right size, however almost all of it is insoluble (last 2 lanes). Addition of protease inhibitor along with 4 μM Skp results in a high amount of soluble protein. Interestingly, the concentration of protease inhibitor does not change the intact protein yield too much. Therefore, it was decided to use 1X concentration in the follow-up experiments. Addition of glutathione to cell-free mixture along with protease inhibitor increased the total and soluble protein amounts, however they are still lower than the reactions with Skp. Combining Skp with glutathione did not have a synergistic effect. In fact, the full-size protein is not observed.

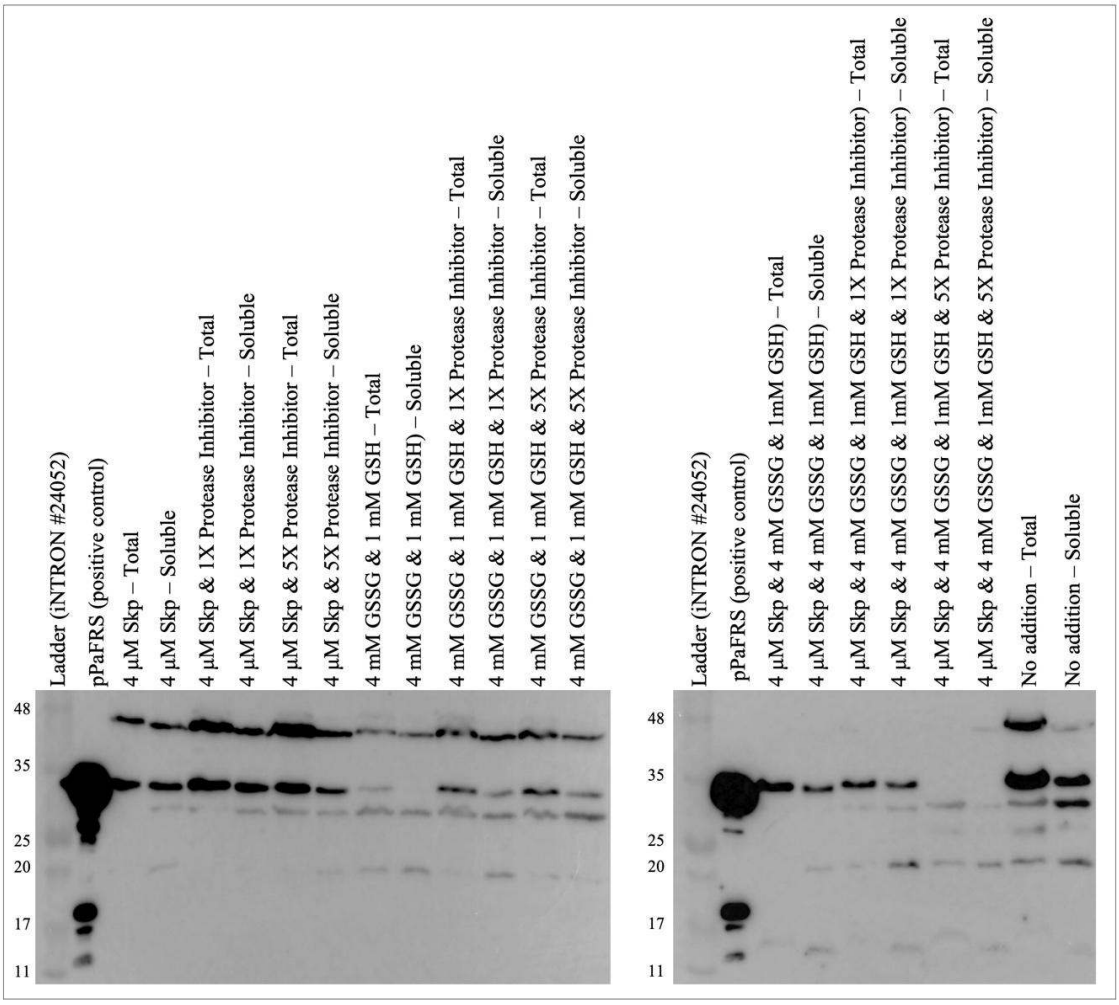


Figure 3.27 Western blot analysis of cell-free HRS production with additions of Skp chaperone, protease inhibitor and glutathione (HRS: 48.3 kDa).

Cell extract preparation from E. coli BL21 and E. coli ARG1 cell lines

The protease inhibitor addition could not fully restore the intact protein. Therefore, it was decided to see which protease was causing the proteolysis. It is known that the 759.T7.opt cell extract contains OmpT protease, and it was suspected that this enzyme was cutting cell-free produced HRS. To test this hypothesis, two cell extracts that lack the OmpT protease were prepared. ARG1 cell line is deficient in OmpT protease only while BL21 cell line is deficient in both OmpT and Lon proteases.

E. coli BL21 cells were grown in LB medium at 37ºC until hr^{-1} value dropped below 0.6 to ensure the culture stays in the log phase. The culture was then harvested at OD_{600} 3.372 and the specific growth rate 0.651 hr^{-1} (Figure 3.28a), and 6 g cell pellet was obtained.

ARG1 cell line was initially grown in M9 media at 37ºC. The culture was harvested when the specific growth rate was 0.609 hr^{-1} and OD_{600} was 0.64 (Figure 3.28b). The obtained amount of ARG1 cell pellet was only 1.81g, so it was very little to lyse the cells in the homogenizer. Therefore, the experiment was repeated with LB medium in duplicates (Figure 3.28c). The total cell pellet weight of two 1 L cultures was 3.44. When combined with the pellet produced from the growth in M9 media, a total of 5.44 g cell pellet was obtained for ARG1.

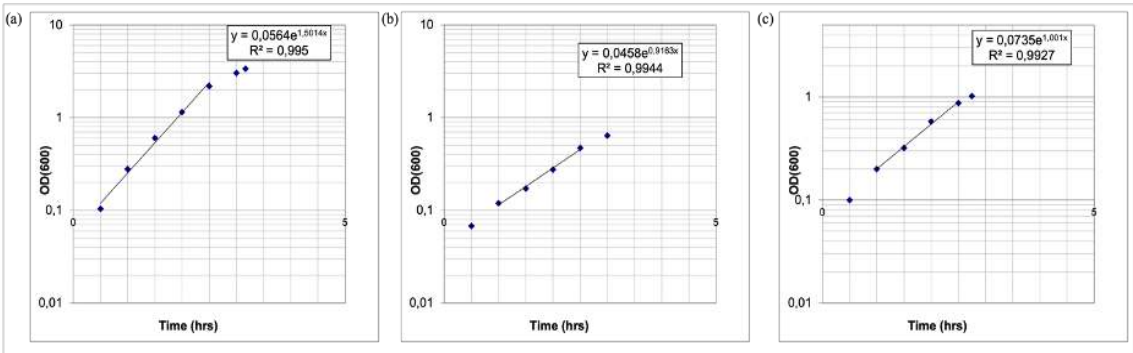


Figure 3.28 Growth curve of (a) *E. coli* BL21 in LB medium. Specific growth rate (μ): 1.50 hr^{-1}. (b) *E. coli* ARG1 in M9 medium. Specific growth rate (μ): 0.9163 hr^{-1}. (c) *E. coli* ARG1 in LB medium. Specific growth rate (μ): 1.001 hr^{-1}.

To prepare cell extracts from the pellets, the S30 cell extract preparation procedure was used. 11mL of ARG1 extract and 9mL of BL21 extract were obtained. The cell extracts were aliquoted and flash frozen at -150ºC (5min). ARG1 and BL21 cell extracts

were then tested for their sfGFP and HRS production efficiency. Given in Figure 3.29, while the CFPS reaction gave similar sfGFP yield with the opt. and BL21 extracts, ARG1 extract resulted in significantly lower sfGFP production.

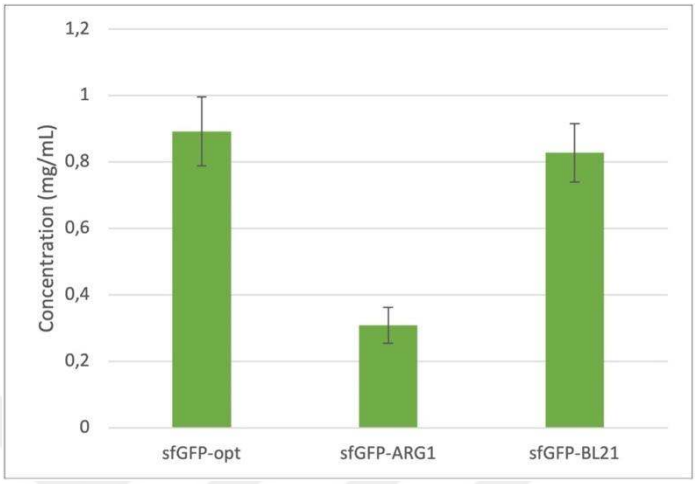


Figure 3.29 Cell-free produced soluble sfGFP yields with different cell extracts. Error bars represent the difference in soluble protein concentration between reaction replicas.

To analyze HRS production with different cell extracts, 4 μM Skp and 1X protease inhibitor additions, a western blot was done for the total and soluble samples. Figure 3.30a and 3.30c are the Western-blot images and Figure 3.30b and 3.30d are their corresponding SDS-PAGE gels. To show the signal strength in Western-blot analysis is correlated with the protein concentration of the loaded sample, serial dilutions of a pure pPaFRS protein were analyzed with Western-blot. As expected, the signal got weaker with the increased dilution rate. The strong signal of the positive control in Figure 3.30c, (6xHis-tagged pPaFRS, 3.29 mg/mL, 1:10 diluted) proves that the western blot method worked well. When ARG1 or BL21 extracts are used in cell-free HRS production instead of opt. extract, the intense band around 35 kDa is not observed, proving that it was indeed OmpT protease in the opt. extract that leads to the major digestion product of HRS. Also, the soluble yield is observed to be higher in the reactions where BL21 extract is used instead of ARG1 extract. Thus, it was decided to continue the CFPS experiments with BL21 extract. The proteolyzed protein around 35 kDa synthesized in the opt. extract can be seen in the SDS-PAGE analysis (Figure 3.30d), suggesting that a high amount of protein is produced with the opt. extract.

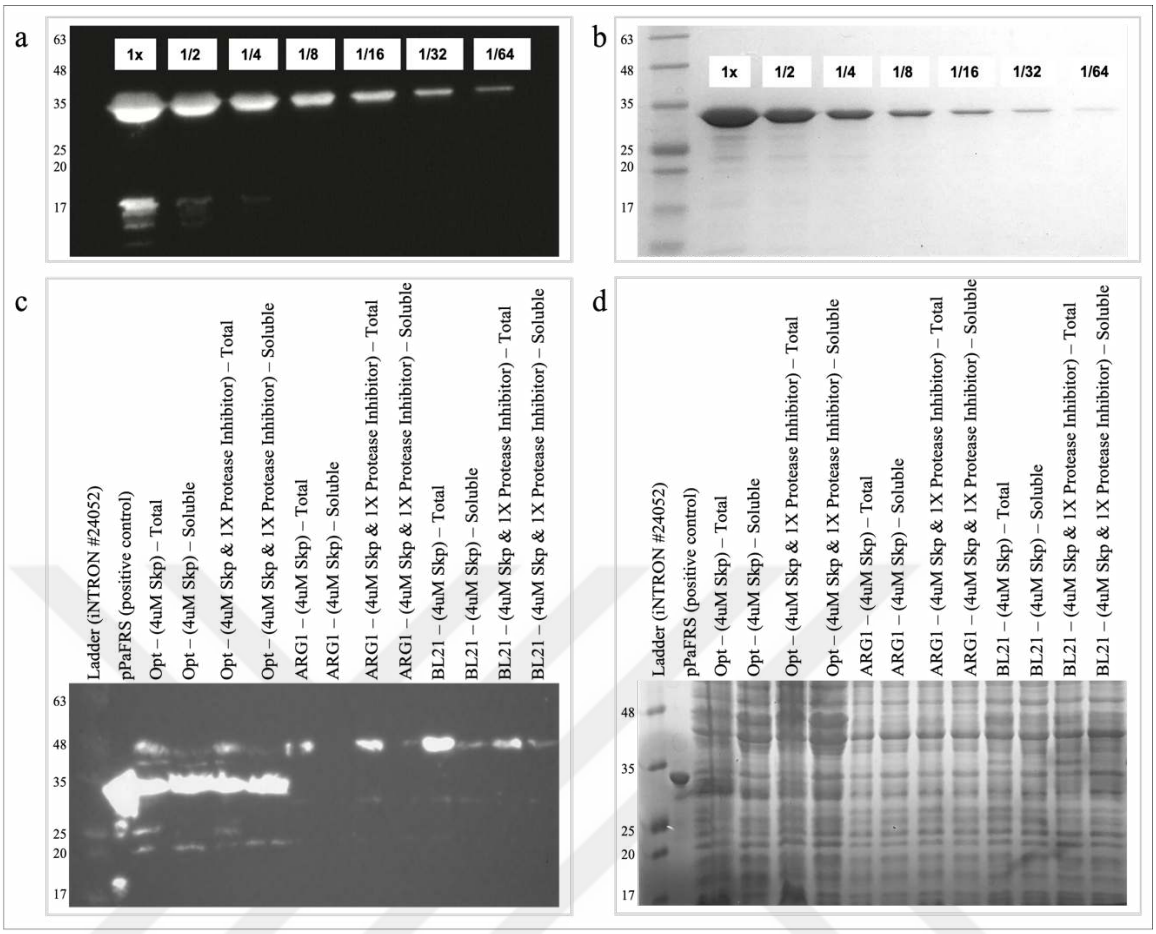


Figure 3.30 Western blot and SDS-PAGE analysis comparison for pPaFRS and HRS. (a) Western-blot analysis of pure pPaFRS protein dilutions. (b) SDS-PAGE analysis of pure pPaFRS protein dilutions (1X pPaFRS: 3.29 mg/mL). (c) Western blot analysis of cell-free HRS production with different cell extracts. (d) of cell-free HRS production with different cell extracts (HRS: 48.3 kDa).

A densitometry analysis was then performed for the Western-blot image of the pPaFRS dilutions shown in Figure 3.30a. From the peaks plotted for each band in the image, an area under the peak versus concentration graph was plotted. The standard curve could be helpful in the future to estimate the amount of protein the Western-blot gel when the signal is not overwhelming.

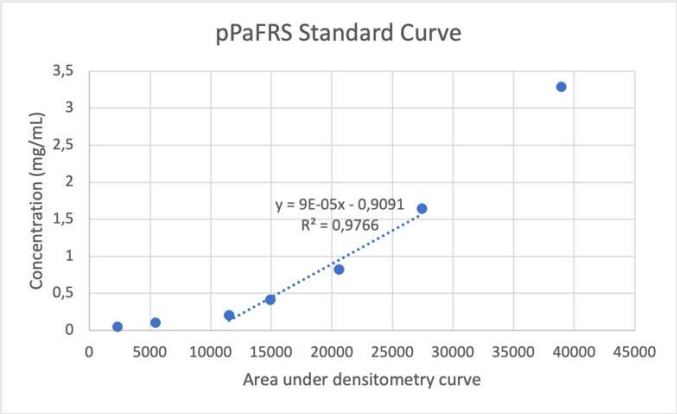


Figure 3.31 Densitometry analysis of the Western-blot image for pPaFRS dilutions.

Since HRS produced with BL21 cell extract and addition of 1X protease inhibitor showed promising results to obtain soluble, intact protein, it was decided to produce HRS with the same conditions in a large-scale CFPS reaction. Before that, magnesium concentration to be used in the reaction was aimed to be optimized. It is known that the magnesium amount in the cell-free reaction is high of importance, and it needs to be adjusted for every cell-extract. Therefore, varying magnesium concentrations, from 8 mM to 20 mM, were tested in a CFPS experiment to produce sfGFP. As a positive control sfGFP was also produced with the opt. extract and 8 mM magnesium as this was the reaction standard before. Figure 3.32 shows that 8 mM and 12 mM magnesium concentrations are proven to be the optimum concentrations when using BL21 extract. As 8 mM magnesium was used in the previous CFPS HRS production trials, it was decided to continue with that concentration.

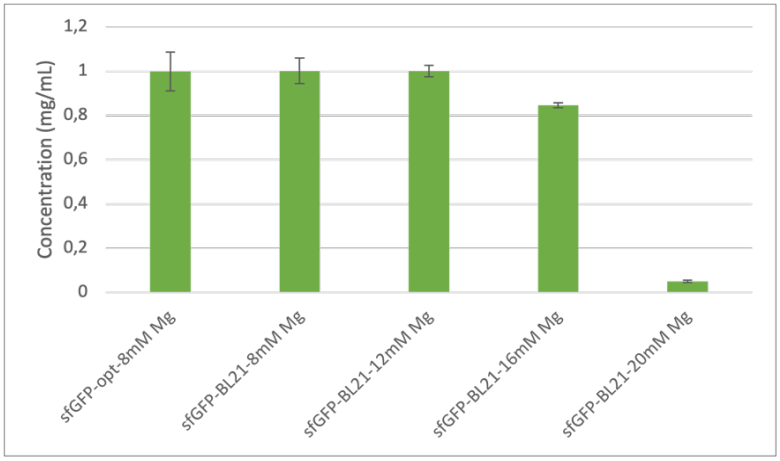


Figure 3.32 BL21 cell extract magnesium concentration optimization for CFPS. Error bars represent the difference in soluble protein concentration between reaction replicas.

After magnesium optimization of the BL21 cell extract, to test the efficiency of it in terms of ncAA insertion, wild-type sfGFP and sfGFP23TAG variant were produced in CFPS. Interestingly, both wild-type sfGFP and sfGFP23.pPaF concentrations were higher in the reactions conducted with the BL21 extract (Figure 3.33).

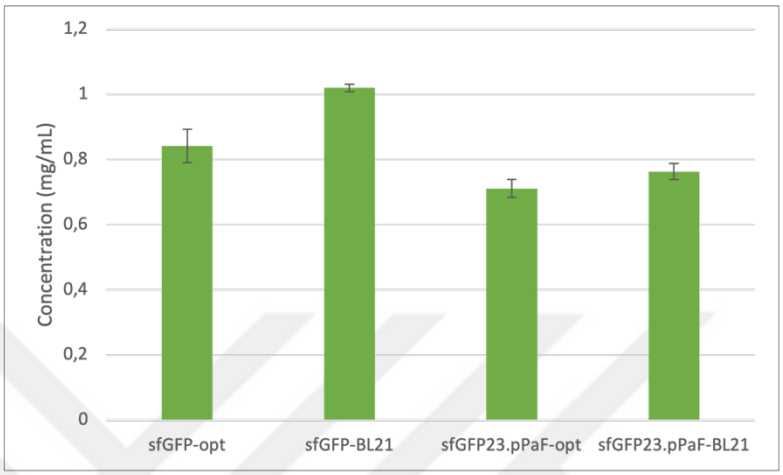


Figure 3.33 Comparison of opt. and BL21 cell extracts in terms of cell-free soluble sfGFP production. Error bars represent the difference in soluble protein concentration between reaction replicas. sfGFPX.Y: **Y** ncAA is inserted at the **X** position of sfGFP.

HRS was then produced in a 3 mL CFPS reaction with BL21 extract, 1X protease inhibitor and 4 µM Skp at 30ºC overnight. After centrifugation of the reaction, the supernatant was filtered and loaded to the His-tag affinity column. FPLC chromatogram given in Figure 3.34a shows two small peaks in the elution area. These elution fractions were dialyzed against a 10 mM potassium phosphate buffer and a Western-blot analysis was done for the purification, shown in Figure 3.34b. The positive control displayed a strong signal, proving there was no problem with the blotting or signal detection. The figure also shows HRS again accumulated in the pellet, and even though there is a band in the desired location for the loaded sample, some of it washed away during sample application. The concentration of the produced HRS was probably lower than the binding capacity of the column. Thus, there is no observed band in the elution fractions or in the dialysis samples.

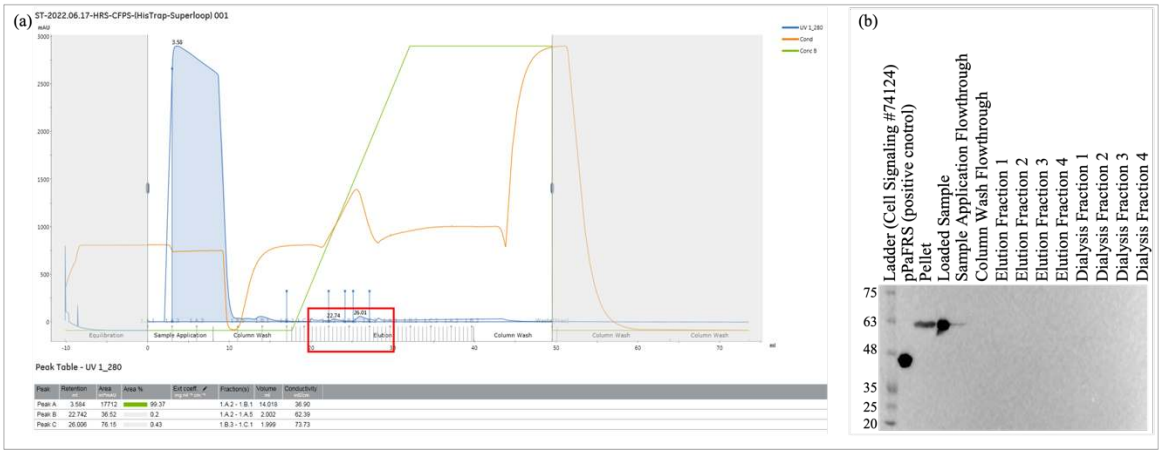


Figure 3.34 Purification of HRS produced in preparative (3mL-scale) CFPS. (a) FPLC Chromatogram obtained after His-tag affinity chromatography. Red box: Elution peak of HRS. Blue line: UV adsorption at 280 nm for protein concentration determination. Orange line: Conductivity for salt concentration measurement (b) Western-blot analysis of chromatography fractions (HRS: 48.3 kDa).

3.2.4 SDM to obtain an OmpT protease resistant HRS

HRS produced with the BL21 cell extract in CFPS could not be purified due to low protein yields. Moreover, even though HRS was subject to proteolysis when it was produced with the opt. extract *in vitro*, the amount of produced protein was much more than the ones produced with other extracts. Therefore, the strategy was changed to making HRS protease resistant instead.

The size of the major proteolysis product is around 35 kDa. So, the possible cut sites of OmpT on HRS were investigated. It is known that OmpT protease cuts proteins from basic dipeptide residues and the 35 kDa fraction should include 6xHis-tag as it is seen in the western-blot membrane. Table 3.3 shows the possible target sited of OmpT protease and the size of the cut fraction that includes the C-terminal 6xHis-tag. K104/K105 site was identified as the most potent site to lead to the major proteolysis product as it gives a product of 36 kDa.

Table 3.3 OmpT protease target sites and the resulting protein fraction sizes.

Possible OmpT cut sites on HRS	Size of the proteolysis product that includes 6xHis-tag (kDa)
K3/K4	47.97

R66/K67	40.68
K70/R71	40.22
K104/K105	36.37
R191/R192	27.15
R192/R193	26.99
R275/K276	17.11
R295/R296	14.81
K403/R404	8.63

Primers were then designed to mutate K105 residue to alanine via SDM. The miniprepped plasmid with the SDM was sent to sequencing but it was seen that the SDM was unsuccessful. Even when the experiment was repeated, it was unsuccessful again. Due to time limitations, the procedure could not be optimized to obtain the HRS.K105Avariant.

3.2.5 Designing an HRS variant with a solubility tag

As another strategy to obtain soluble HRS by *in vivo* production, a fusion protein that contains a solubility tag called SmbP, originated from *N. europaea,* as well as an enterokinase cleave site at the upstream of HRS gene was designed. Ribosome binding site (RBS) and NdeI restriction site were added at the upstream of SmbP sequence and after the enterokinase cleavage site, SalI restriction site was added. At the C-terminal of the HRS gene, HindIII restriction site was placed. The designed fusion protein (Figure 3.35) was ordered from Twist Bioscience in an expression vector.

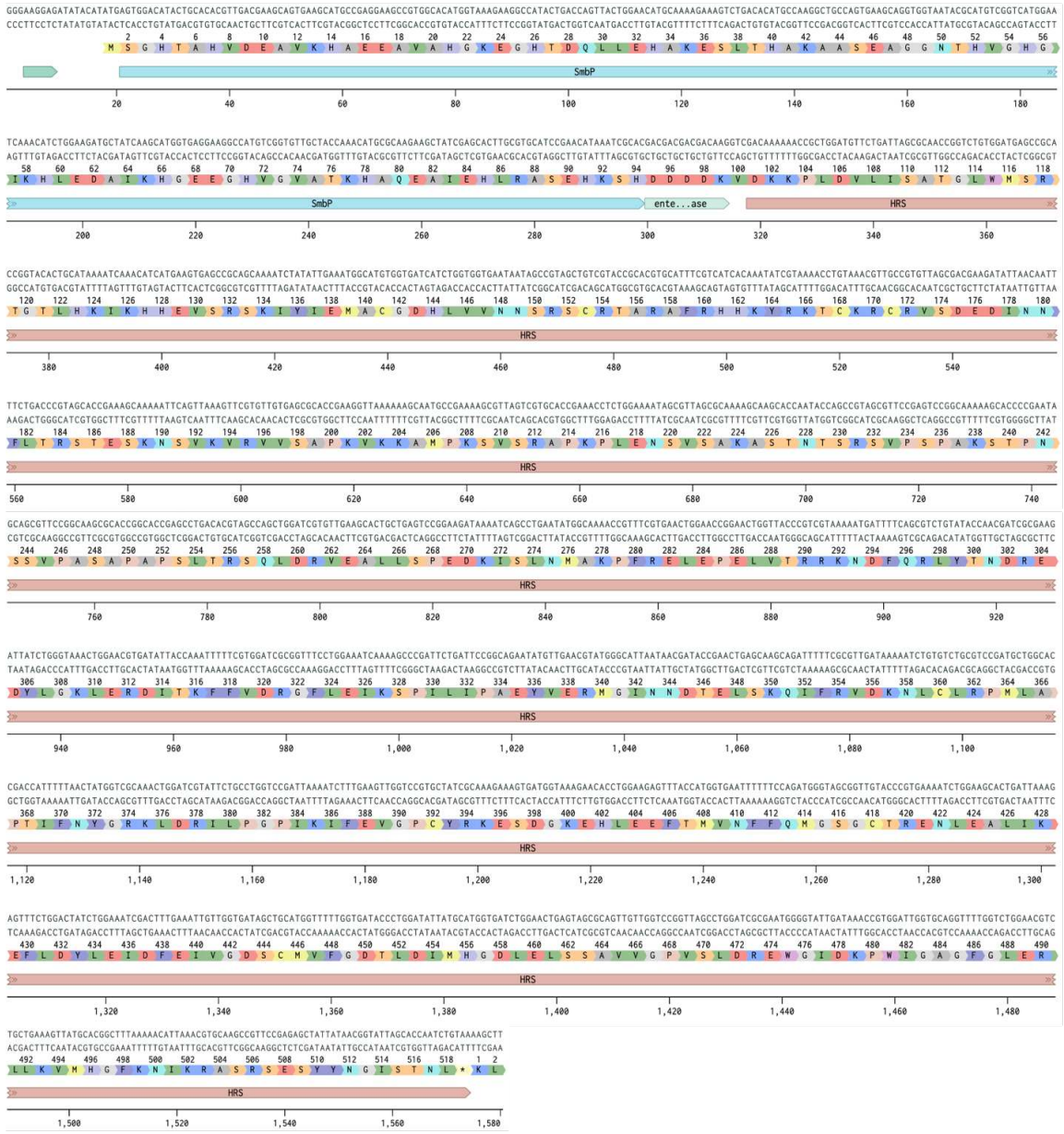


Figure 3.35 The designed HRS fusion protein with SmbP solubility tag.

3.3 Production, Purification and Analysis of CpI Holo-Hydrogenase

3.3.1 E. coli BL21 (DE3) ΔiscR pACYC HydEFG competent cell preparation

To transform pET21b sCpI sIIC StrepII plasmid into a cell line containing maturase enzyme genes, competent cells were prepared from the E. coli BL21 (DE3) ΔiscR pACYC HydEFG cell line. The cells were harvested at OD_{600} 0.476, as it can be seen from Figure 3.36.

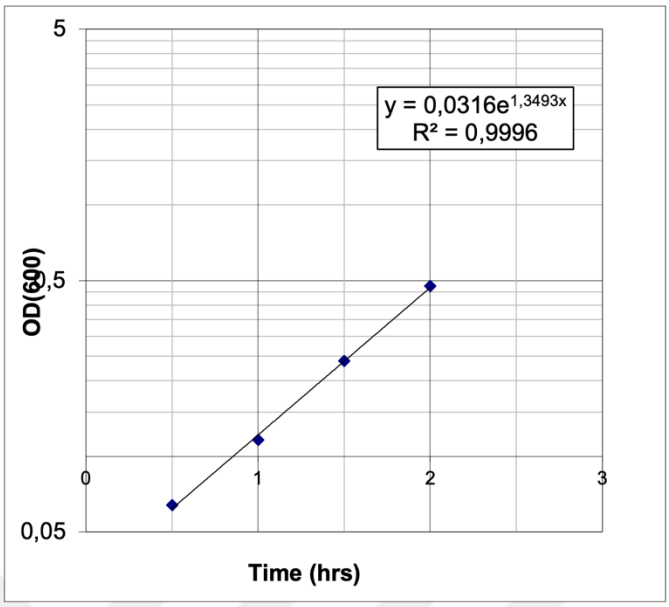


Figure 3.36 Growth curve of *E. coli* BL21 (DE3) ΔiscR pACYC HydEFG for competent cell preparation. Specific growth rate (μ): 1.34 hr^{-1}.

3.3.2 Production and purification of CpI Hydrogenase in vivo under the T7 promoter

The plasmid containing CpI Hydrogenase gene was transformed to the *E. coli* BL21 (DE3) ΔiscR pACYC HydEFG competent cells described in the previous section. At OD_{600} 0.368, the culture flask was cycled inside an anaerobic glovebox and 5 mM L-cysteine along with 10 mM sodium fumarate. The color of the culture became darker with the addition of 5 mM L-cysteine. Production of the wild-type CpI enzyme was induced anaerobically at room temperature, with 0.25 mM IPTG *in vivo* at OD_{600} 0.436 (Figure 3.37a) . After an overnight protein production, the color of the culture turned full black, as shown in Figure 3.37b. From 1L culture, 2.4g of wet cell pellet was obtained.

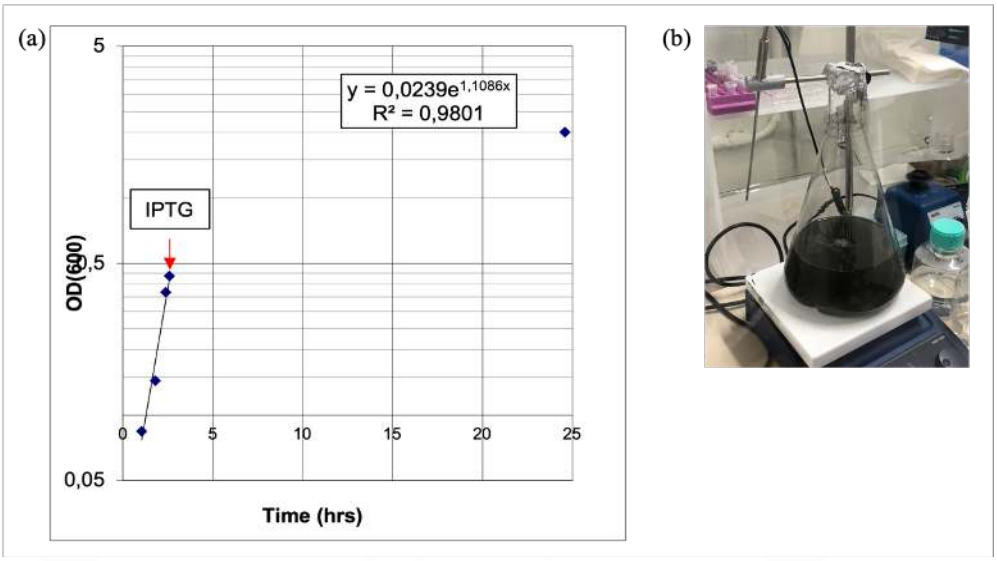


Figure 3.37 (a) Growth curve of *E. coli* BL21 (DE3) ΔiscR pACYC HydEFG pET21b sCpI sIIC StrepII for CpI hydrogenase production. Specific growth rate (μ): 1.11 hr^{-1}. Red arrow: Induction point with IPTG. (b) Culture broth after induction. The deep black color is mainly due to the production of maturation enzymes.

When purifying the Strep-tagged hydrogenase using StrepTactin affinity chromatography, 1 gram cell pellet was used from the total 2.4 grams. During the purification process, the resin just flowed through the support of the column when the resin was poured into the column. Therefore, the resin was transferred to another column containing a different support. The leakage of the resin was minimum then. When the column was washed with wash buffer after sample application, most of the red-brownish color was collected at the flowthrough (Figure 3.38c), meanwhile the elution sample had a very faint color as can be seen in Figure 3.38d. The volume of the elution fraction is 3 mL.

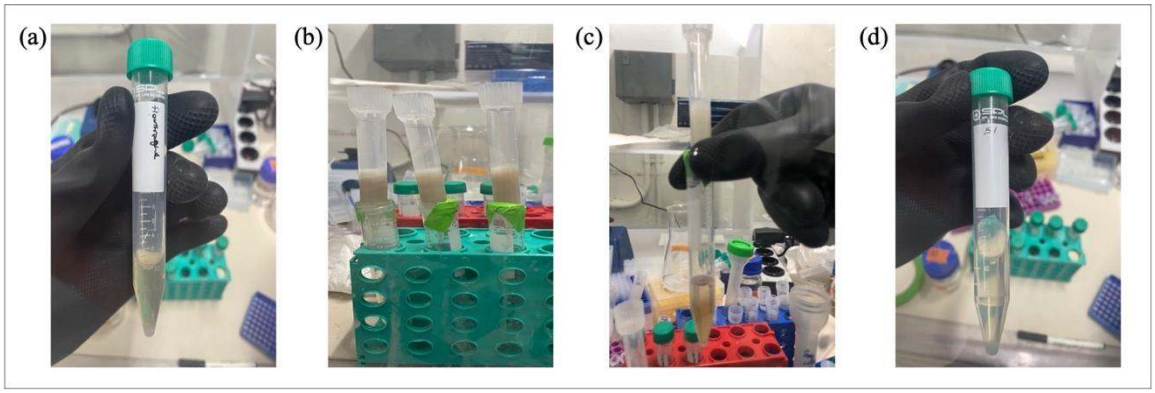


Figure 3.38 Anaerobic purification of CpI hydrogenase by StrepTactin affinity chromatography. (a) Flow-through fraction. (b) Wash step. (c) Column wash flowthrough. (d) Elution fraction. 3 mL Cytiva StrepTactin resin was packed inside a 10 mL disposable column.

The elution fraction has a distinct, clear band for CpI hydrogenase in the SDS-PAGE gel shown in the Figure 3.39. The sample application and column wash flowthroughs also contain the band for the protein of interest along with many other proteins.

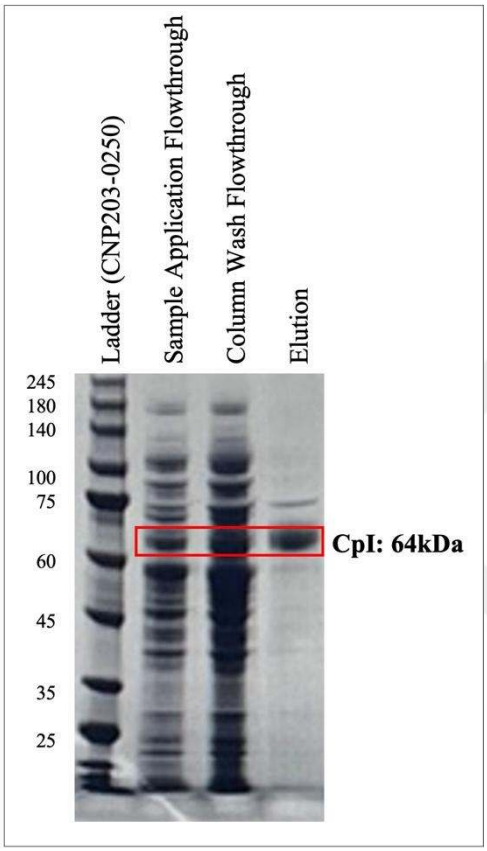


Figure 3.39 SDS-PAGE analysis of CpI Hydrogenase purification by StrepTactin affinity chromatography.

In order to determine the concentration of the purified protein, Bradford assay was performed. Table 3.4 shows the calculated concentration and molarity of CpI Hydrogenase before and after concentrating the sample using Pierce Protein concentrator (PES, 10K MWCO, Thermo Scientific Cat. No: 88513). After purifying the protein, the concentration was calculated to be 0.23 mg/mL +/- 0.01 and the molarity was 3.57 µM. Since the elution fraction volume was 3 mL, 0.7 mg protein was purified from 1 g of cell. From 1 liter culture, with 2.4g wet cell in total, the yield would be 1.67 mg.

As the protein would be used in activity assays in the future, it was concentrated. 0.5 mL of the elution fraction was taken for concentrating with Pierce Protein

concentrator (PES, 10K MWCO, Thermo Scientific Cat. No: 88513). The average concentration of the sample was determined to be 0.42 mg/mL with a molarity of 6.52 µM, 2-fold of the initial sample (Table 3.4).

Table 3.4 Concentration determination of CpI Hydrogenase with Bradford assay.

Sample	Average Concentration (mg/mL)	Standard deviation	Molarity (µM)
Elution Fraction (before concentrating)	0.232	0.01	3.56
Elution Fraction (after concentrating)	0.423	0.04	6.52

Alternatively, the protein production was also performed anaerobically using Strep-Tactin®XT® 4Flow trial kit (IBA Lifesciences). When the column was loaded with the sample, a brown color that can be seen in Figure 3.40a was observed. As the Buffer BXT washed through the column, the protein ring that can be clearly seen in the Figure 3.40b) and Figure 3.40c moved downwards. After a total of 3 CV Buffer BXT wash, the column was cleared of the brown color (Figure 3.40d). Elution Fraction 2 had a darker brown color compared to the Elution Fraction 1 and 3.

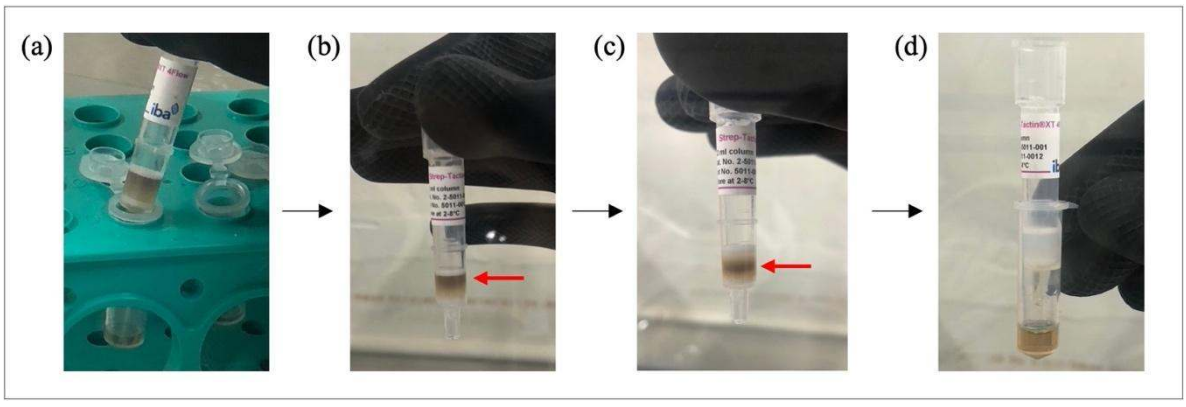


Figure 3.40 Purification of CpI hydrogenase by StrepTactin affinity chromatography using IBA Strep-Tactin®XT® gravity flow column (CV = 0.2 mL). (a) The column after sample application. (b) The column after the wash step. (c) The column after passing 0.6 CV of elution buffer. (d) The column after 1.6 CV elution buffer & Elution Fraction 2. Red arrow: Movement of the adsorbed protein through the column.

Figure 3.41 shows the SDS-PAGE analysis of the purification of CpI Hydrogenase using IBA Strep-Tactin®XT® gravity flow column. Even though the gel did

not run quite well due to the SDS Running buffer being old, Elution Fraction 2 and 3 display the band for the protein of interest clearly. Column wash flowthrough does not show a sharp band for CpI hydrogenase, suggesting that most of it stayed bound to the column at this step.

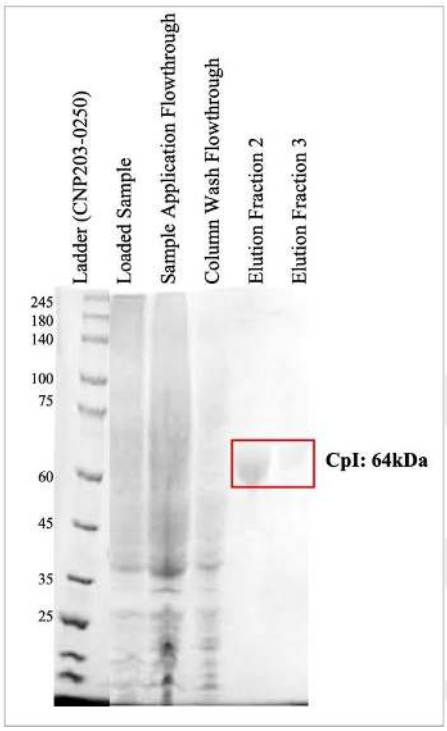
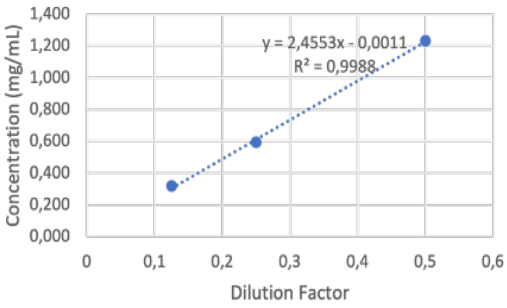


Figure 3.41 SDS-PAGE analysis of CpI Hydrogenase purification by StrepTactin affinity chromatography using StrepTactin®XT® gravity flow column (IBA).

To determine the concentration of the elution fractions, Bradford assay was performed. Several dilutions of the fractions were prepared as can be seen in Table 3.5 and Table 3.6. From Table 3.5, a dilution factor versus concentration graph was drawn for EF2. When X is assumed to be 1 on the equation of the graph to calculate 1X protein concentration, Y= 2.4553-0,0011 = 2.4542 is obtained. From the 65 kDa molecular weight of the protein and 2.45 mg/mL concentration, molarity was calculated to be 37.76 μM.

Table 3.5 Concentration determination for EF2 of CpI Hydrogenase with Bradford assay.

Dilution factor	Average Concentration (mg/ml)	Standard deviation
0.5	1.23	0.07
0.25	0.59	0.07
0.125	0.32	0.05



From the dilution factor versus concentration graph drawn from Table 3.6, 1X concentration of EF3 was calculated to be 0.38 mg/mL. Molarity was then calculated to be 5.83 µM from concentration and molecular weight.

Table 3.6 Concentration determination for EF3 of CpI Hydrogenase with Bradford assay.

Dilution factor	Average Concentration (mg/ml)	Standard deviation
1	0.38	0.00
0.5	0.13	0.00
0.25	0.04	0.00

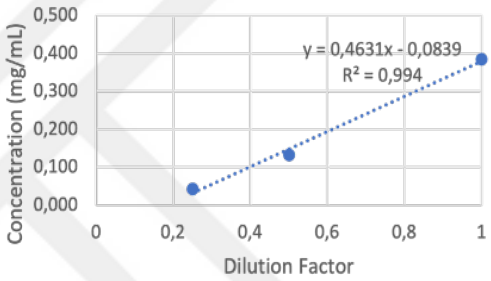


Table 3.7 shows the summary of the calculations explained previously. As the molarity of EF2 is high enough for activity assays, it was not concentrated and used in the following experiments. The volume of EF2 was 0.32 mL, thus it contained 0.79 mg protein. EF3 was 0.16 mL, so it had 0.06 mg protein. In total 0.84 mg protein was purified from 2 grams of wet cell. The 1-liter culture the protein was purified from had 5.2 grams of wet cell. If the protein amount is multiplied with the total amount of cells to the amount of cell used in the purification (5.2g / 2g = 2.6), approximately 2.19 mg protein was yielded from 1-liter culture.

Table 3.7 Calculated concentrations and molarities for pure EF2 and EF3 of CpI Hydrogenase.

Sample	Average Concentration (mg/mL)	Molarity (µM)
Elution Fraction 2	2.45	37.76
Elution Fraction 3	0.37	5.83

3.3.3 Production and purification of Rice root FNR

E.coli BL21 (DE3) pET21B RrFNR cell line was used for a small-scale production of RrFNR in LB medium. 0.5 mM IPTG was used for induction of the protein at 30°C overnight. The culture was harvested at OD_{600} 1.588 (Figure 3.42) and resulted 0.51g cell pellet.

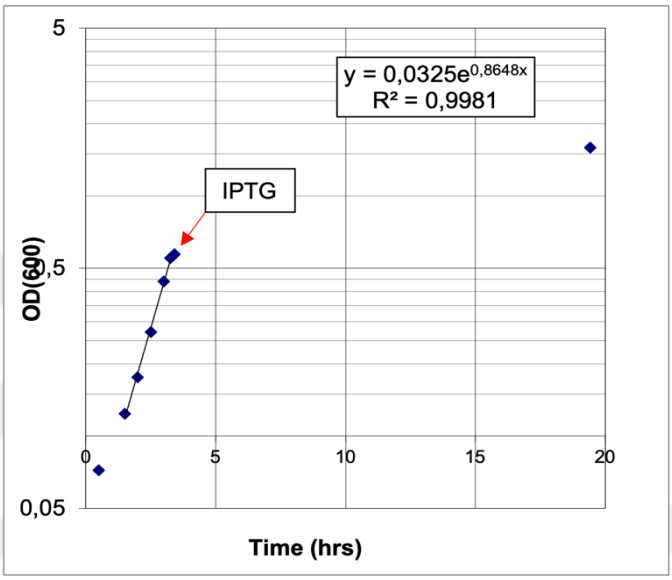


Figure 3.42 Growth curve of small-scale *E. coli* BL21 (DE3) pET21b RrFNR. Specific growth rate (μ): 0.86 hr^{-1}. Red arrow: Induction point with IPTG.

In an initial purification trial of RrFNR using CM Sepharose® Fast Flow resin, the protein did not bind to the column probably due to high ionic strength of the sample as the pH of the lysate was not adjusted to 5.5 prior to loading it to the column. In the following trials, sample pH was always adjusted before purification. However, addition of acid would make the suspension cloudy, therefore a centrifugation at 6,500 rpm and 30 minutes would be performed before loading to the column.

From the 0.51 g cell pellet, 0.25 g cell was used for the purification initially. When the cell lysate was loaded to the column, it gave a faint yellow color that can be seen in Figure 3.43a. As the salt concentration of the loaded elution buffer increased, the yellow protein ring slowly went downwards.

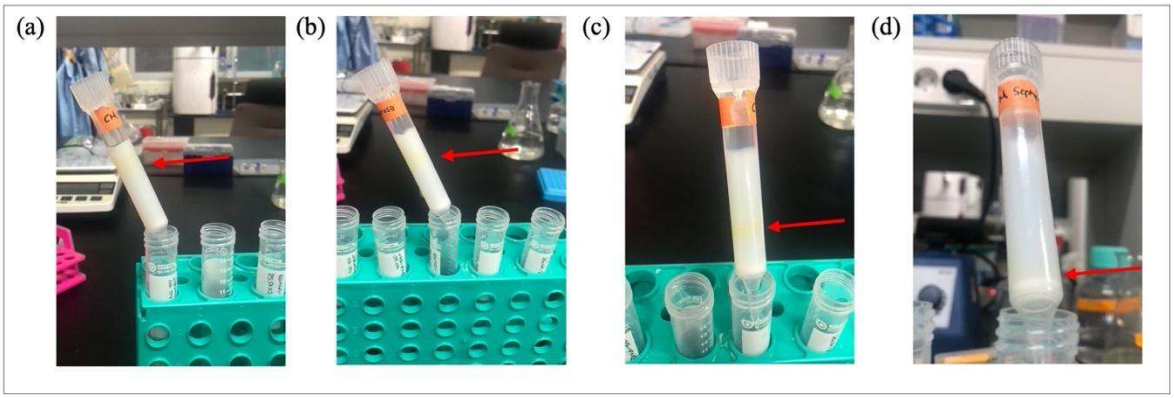


Figure 3.43 Purification of RrFNR by cation exchange chromatography. (a) Sample application. (b) 150 mM NaCl Elution. (c) 250 mM NaCl Elution. (d) 350 mM NaCl Elution.

The resulting elution fractions had a faint yellow color as well, as RrFNR has a natural yellow color. EF2 and EF3 were pooled together, making a total of 13 mL sample. This pooled sample was loaded into Amicon® Stirred Cell concentrator and the volume was initially dropped to ~4 mL. From 4 mL, two 5X buffer exchanges were done with 25mM Tris, pH 8.0 buffer. Then using a 0.5 mL Pierce Protein concentrator, the volume was reduced to 0.2 mL. After adding 5% glycerol to the final sample, it was stored at -80ºC. As the purification was successful, the rest of the cell pellet was purified with the exact protocol. Figure 3.44 shows the purified samples clearly (the last 2 bands). Almost all RrFNR bound to the column, as the band in the sample application flowthrough is very faint. There is no visible RrFNR band in the EF1 sample, suggesting the protein started eluting at 250 mM NaCl concentration. The first purification sample was labeled as RrFNR Lot1 and the second one was labeled as RrFNR Lot2.

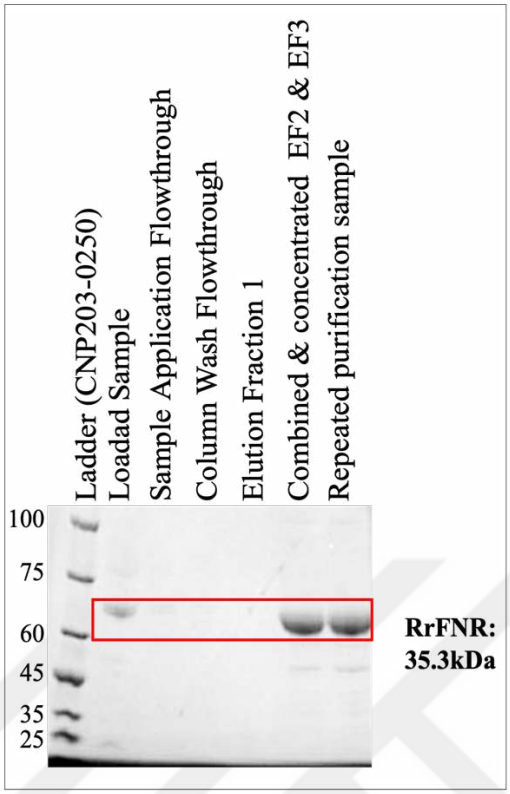
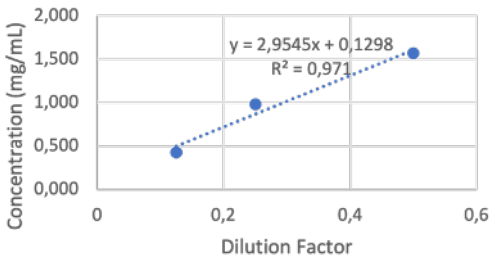


Figure 3.44 SDS-PAGE analysis of RrFNR purification by cation exchange chromatography. Red box: expected sample location on the gel.

To determine the concentration of the two RrFNR lots, Bradford assay was done. Samples ranging from 0.5 to 0.125 dilution factor were prepared as can be seen in Table 3.8. A dilution factor versus concentration graph was drawn for RrFNR Lot1. When X on the equation was assumed to be 1, Y= 2.954+0.1298 = 2.96748 was obtained. This means that the concentration of 1X RrFNR Lot1 is 2.97 mg/mL. From this result and from the molecular weight of the protein, which is 35.3 kDa, the calculated molarity is 84.04 µM.

Table 3.8 Concentration determination for RrFNR Lot1 with Bradford assay.

Dilution factor	Average Concentration (mg/ml)	Standard deviation
0.5	1.57	0.03
0.25	0.98	0.00
0.125	0.43	0.00



In a similar way, several dilutions of RrFNR Lot2 were prepared, ranging from 0.5X to 0.125X. From Table 3.9, a dilution factor versus concentration graph was plotted and the 1X concentration of RrFNR Lot2 was calculated to be 2.83 mg/mL. Then the molarity was calculated as 80.16 µM from concentration and molecular weight.

Table 3.9 Concentration determination for RrFNR Lot2 with Bradford assay.

Dilution factor	Average Concentration (mg/ml)	Standard deviation
0.5	1.51	0.01
0.25	0.98	0.03
0.125	0.50	0.04

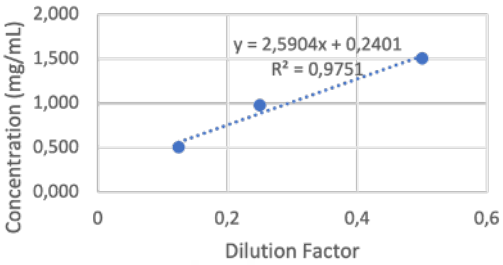


Table 3.10 shows the summary of the calculations above. The volume of each lot was 0.2 mL. In this case, Lot1 had 0.593 mg RrFNR meanwhile Lot2 had 0.566 mg. In total, 1.16 mg protein was obtained from a 100 mL culture. When proportioned to 1000 mL, the yield is approximately 11.6 mg RrFNR per liter culture.

Table 3.10 Calculated concentrations and molarities for RrFNR Lot 1 and Lot 2.

Sample	Average Concentration (mg/mL)	Molarity (µM)
RrFNR Lot1	3	84
RrFNR Lot2	2.8	80.2

Alternatively, RrFNR was also produced in large-scale with 300 mL TB medium. After an overnight induction at 30°C with 0.5 mM IPTG, the culture was harvested at OD_{600} 19.16 (Figure 3.45). The culture resulted in 7.3g cell pellet.

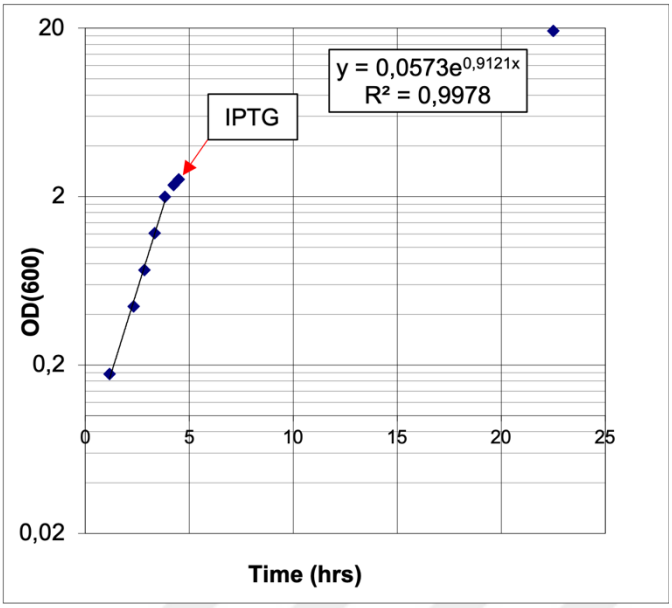


Figure 3.45 Growth curve of large-scale *E. coli* BL21 (DE3) pET21b RrFNR in TB medium. Specific growth rate (μ): 0.91 hr^{-1}. Red arrow: Induction point with IPTG.

As the amount of the cells and thus the lysate was more in the large-scale production, a 20 mL gravity flow column containing 10 mL CM Sepharose® Fast Flow resin was used in the purification of RrFNR. Also different from the previous purification trials, the lysate was diluted in wash buffer with a ratio of 1:1.5 instead of 1:5 due to the already large sample volume. The purified protein was also not concentrated with Pierce Protein concentrator (0.5 mL).

When the sample was loaded to the column, an overwhelming yellow color was observed. Just like the previous trials, the yellow color went downwards as the NaCl concentration of the elution buffer increased (Figure 3.46). After 1.6 CV 350 mM NaCl containing elution buffer (EB3) was run through the column, the yellow color disappeared from the column.

The Elution Fraction 2 and 3 were combined, making a total of 31 mL volume. The combined sample was then loaded into the Amicon stirred-cell concentrator, and the volume was reduced to ~10 mL. From 10 mL, two 5X buffer exchanges were done with the exchange buffer. To the final 9 mL volume, 5% glycerol was added, and the sample was stored at -80ºC. The concentrated sample was bright yellow in color as can be seen in Figure 3.46f.

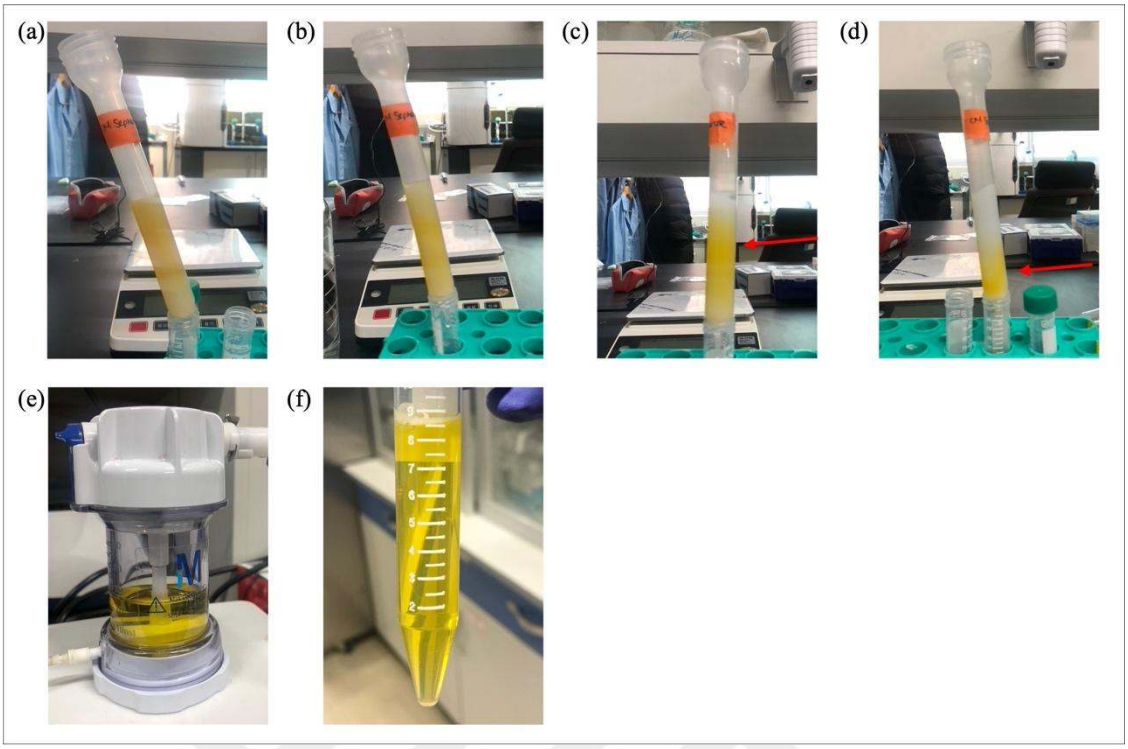


Figure 3.46 Purification of RrFNR using cation chromatography. (a) After sample application. (b) After column wash. (c) After 150mM NaCl Elution. (d) After 250mM NaCl Elution. (e) Concentrating using Amicon stirred-cell concentrator (10kDa). (f) Concentrated pure RrFNR.

For the purification trial of RrFNR from the large-scale production, an SDS-PAGE analysis was performed and can be seen in Figure 3.47. The loaded sample shows a high amount of RrFNR. However, a small amount of the RrFNR loaded to the column did not bind to the resin as the sample application and column wash flowthroughs show a band around 35 kDa. Almost no RrFNR is observed in EF1, suggesting elution starts at 250 mM NaCl. EF3 is almost free from impurities and lysed protein, unlike EF2, which displays many other bands than RrFNR. Since these 2 fractions were combined, the final concentrated sample also shows these impurities. For this reason, the final sample was purified once more.

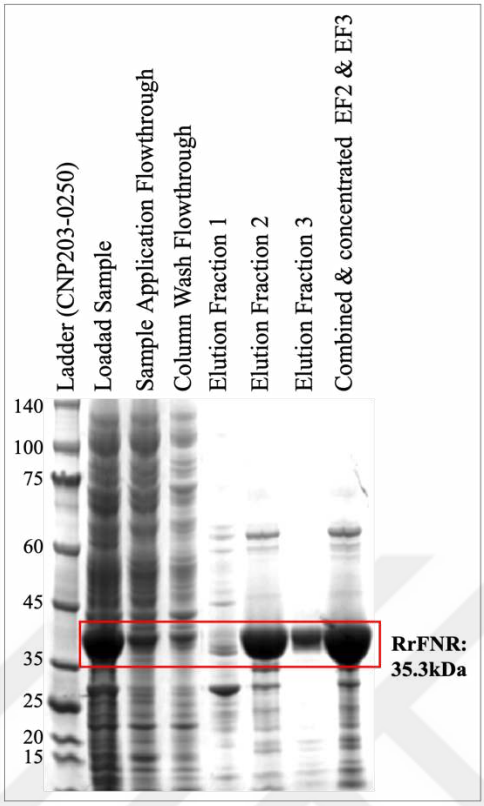


Figure 3.47 SDS-PAGE analysis of RrFNR chromatography fractions.

For the second purification, 5 elution buffer with NaCl concentrations ranging from 150 mM to 350 mM with 50 mM increments were prepared. For each buffer except for the one with 150 mM NaCl, 4 fractions with a volume of 2.5 mL were eluted. When an SDS-PAGE analysis was performed for the purification (Figure 3.48), the fractions with least impurities were chosen to concentrate. 250 mM Fraction 4 and 300 mM Fractions 1-3 were pooled together to a total volume of 10 mL and concentrated as well as buffer exchanged using the Amicon® Stirred Cell concentrator. After adding 5% glycerol, the sample was called RrFNR Lot3 and stored in -80ºC.

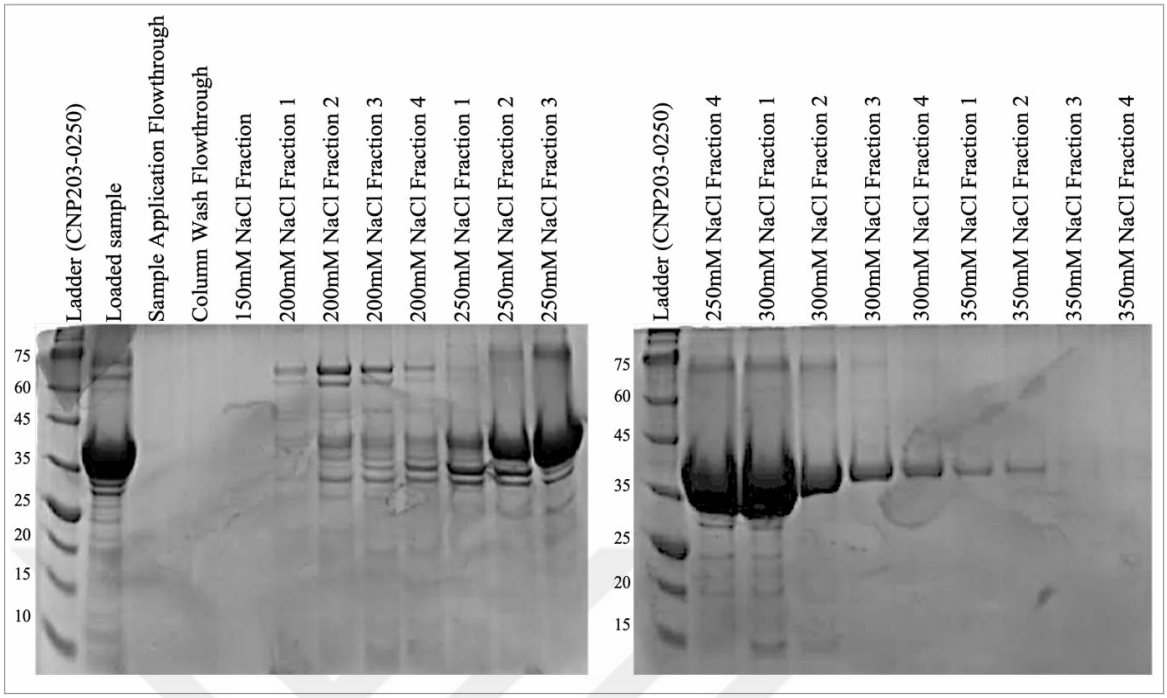


Figure 3.48 Chromatography fractions of the large-scale RrFNR second purification.

It was expected to have a high concentration of RrFNR after the purification from large-scale production in TB medium, and the BSA sample used in drawing the standard curve in Bradford assay was only 0.5 mg/mL in concentration. Therefore, the dilutions prepared of the pure RrFNR to be used in the Bradford assay ranged from 0.125X to 0.008X. A dilution factor versus concentration graph was plotted from the data in Table 3.11. When the x-value is assumed to be 1 in the equation shown in graph, y-value is calculated to be 5.7066. This means 1X concentration of RrFNR Lot 3 is 5.71 mg/mL. To calculate the molarity, concentration was divided to the molecular weight of RrFNR. This calculation gives 161.61 µM.

Table 3.11 Concentration determination of RrFNR dilutions with Bradford assay.

Dilution factor	Concentration (mg/mL)	Standard deviation
0.125	0.67	0.02
0.063	0.38	0.02
0.031	0.15	0.01
0.016	0.04	0.01
0.008	0.01	0.02

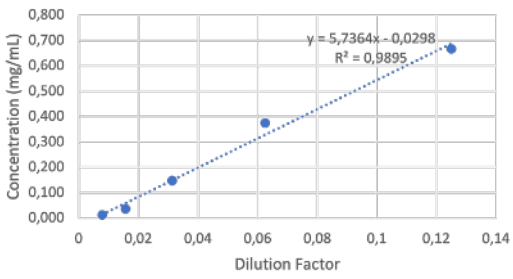


Table 3.12 shows the summary of the calculations above. The final volume of the sample was 8 mL. In total, 45.656 mg protein was obtained from a 300 mL culture. When proportioned to 1000 mL, the purification yield is 152.187 mg RrFNR per liter culture.

Table 3.12 Calculated concentration and molarity of RrFNR Lot3.

Sample	Average Concentration (mg/mL)	Molarity (μM)
RrFNR Lot 3	5.71	161.6

3.3.4 Maturase lysate preparation for in vitro CpI production

In order produce CpI hydrogenase holoenzyme in cell-free, a cell extract that contains the maturase enzymes used in maturation process was required. A cell line with the genes encoding these maturase enzymes, *E. coli* BL21 (DE3) ΔiscR pACYC HydEFG, was grown in initially 50 mL LB medium supplemented with necessary antibiotics, 100 mM MOPS/KOH pH 7.8, 0.5% w/v glucose, and 2 mM ferric ammonium citrate. The culture was grown at 28ºC until OD_{600} reached 0.528 and then it was cycled into an anaerobic glove box. After the anaerobic additions, 10 mM sodium fumarate and 2 mM L-cysteine, the protein production was induced with 0.5 mM IPTG and was left overnight. The harvest OD_{600} was 2.548 (Figure 3.49).

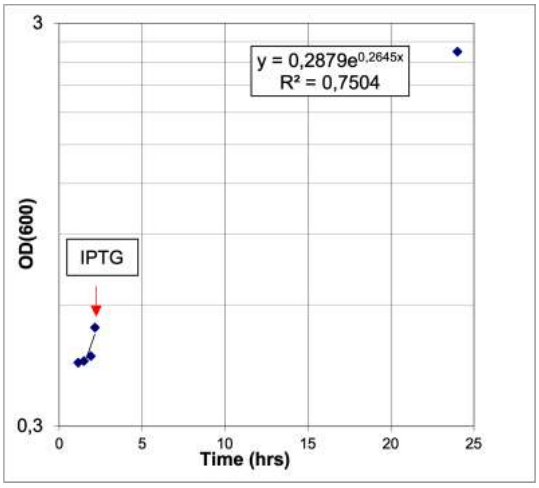


Figure 3.49 Growth curve of small-scale *E. coli* BL21 (DE3) ΔiscR pACYC HydEFG. Specific growth rate (μ): 0.26 hr^{-1}. Red arrow: Induction point with IPTG.

An SDS-PAGE gel was run to determine whether the lysate prepared contained the necessary maturase enzymes. Figure 3.50 shows that these enzymes are present on the lysate when the sample is compared to the before induction sample. HydG is observed around 50 kDa, HydF around 45 kDa, and HydE around 40 kDa as expected. After induction sample shows very faint bands.

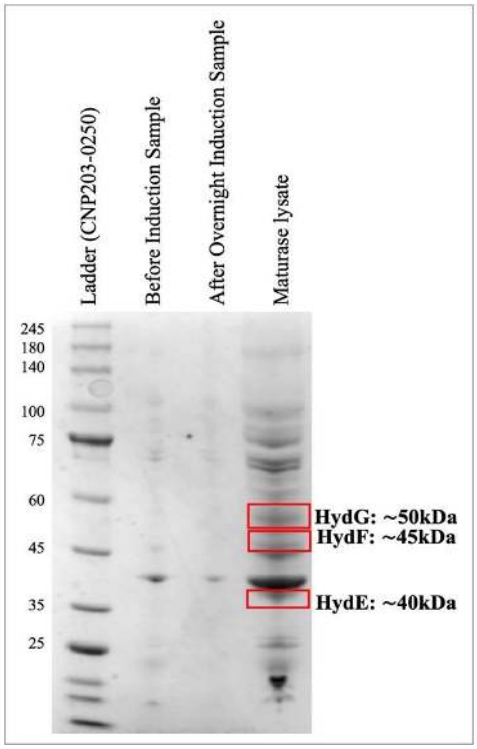


Figure 3.50 SDS-PAGE analysis of maturase lysate preparation. The cell lysate contains the HydE, HydF, and HydG enzymes that take part in the maturation process of CpI apo-hydrogenase into holoenzyme.

After the successful maturase lysate preparation from small-scale production, a large-scale production was done. The *E. coli* BL21 (DE3) ΔiscR pACYC HydEFG culture was induced with 0.5 mM IPTG at OD_{600} 0.576 and harvested at OD_{600} 2.5 (Figure 3.51).

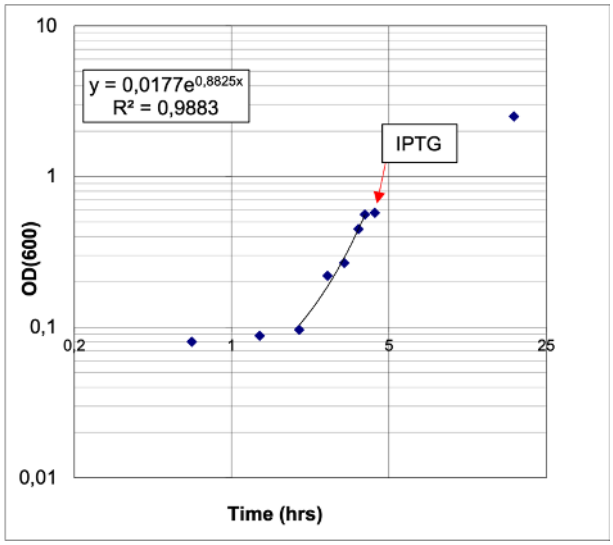


Figure 3.51 Growth curve of large-scale *E. coli* BL21 (DE3) ΔiscR pACYC HydEFG. Specific growth rate (μ): 0.88 hr^{-1}. Red arrow: Induction point with IPTG.

To determine the presence of the maturase enzymes in the prepared lysate, an SDS-PAGE gel was run (Figure 3.52). After induction sample shows more distinct and clear bands for the proteins of interest when compared to the lane of the lysate. All enzymes precipitated in the pellet as can be seen from the thin lanes in the pellet sample.

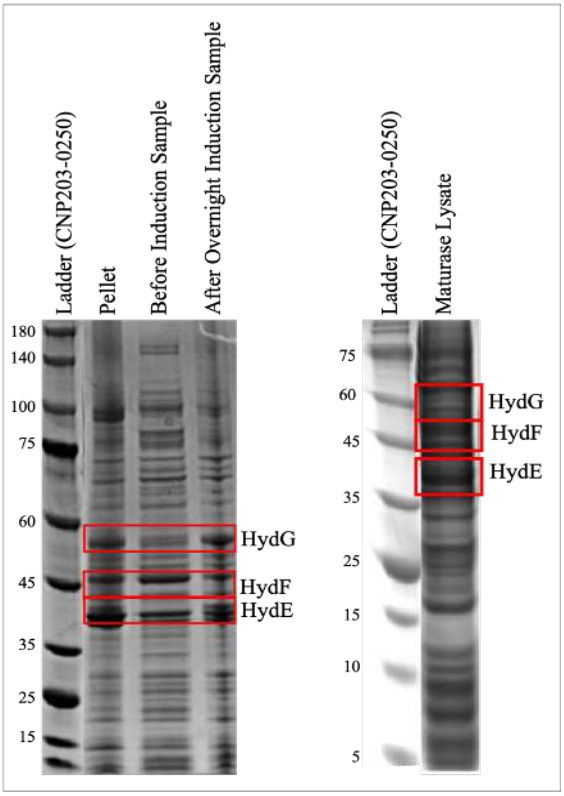


Figure 3.52 SDS-PAGE analysis of large-scale maturase lysate preparation. The cell lysate contains HydE, HydF, and HydG maturases.

Due to the low amount of the maturation enzymes present in the lysates, the enzymes were produced separately, instead of producing them all together. The cultures were transferred inside the anaerobic glovebox when OD_{600} reached ~0.55. After finishing anaerobic additions, the cultures were incubated at room temperature for 15 minutes, then induced with 0.5mM IPTG.

E. coli BL21 (DE3) ΔiscR pACYC HydE was harvested at OD_{600} 2.192, HydF at OD_{600} 1.672, HydG at OD_{600} 2.416. Separate lysates were prepared from the pellets, as described in Section 2.3.4. The SDS-PAGE analysis of the independent maturase growth experiments (Figure 3.53) shows that when the enzymes were produced separately, they display more intense and distinct bands on the gel. As expected, HydF maturase is observed around 45 kDa, HydG around 50 kDa, and HydE around 40 kDa.

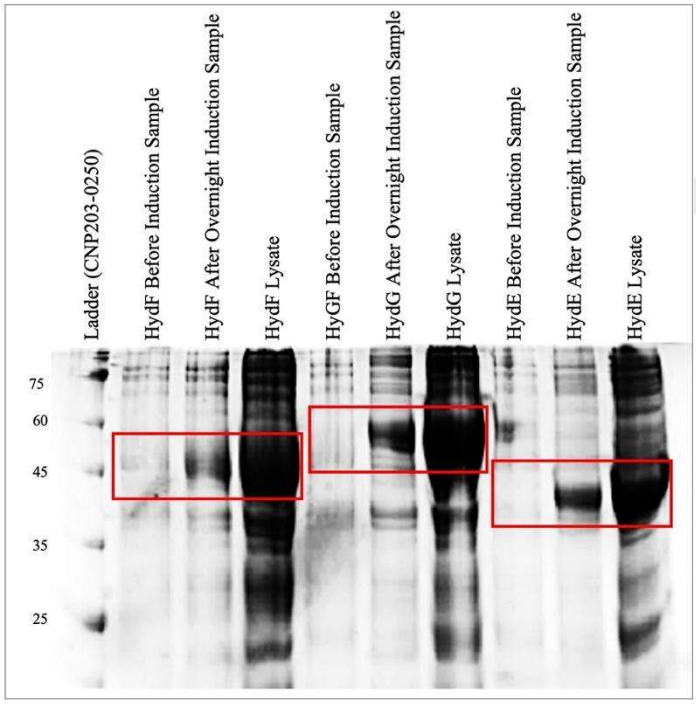


Figure 3.53 SDS-PAGE analysis of different maturase lysate preparations. In this case, each lysate contains a different maturation enzyme (maturase).

3.3.5 Gas Chromatography to detect hydrogen production activity of CpI Hydrogenase

To detect the hydrogen produced by CpI Hydrogenase and the diffused oxygen amount to the reaction vials, gas chromatography (GC) was used. Before the activity assays, calibration curves for hydrogen and oxygen were drawn.

Drawing calibration curves for hydrogen and oxygen

When the oven temperature, column temperature, and carrier gas flow rate were adjusted in the GC system, the retention time of hydrogen was 1 minute and of oxygen was approximately 1.5 min. To draw calibration curves for hydrogen and oxygen, varying percentages of the gas of interest containing vials were prepared as described in Table 2.9 and 2.10. The percentage of hydrogen in the vials ranged from 0 to 1.9 and oxygen from 0 to 21. After preparing the dilutions, 50 µL for hydrogen and 100 µL for oxygen gas mixture was taken from each vial and injected to the GC system. To calculate the volume of hydrogen or oxygen loaded to the system, the following formula was used:

$$Volume\ of H2/O2\ loaded = Total\ volume\ injected\ x\ Percentage\ of\ H2/O2\ in\ the\ vial$$

From this point on, ideal gas law was assumed and the moles of H_2 and O_2 were calculated according to the ideal gas law below:

$$n = P\ x\ V / R\ x\ T$$

where T (temperature) is 293.15K (20ºC), P (pressure) was assumed to be 1 atm and R (ideal gas constant) is 0.0821 L atm mol^{-1} K^{-1}. The resulted mole was converted to nanomoles for hydrogen and µmoles for oxygen. With the measured area under the peak on YL software, a mole versus area curve was drawn for each gas (Figure 3.54).

Table 3.13 shows the calculated volumes of hydrogen in the 50 µL sample injected to the system with the first formula above, calculated mole of the hydrogen in the 50 µL sample with the second formula, average of the measured area replicas, and standard deviation of the areas for each dilution of H_2 according to the equations above.

Table 3.13 H_2 mole calculations using ideal gas law and the measured areas of H_2 dilutions.

Vial	H2%	H2 volume in 50 µL sample (µL)	H2 mole (nmoles)	Measured average area (mV*s)	Standard deviation
1	1.90	0.95	39.47	622.93	19.17
2	1.73	0.86	35.88	585.48	13.49
3	1.52	0.76	31.58	516.70	5.82
4	0.38	0.19	7.89	91.7	1.26
5	0.17	0.09	3.59	58.41	3.71
6	0	0.00	0.00	0.00	0.00

From Table 3.13, a standard curve showing moles versus area under the observed GC peak was drawn, which can be seen in Figure 3.54a. The resulting graph shows a clean, straight line with an R-square value of 0.997, suggesting the measured areas and the moles of the sample are consistent with each other.

A calibration curve for oxygen was drawn with the measurements of O_2 dilutions taken by Kim Seongwon of Dr. Jamin Koo Lab, Hongik University (Figure 3.54b). Table 3.14 shows the calculated volume of oxygen in the 100 µL sample injected to the GC system, mole of oxygen injected, average of the area of the obtained peaks and standard deviation of the replicas.

Table 3.14 O_2 mole calculations using ideal gas law and the measured areas of O_2 dilutions.

Vial	O2 %	O2 volume in 100 µL sample (µL)	O2 moles (µmoles)	Measured average area (mV*s)	Standard deviation
1	21	21	0.87	202.92	0.08
2	16.8	16.8	0.7	164.19	0.88
3	4	4	0.17	39.84	0.40
4	2	2	0.08	21.99	1.11
5	1	1	0.04	11.83	0.00
6	0.5	0.5	0.02	7.7	0.00

From Table 3.13, a calibration curve was drawn with moles of oxygen at the x-axis and area at the y-axis. Even though both curves display good R-square values, the biggest issue with them was a lack of dilution between 1.52% and 0.38% for hydrogen and between 16.8% and 4% for oxygen. It was assumed that the pressure inside the 2 mL GC vials would change if there was an addition of more than 0.5 mL gas in any case. Also, if any large-volume gas were to be removed from a vial, the pressure inside could drop, leading to diffusion of air towards inside. To avoid this issue, no more than 0.5 mL of gas was injected to any vial and no gas was ejected except for GC measurements.

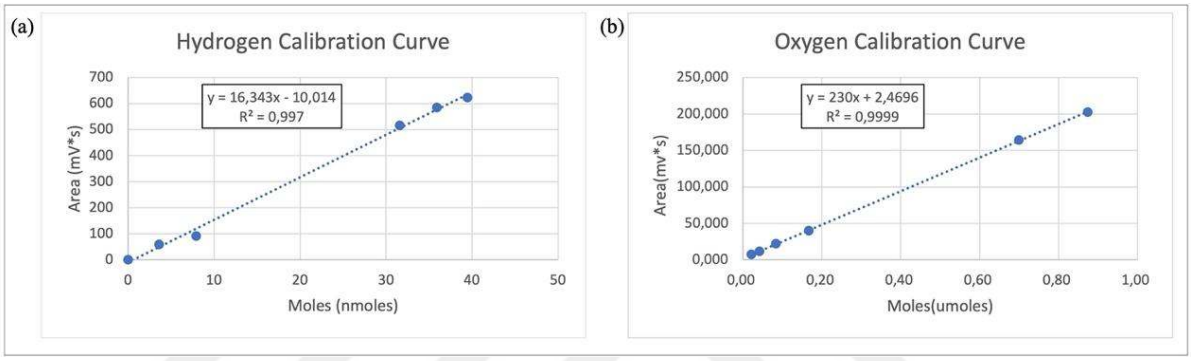


Figure 3.54 Gas chromatography (GC) calibration curves. (a) Hydrogen calibration curve. (b) Oxygen calibration curve.

NADPH and DTH assays to measure hydrogen production activity of in vivo produced CpI

Once the calibration curves were established, NADPH and DTH assays were performed to test the hydrogen production activity of the purified hydrogenase samples. Figure 3.55a and 3.55c show that in both assays, hydrogen production was observed. For the DTH assay, the rate of the hydrogen production was calculated for the first 30 minutes of the reaction as it was known the assay does not last more than that. The rate was calculated to be 36.11 nmole/min. For the NADPH assay, it was 3.73 nmole/min for 60 minutes.

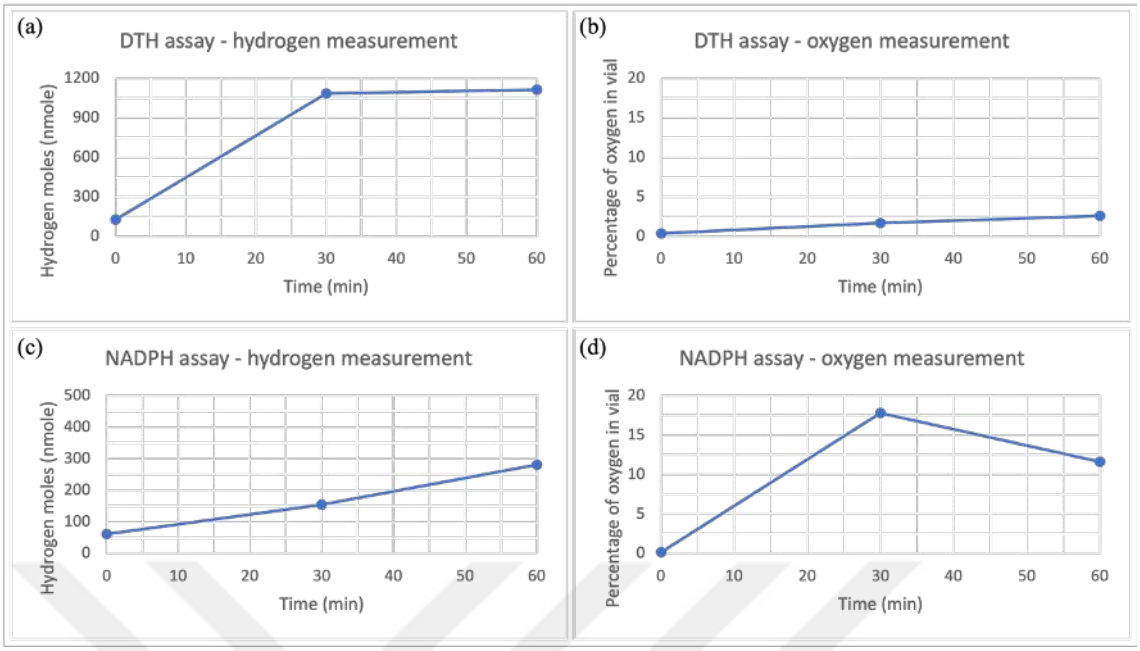


Figure 3.55 Gas chromatography results from DTH and NADPH assays. (a) DTH assay hydrogen production curve. (b) DTH assay oxygen diffusion curve. (c) NADPH assay hydrogen production curve. (d) NADPH assay oxygen diffusion curve.

The NADPH assay above was performed with small-scale RrFNR produced previously. When the lot was switched to large-scale RrFNR, the assay did not work after several trials, so it was assumed that the new batch was not actively working. However, an air sample was injected into the system as a control experiment, and surprisingly no peak was observed even though air has 21% oxygen. This result led to the suspicion of something wrong with the GC setup. The oven and column temperatures were tried to be adjusted, and the nitrogen source was checked for its pressure, but still, no peak was observed on the GC system for oxygen when air was sampled. Then a hydrogen-containing vial taken from an anaerobic glovebox containing 2% hydrogen was also sampled, but there was still no peak in the system. The septa of the GC system were also changed to see if it was causing leakage, thus leading to no peak, but the problem still continued. Another GC syringe was used to inject the sample into the GC system as a last resort. Interestingly, both oxygen and hydrogen peaks appeared on the chromatogram with the new syringe. After investigating the syringes, it was noticed that the small rubber particles from the septa caused clogging of the syringe needle when the septa were punctured while injecting the sample into the system. From this point, extra care was taken to the syringes to ensure they were not clogged before sampling.

In the following trials it was noticed that NADPH assay efficiency was dropped as the assay started to not last for 1 hour but only 30 minutes. It was suspected that the CpI hydrogenase, which was stored at room temperature inside the anaerobic glovebox, got inactivated due to the increase of oxygen pressure inside the glovebox above 100 ppm more than a few times. A DTH assay was performed and the rate of the hydrogen production of CpI hydrogenase was compared to the rate of the first assay after CpI purification. In the first 30 minutes, the hydrogen production rate was 8.574 nmoles/min for the first 30 minutes. While previously it was 36.11 nmoles/min. The rate dropped to 1/4th of the original rate. This suggests that the *in vivo* CpI indeed got degraded over time as the oxygen pressure in the glovebox increased over 100ppm several times. An assumption was made that 25% of the CpI hydrogenase was inactivated based on the results. Since the inactivation of hydrogenase is irreversible, on the following experiments, 4-fold more enzyme was added to the assays to compensate for the lost ones.

An oxygen tolerance assay was also performed for the wild-type CpI hydrogenase. As shown in Figure 3.56a, after introducing 5% oxygen to the reaction vial after 10 minutes, there was no hydrogen production was observed after the addition of 5% oxygen.

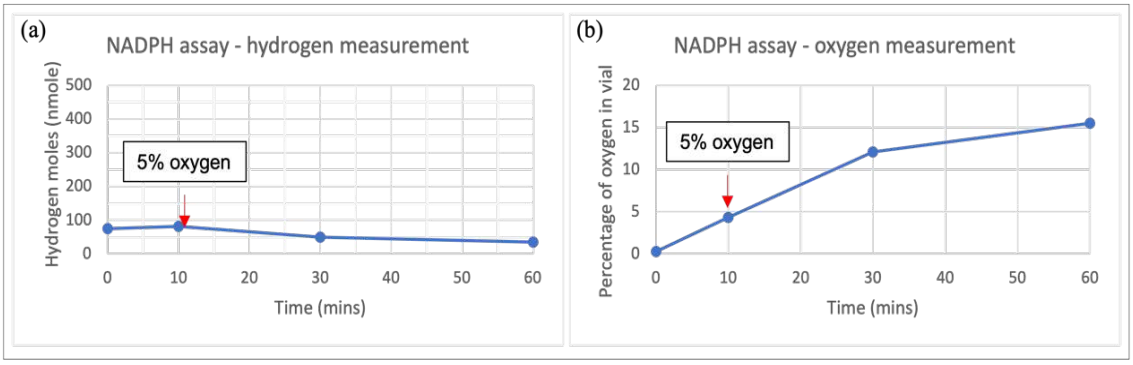


Figure 3.56 Gas chromatography NAPDH oxygen addition assay results. (a) NADPH assay hydrogen production curve. (b) NADPH assay oxygen diffusion curve. Red arrow: 5% oxygen injection point to the reaction vial.

3.3.6 Cell-free production of wild-type CpI hydrogenase

Figure 3.57 shows the SDS-PAGE analysis of the comparison of ΔiscR (from *E. coli* BL21 (DE3) ΔiscR cell line) and maturase (from large-scale *E. coli* BL21 (DE3) ΔiscR pACYC HydEFG production) lysates in terms of wild-type sfGFP and CpI Hydrogenase production. sfGFP-ΔiscR sample displays a distinct band for sfGFP (26

kDa). However, when it was tried to produce with the maturase lysate, there was no observed band for sfGFP around 26 kDa. CpI hydrogenase on the other hand, did not show a visible band on the gel when it was tried to be produced either with ΔiscR or maturase lysate.

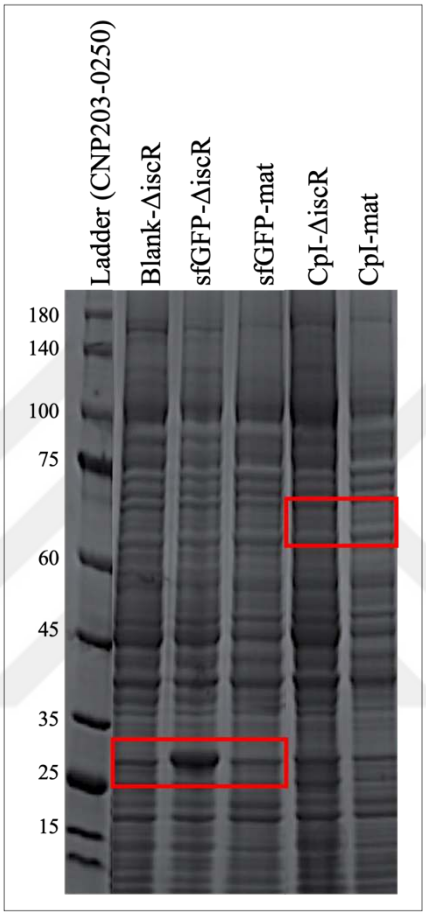


Figure 3.57 SDS-PAGE analysis of CFPS sfGFP and CpI Hydrogenase production, comparing ΔiscR and maturase lysates.

As a quantitative assay, the fluorescence intensities of the CFPS produced sfGFPs in different extracts were measured and concentrations were calculated from a standard curve plotted with a pure sfGFP sample. Figure 3.58 shows that 0.593 mg/mL sfGFP was produced with ΔiscR lysate and 0.176 mg/mL sfGFP was produced with the maturase lysate. In terms of protein production efficiency, maturase lysate is only 29.7% of the ΔiscR lysate.

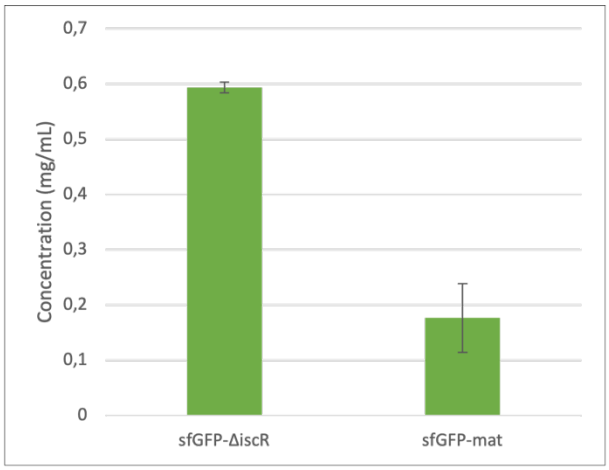


Figure 3.58 Comparison of ΔiscR and maturase lysates in terms of soluble sfGFP production efficiency. Error bars represent the difference in soluble protein concentration between reaction replicas.

When cell-free hydrogenase production was tried and an NADPH assay was performed to measure hydrogen productivity of the produced enzyme, there was no observed peak for hydrogen on the GC chromatogram, meaning the production trial was unsuccessful. The reaction was repeated with another hydrogenase plasmid, but the outcome was still the same. At first, the positive control, CFPS sfGFP production also did not work. Therefore, it was suspected that there was something wrong with the cell-free setup and a troubleshooting was performed with ΔiscR cell extract as it was more efficient than the maturase lysate in terms of sfGFP production. By testing different lots of some of the key reagents, PEP and the gene of interest in its vector, it was observed that the PEP lot used in the previous trials lowered the yield of the reactions significantly, possibly due to high freeze-thaw cycles. pY71 sfGFP plasmid was also suffering from the same condition. With a fresh PEP stock and a new batch of sfGFP plasmid, the yield increased significantly (Figure 3.59).

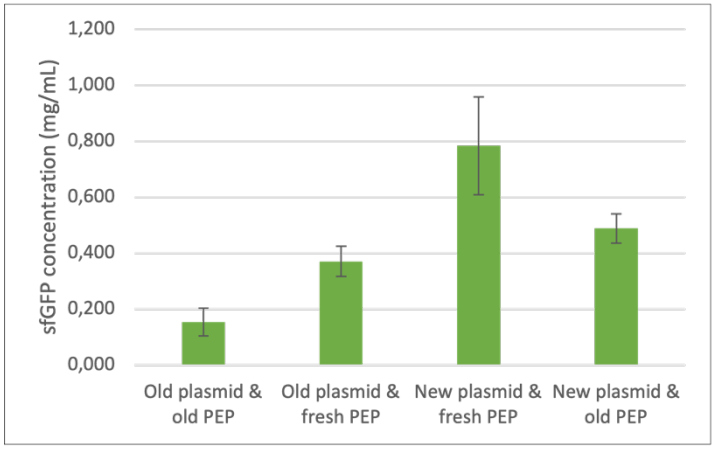


Figure 3.59 CFPS troubleshooting comparing old and new pY71 sfGFP plasmids and PEP stocks in terms of soluble sfGFP production efficiency. Error bars represent the difference in soluble protein concentration between reaction replicates.

With the new working setup, the wild-type hydrogenase enzyme was tried to be produced once more. However, there was still no peak on the GC chromatogram, but the positive control was working fine and the positive controls, which are air sample, and 2% hydrogen anaerobic glovebox sample were working fine. Therefore, this time it was suspected that the maturase lysate that was used in CFPS process was unable to produce active hydrogenase. This could be due to the insufficient maturase enzymes in the lysate since they were produced all together. Therefore, all the maturases were grown in different cultures and the maturase lysate containing all 3 enzyme was prepared by mixing the lysates in the ratio shown in Table 3.15.

Table 3.15 The ratios of the lysates required to prepare the final maturase lysate.

Lysate	Ratio	Volume (µL)
HydF	1	25
HydG	4	100
HydE	1	25
ΔiscR	1	25

The new maturase lysate was used to produce the wild-type hydrogenase, however still there was no observed hydrogen peak on the chromatogram, indicating there was no active hydrogenase present. After the experiment, while putting the SAM into the -20C freezer, it was noticed that its expiry date was way passed. This was probably the reason

why the CFPS production experiments did not work as SAM is crucial to the hydrogenase formation because it has a key role in iron-sulfur cluster formation. From this point on, our project partner, Dr. Jamin Koo continued to the *in vitro* hydrogenase production experiments as well as activity assays.

Chapter 4:
DISCUSSION

Preparations for wild-type and mutant CpI Hydrogenase Production in CFPS

In the first part of the thesis, the aim was to produce and purify all enzymes, templates, and plasmids required for noncanonical amino acid insertion to CpI holoenzyme with CFPS. Towards this goal, 5 the aminoacyl tRNA synthetase genes (CNFRS, pAFRS, pIFRS, OMeYRS, HRS) were cloned into pY71 vector with optimizations made in the process. Before this, an N-terminal NdeI restriction site, C-terminal SalI restriction site and 6xHis affinity marker sequence were genetically incorporated into the synthetase genes by PCR. After the successful restriction-ligation cloning, confirmed by Macrogen sequence analysis, it was noticed that there was an extra methionine codon at the translation start site due to a design mistake in the primers used in the insertion of NdeI restriction site. Therefore, this codon was removed from the gene sequence via SDM.

M. jannaschii-derived synthetases (CNFRS, pAFRS, pIFRS, OMeYRS, pPaFRS, pAMFRS) with C-terminal 6xHis-tags were produced with the exact procedure, as they are closely related. Protein production was conducted under the T7 promoter with BL21(DE3)pLysS cell line at 37^{o}C, followed by an IMAC purification with imidazole being the competitive reagent of the elution. The purified proteins were dialyzed against a 10 mM potassium phosphate buffer (pH 7.4) with 0.05% (v/v) Tween-20. The only exception was pAMFRS, which was dialyzed against a PBS buffer (pH 7.1) as it formed aggregates in the potassium phosphate buffer at the said pH. The reason for this was discussed as the isoelectric point of pAMFRS , 7.76, being too close to the pH of the said buffer, promoting interprotein protein-hydrophobic interactions and thus leading to the formation of aggregates.

The optimized o-tRNA template used with the *M.jannaschii*-derived aaRSs to incorporate ncAAs was synthesized by a large-scale PCR of a plasmid from the lab stocks and the resulting PCR product was purified using ethanol precipitation. Another optimized o-tRNA template derived from *M.barkeri* that is compatible with *Mb*aaRSs was built using PCR assembly. After confirmation of production of the desired o-tRNA

templates by agarose gel electrophoresis, the templates were amplified via PCR. To ensure high promoter strength, a hammerhead ribozyme sequence was also part of the template sequences.

Purified aaRS and o-tRNA components were tested for their efficiency in inserting ncAAs into sfGFP151TAG. Interestingly sfGFP151.CNF variants showed no fluorescence activity. However, when CNF was inserted into 23TAG or 39TAG locations, fluorescence measurement could be done. The reason why only CNFRS enzymes were low in activity with 151TAG could be due to an incompatibility of the CNF ncAA with the area surrounding sfGFP151 location, causing misfolding, or insolubility. Moreover, the enzyme activity could simply be lower compared to the other synthetase enzymes. The o-tRNA lot could also affect the fluorescence results. Regardless, in this section, *all M.jannaschii*-derived aaRSs that will be used in the project were synthesized in the active form.

M.barkeri-derived aminoacyl tRNA synthetase HRS production

M.barkeri-derived HRS, after the successful restriction-ligation cloning, was initially produced and purified with the traditional *in vivo* production and IMAC purification method used to obtain pure *M.jannaschii*-derived aaRSs. However, even though the full-size protein was produced, it could not be purified as the enzyme formed insoluble protein aggregates and accumulated in the pellet. Hence, to combat aggregation and increase the solubility of the enzyme, *in vivo* expression of the enzyme was studied, with decreased temperatures, decreased IPTG concentrations, and additions of 2% ethanol and/or 15% glycerol. None of the tried methods increased the soluble protein amount. Therefore, *in vivo* production was abandoned and HRS was synthesized with CFPS.

Chaperones and changes in the redox state of the environment can enhance protein solubility. While the protein solubility increased with 4 μM Skp addition and also with 4mM GSSG & 1mM GSH additions, the produced protein was not intact. A 35 kDa truncated protein was produced in the CFPS reactions. Suspecting protease activity, protease inhibitor cocktails were inserted to the CFPS system with varying concentrations. While the addition of a protease inhibitor cocktail increased the intact

protein amount, the tested concentrations of the inhibitor did not cause any noticeable change. Therefore, the low concentration, 1X, was used in the follow-up experiments. However, the protease inhibitor alone was not sufficient to suppress the proteolytic activity as combining 1X inhibitor cocktail with 4 μM Skp still resulted in the proteolyzed product.

It was suspected that the OmpT protease inside the opt. cell extract caused the proteolysis of *in vitro* produced HRS. To test this hypothesis, cell extracts were prepared from the BL21 cell line which lacks OmpT and Lon proteases, and the ARG1 cell line which lacks the OmpT protease. When these extracts were used in CFPS HRS production instead of opt. extract, the intense band around 35 kDa was not observed, proving it was indeed OmpT that resulted in the main proteolysis product of HRS. Compared to ARG1extract, BL21 was more efficient in terms of soluble protein production and wild-type sfGFP production.

HRS was then produced in a 3 mL large-scale CFPS reaction that harbors 4 μM Skp, 1X protease inhibitor cocktail with BL21 cell extract. When purified using His-tag affinity chromatography, intact soluble protein in the loaded sample was observed in the western-blot analysis. However, there was no protein band in the expected or any other area in the elution sample. One hypothesis to explain this is that the concentration of the intact HRS in the loaded sample was much lower than the binding capacity of the column. As the purification of the cell-free produced HRS failed, other strategies were searched.

With the discoveries from previous experiments in mind, the target sites of OmpT protease on HRS were investigated, and 9 possible OmpT target sites were determined on the enzyme. From these sites, K104/K105 site was selected as the most potent site as the lysis from this cut site results in s 36 kDa truncated protein. To make HRS an OmpT-resistant protein, primers were designed to produce HRS.K105A variant via SDM. Alanine was chosen as the substrate as it is the smallest amino acid and will not cause any harm to the protein function. The primers were ordered for the future SDM and CFPS HRS.K105A variant production trials with the opt. extract to test whether the engineered enzyme is OmpT-resistant as planned.

Finally, a recent publication suggested that a solubility tag called SmbP could assist in the proper folding of an *M.barkeri* derived aaRS when it is genetically fused to the gene of the protein. With this finding, a fusion protein was designed that contains the SmbP sequence, enterokinase cleave site to remove the tag after production and purification of the enzyme and HRS gene. The designed fusion protein was ordered from Twist Biosciences for future *in vivo* production trial of the enzyme.

Production, Purification, and Analysis of CpI Holo-Hydrogenase

In the last part of the thesis, CpI holoenzyme was produced, and after its purification, hydrogen activity assessment was done with two separate assays. The plasmid harboring the CpI gene was transformed into BL21(DE3)*ΔiscR* pACYC HydEFG cell line, which contains the maturation enzymes that are required in CpI holoenzyme formation. *In vivo* protein production was performed under T7 promoter with 0.5 mM IPTG anaerobically. The produced CpI holoenzyme was then purified with StrepTactin affinity chromatography as the CpI has a C-terminal StrepII affinity tag. The purified protein was concentrated if necessary and stored inside the anaerobic glove box.

For *in vitro* production of CpI holoenzyme, a lysate with the maturase enzymes is required. Therefore, BL21(DE3)*ΔiscR* pACYC HydEFG cell line was grown aerobically at 28^{o}C until the induction point with the additions necessary for maturase synthesis before being transferred to an anaerobic glovebox. After anaerobic additions, 0.5 mM IPTG was used to induce maturase production under the T7 promoter. The culture was harvested after overnight growth, and the cell extract was prepared using chemical lysis. Because the maturase amount in the obtained lysate was small, and no CFPS-produced CpI could be synthesized with the prepared lysate, the maturation enzymes were produced in separate cultures, and the cell lysates made with them were combined in a specific ratio for CFPS.

RrFNR, one of the NADPH-driven assay components, was produced with the BL21(DE3) cell line. The culture was grown at 30^{o}C, and protein production continued overnight after induction with 0.5 mM IPTG under the T7 promoter. Cation exchange chromatography was used for the purification of RrFNR. In the initial purification trial, the protein did not bind to the column at all, as the distinct yellow color of the protein

simply flowed through the column. It was argued that since the pH of the lysate was high due to the lysis buffer (pH 8), the ionic strength of the buffer was high. Indeed, when the pH of the loaded sample was adjusted to 5.5 in the following trial, the protein was easily purified. The pure sample was then concentrated due to its high volume. The yield of the small-scale protein production (50 mL, LB medium), was 11.6 mg/L. In the large-scale production (300 mL, TB medium), the yield was calculated as 152.19 mg/L, which is a significantly high result.

After producing and preparing all necessary components for the NADPH- and DTH-driven assays, the hydrogen production activity of the anaerobically produced and purified CpI holoenzyme was measured. In the NADPH-driven assay, the rate of the reaction was calculated as 3.73 nmole/min in 60 minutes. The rate of the reaction was a little below the expected 10nmoles/min rate but there was a high amount of oxygen diffusion inside the vial, possibly inactivating CpI. In the DTH-driven assay, the rate of the reaction was 32.03 nmole/min for 30 minutes. It was expected to see a higher reaction rate in the DTH-driven assay due to the high reducing power of dithionite. The blank sample not shown here did not display any hydrogen peak on the gas chromatography as expected.

The oxygen tolerance of the wild-type CpI holoenzyme was also assessed with the NADPH-driven assay by the addition of 5% (v/v) oxygen 10 minutes after the reaction started. As expected, no increase in the hydrogen levels was observed after oxygen exposure of CpI. With these experiments, it was proven that the anaerobically produced CpI is indeed active and able to form hydrogen by combining electrons with protons.

In the last part of this section, CpI holoenzyme was also tried to be synthesized *in vitro*, with the prepared maturase lysates. However, it could not be synthesized no matter how many components were tried to be improved. This suggests that CpI was either not being synthesized or synthesized as the apoenzyme that has no hydrogen production activity. Later it was realized that the problem was one of the crucial chemicals, SAM, being expired. SAM is a crucial chemical in the maturation of CpI. Therefore, its absence must have prevented CpI enzyme from maturing into holoenzyme.

Chapter 5:
CONCLUSION

Preparations for wild-type and mutant CpI Hydrogenase Production in CFPS

To incorporate ncAAs into proteins, an aminoacyl-tRNA synthetase and its orthogonal tRNA component are required. M. jannaschii-derived aaRSs can insert tyrosine and phenylalanine analogs. All the aaRS enzymes planned to be synthesized in the scope of this thesis were indeed produced *in vivo* and purified, with HRS being the only exception. However, important progress was also made with HRS as well. The optimized o-tRNA templates were also produced.

Except for pAMFRS, the production and purification of the aaRS enzymes were rather simple as one protocol was applied to all of them. For pAMFRS, a different dialysis buffer was used to prevent the enzyme from forming aggregates.

M.barkeri-derived aminoacyl tRNA synthetase HRS production

HRS is an *M.barkeri*-derived aminoacyl-tRNA synthetase that is engineered for the incorporation of histidine analogs into proteins. While *M.barkeri*-derived synthetase/tRNA pairs can incorporate more than 100 ncAAs including pyrrolysine, phenyl, histidine, or tryptophan analogs efficiently and are orthogonal in both bacterial and eukaryotic systems, few attempts were made to purify the aaRS enzymes. The enzyme has no problem with solubility when it is synthesized and used *in vivo*, however its solubility significantly decreases in studies conducted *in vitro*. It was reported that the hydrophobic N-terminal domain, responsible of recognition, binding, and charging of its orthogonal tRNA component, is related to the low solubility of the enzyme.

Unlike the straightforward production and purification of *M.jannaschii*-derived aaRS enzymes, production trials of HRS were more complicated. In the thesis, many conditions and variables were tested both *in vivo* and *in vitro* to obtain purified HRS enzyme. As the *M.barkeri*-derived aaRS/tRNA pairs have so much potential for genetic code expansion, the aim was to use the pure protein for ncAA incorporation in CFPS systems.

One of the key findings of the HRS production experiments is that the enzyme is a target of the OmpT protease found in 759.T7.Opt cell extract. If in the future, an OmpT-resistant HRS can be synthesized, by mutating the target site of the protease, it would be possible to produce the enzyme in CFPS and purify it. Another important point is that a solubility tag, such as SmbP when fused with the *M. barkeri*-derived synthetase genes, increases the chance to produce the protein soluble and purify it accordingly.

Production, Purification, and Analysis of CpI Holo-Hydrogenase

CpI holoenzyme is a difficult protein to work with due to its highly oxygen-sensitive nature. However, it is one of the fastest known hydrogenases in terms of H_2 production, and if the photosynthetic hydrogen production design with the oxygen-tolerant CpI variants comes to fruition, a sustainable, high-density energy source can be produced without contributing to the carbon footprint.

In the thesis, significant progress was made toward the production of the planned CpI variants in CFPS systems. CpI holoenzyme was produced anaerobically *in vivo* and purified. RrFNR, used in the NADPH-driven hydrogen production pathway was also produced and purified. Then the pure CpI holoenzyme was proved to be active with NADPH- and DTH-driven hydrogen production assays. A gas chromatography protocol was established, and the calibration curves for H_2 and O_2 were drawn. For the production of cell-free CpI, the cell extracts that contain the maturation enzymes were prepared. Even though CpI could not be produced in the scope of this thesis, the reason behind it was determined, that SAM used in the CFPS reactions was expired. Every reagent, enzyme, lysate, and procedure necessary for the cell-free production and activity assessment of CpI holoenzyme was prepared.

REFERENCES

[1] Koo, J., & Swartz, J. R. (2018). System analysis and improved [FeFe] hydrogenase O2 tolerance suggest feasibility for photosynthetic H2 production. *Metabolic Engineering*, *49*, 21–27. https://doi.org/10.1016/j.ymben.2018.04.024

[2] Chaubey, R., Sahu, S., James, O. O., & Maity, S. (2013). A review on development of industrial processes and emerging techniques for production of hydrogen from renewable and sustainable sources. *Renewable and Sustainable Energy Reviews*, *23*, 443–462. https://doi.org/10.1016/j.rser.2013.02.019

[3] Frey, M. (2002). Hydrogenases: Hydrogen-Activating Enzymes. *ChemBioChem*, *3*(2-3), 153–160. https://doi.org/3.0.co;2-b">10.1002/1439-7633(20020301)3:2/3<153::aid-cbic153>3.0.co;2-b

[4] Cornish, A. J., Gärtner, K., Yang, H., Peters, J. W., & Hegg, E. L. (2011). Mechanism of Proton Transfer in [FeFe]-Hydrogenase from Clostridium pasteurianum. *Journal of Biological Chemistry*, *286*(44), 38341–38347. https://doi.org/10.1074/jbc.m111.254664

[5] Adams, M. W. W. (1990). The structure and mechanism of iron-hydrogenases. *Biochimica et Biophysica Acta (BBA) - Bioenergetics*, *1020*(2), 115–145. https://doi.org/10.1016/0005-2728(90)90044-5

[6] Therien, J. B., Artz, J. H., Poudel, S., Hamilton, T. L., Liu, Z., Noone, S. M., Adams, M. W. W., King, P. W., Bryant, D. A., Boyd, E. S., & Peters, J. W. (2017). The Physiological Functions and Structural Determinants of Catalytic Bias in the [FeFe]-Hydrogenases CpI and CpII of Clostridium pasteurianum Strain W5. *Frontiers in Microbiology*, *8*. https://doi.org/10.3389/fmicb.2017.01305

[7] Adams, M. W., & Mortenson, L. E. (1984). The physical and catalytic properties of hydrogenase II of Clostridium pasteurianum. A comparison with hydrogenase I. *The Journal of Biological Chemistry*, *259*(11), 7045–7055. https://pubmed.ncbi.nlm.nih.gov/6327705/

[8] Koo, J., Shiigi, S., Rohovie, M., Mehta, K., & Swartz, J. R. (2016). Characterization of [FeFe] Hydrogenase O2 Sensitivity Using a New, Physiological Approach. *Journal of Biological Chemistry*, *291*(41), 21563–21570. https://doi.org/10.1074/jbc.m116.737122

[9] Broderick, J. B., Byer, A. S., Duschene, K. S., Duffus, B. R., Betz, J. N., Shepard, E. M., & Peters, J. W. (2014). H-Cluster assembly during maturation of the [FeFe]-hydrogenase. *JBIC Journal of Biological Inorganic Chemistry*, *19*(6), 747–757. https://doi.org/10.1007/s00775-014-1168-8

[10] Kuchenreuther, J. M., Britt, R. D., & Swartz, J. R. (2012). New Insights into [FeFe] Hydrogenase Activation and Maturase Function. *PLoS ONE*, *7*(9), e45850. https://doi.org/10.1371/journal.pone.0045850

[11] Stripp, S. T., Goldet, G., Brandmayr, C., Sanganas, O., Vincent, K. A., Haumann, M., Armstrong, F. A., & Happe, T. (2009). How oxygen attacks [FeFe] hydrogenases from photosynthetic organisms. *Proceedings of the National Academy of Sciences*, *106*(41), 17331–17336. https://doi.org/10.1073/pnas.0905343106

[12] Goldet, G., Brandmayr, C., Stripp, S. T., Happe, T., Cavazza, C., Fontecilla-Camps, J. C., & Armstrong, F. A. (2009). Electrochemical Kinetic Investigations of the Reactions of [FeFe]-Hydrogenases with Carbon Monoxide

and Oxygen: Comparing the Importance of Gas Tunnels and Active-Site Electronic/Redox Effects. *Journal of the American Chemical Society*, *131*(41), 14979–14989. https://doi.org/10.1021/ja905388j

[13] Swanson, K. D., Ratzloff, M. W., Mulder, D. W., Artz, J. H., Ghose, S., Hoffman, A., White, S., Zadvornyy, O. A., Broderick, J. B., Bothner, B., King, P. W., & Peters, J. W. (2015). [FeFe]-Hydrogenase Oxygen Inactivation Is Initiated at the H Cluster 2Fe Subcluster. *Journal of the American Chemical Society*, *137*(5), 1809–1816. https://doi.org/10.1021/ja510169s

[14] Heidary, N., Utesch, T., Zerball, M., Horch, M., Millo, D., Fritsch, J., Lenz, O., von Klitzing, R., Hildebrandt, P., Fischer, A., Mroginski, M. A., & Zebger, I. (2015). Orientation-Controlled Electrocatalytic Efficiency of an Adsorbed Oxygen-Tolerant Hydrogenase. *PloS One*, *10*(11), e0143101. https://doi.org/10.1371/journal.pone.0143101

[15] Fritsch, J., Scheerer, P., Frielingsdorf, S., Kroschinsky, S., Friedrich, B., Lenz, O., & Spahn, C. M. T. (2011). The crystal structure of an oxygen-tolerant hydrogenase uncovers a novel iron-sulphur centre. *Nature*, *479*(7372), 249–252. https://doi.org/10.1038/nature10505

[16] Liu, C. C., & Schultz, P. G. (2010). Adding New Chemistries to the Genetic Code. *Annual Review of Biochemistry*, *79*(1), 413–444. https://doi.org/10.1146/annurev.biochem.052308.105824

[17] Ambrogelly, A., Palioura, S., & Söll, D. (2006). Natural expansion of the genetic code. *Nature Chemical Biology*, *3*(1), 29–35. https://doi.org/10.1038/nchembio847

[18] Johnson, J. A., Lu, Y. Y., Van Deventer, J. A., & Tirrell, D. A. (2010). Residue-specific incorporation of non-canonical amino acids into proteins: recent developments and applications. *Current Opinion in Chemical Biology*, *14*(6), 774–780. https://doi.org/10.1016/j.cbpa.2010.09.013

[19] Fang, K. Y., Lieblich, S. A., & Tirrell, D. A. (2018). Incorporation of Non-Canonical Amino Acids into Proteins by Global Reassignment of Sense Codons. *Methods in Molecular Biology*, 173–186. https://doi.org/10.1007/978-1-4939-7893-9_13

[20] Koch, N. G. (2021). Efficient Unnatural Protein Production by Pyrrolysyl-tRNA Synthetase With Genetically Fused Solubility Tags. Frontiers in Bioengineering and Biotechnology, 9. https://doi.org/10.3389/fbioe.2021.807438

[21] Kavran, J. M., Gundllapalli, S., O'Donoghue, P., Englert, M., Söll, D., & Steitz, T. A. (2007). Structure of pyrrolysyl-tRNA synthetase, an archaeal enzyme for genetic code innovation. *Proceedings of the National Academy of Sciences*, *104*(27), 11268–11273. https://doi.org/10.1073/pnas.0704769104

[22] Herring, S., Ambrogelly, A., Gundllapalli, S., O'Donoghue, P., Polycarpo, C. R., & Söll, D. (2007). The amino-terminal domain of pyrrolysyl-tRNA synthetase is dispensable in vitro but required for in vivo activity. *FEBS Letters*, *581*(17), 3197–3203. https://doi.org/10.1016/j.febslet.2007.06.004

[23] Thomas, J. G., & Baneyx, F. (1997). Divergent Effects of Chaperone Overexpression and Ethanol Supplementation on Inclusion Body Formation in RecombinantEscherichia coli. *Protein Expression and Purification*, *11*(3), 289–296. https://doi.org/10.1006/prep.1997.0796

[24] Sawano, H., Koumoto, Y., Ohta, K., Sasaki, Y., Segawa, S., & Tachibana, H. (1992). Efficient in vitro folding of the three-disulfide derivatives of hen

lysozyme in the presence of glycerol. *FEBS Letters*, *303*(1), 11–14. https://doi.org/10.1016/0014-5793(92)80466-t

[25] Schiffrin, B., Calabrese, A. N., Devine, P. W. A., Harris, S. A., Ashcroft, A. E., Brockwell, D. J., & Radford, S. E. (2016). Skp is a multivalent chaperone of outer-membrane proteins. *Nature Structural & Molecular Biology*, *23*(9), 786–793. https://doi.org/10.1038/nsmb.3266

[26] Wang, J. D., Michelitsch, M. D., & Weissman, J. S. (1998). GroEL-GroES-mediated protein folding requires an intact central cavity. *Proceedings of the National Academy of Sciences*, *95*(21), 12163–12168. https://doi.org/10.1073/pnas.95.21.12163

[27] Chakravarthi, S., Jessop, C. E., & Bulleid, N. J. (2006). The role of glutathione in disulphide bond formation and endoplasmic-reticulum-generated oxidative stress. *EMBO Reports*, *7*(3), 271–275. https://doi.org/10.1038/sj.embor.7400645

[28] Barney, B. M., LoBrutto, R., & Francisco, W. A. (2004). Characterization of a Small Metal Binding Protein from Nitrosomonas europaea. *Biochemistry*, *43*(35), 11206–11213. https://doi.org/10.1021/bi049318k

[29] Xiao, H., Peters, F. B., Yang, P.-Y., Reed, S., Chittuluru, J. R., & Schultz, P. G. (2014). Genetic Incorporation of Histidine Derivatives Using an Engineered Pyrrolysyl-tRNA Synthetase. *ACS Chemical Biology*, *9*(5), 1092–1096. https://doi.org/10.1021/cb500032c

[30] Schrader, J. M., Chapman, S. J., & Uhlenbeck, O. C. (2009). Understanding the Sequence Specificity of tRNA Binding to Elongation Factor Tu using tRNA Mutagenesis. *Journal of Molecular Biology*, *386*(5), 1255–1264. https://doi.org/10.1016/j.jmb.2009.01.021

[31] Fan, C., Xiong, H., Reynolds, N. M., & Söll, D. (2015). Rationally evolving tRNAPylfor efficient incorporation of noncanonical amino acids. *Nucleic Acids Research*, *43*(22), e156–e156. https://doi.org/10.1093/nar/gkv800

[32] Jewett, M. C., & Swartz, J. R. (2004). Mimicking the Escherichia coli cytoplasmic environment activates long-lived and efficient cell-free protein synthesis. *Biotechnology and Bioengineering*, *86*(1), 19–26. https://doi.org/10.1002/bit.20026

[33] Des Soye, B. J., Gerbasi, V. R., Thomas, P. M., Kelleher, N. L., & Jewett, M. C. (2019). A Highly Productive, One-Pot Cell-Free Protein Synthesis Platform Based on Genomically Recoded Escherichia coli. *Cell Chemical Biology*, *26*(12), 1743-1754.e9. https://doi.org/10.1016/j.chembiol.2019.10.008

[34] Johnson, D. C., Dean, D. R., Smith, A. D., & Johnson, M. K. (2005). Structure, function, and formation of biological iron-sulfur clusters. *Annual Review of Biochemistry*, *74*, 247–281. https://doi.org/10.1146/annurev.biochem.74.082803.133518

[35] Barras, F., Loiseau, L., & Py, B. (2005, January 1). *How Escherichia coli and Saccharomyces cerevisiae Build Fe/S Proteins* (R. K. Poole, Ed.). ScienceDirect; Academic Press. https://www.sciencedirect.com/science/article/pii/S006529110550002X?via%3Dihub

[36] Kriek, M., Peters, L., Takahashi, Y., & Roach, P. L. (2003). Effect of iron–sulfur cluster assembly proteins on the expression of Escherichia coli lipoic acid

synthase. *Protein Expression and Purification*, *28*(2), 241–245. https://doi.org/10.1016/s1046-5928(02)00680-0

[37] Gräwert, T., Kaiser, J., Zepeck, F., Laupitz, R., Hecht, S., Amslinger, S., Schramek, N., Schleicher, E., Weber, S., Haslbeck, M., Buchner, J., Rieder, C., Arigoni, D., Bacher, A., Eisenreich, W., & Rohdich, F. (2004). IspH Protein of *Escherichia coli*: Studies on Iron−Sulfur Cluster Implementation and Catalysis. *Journal of the American Chemical Society*, *126*(40), 12847–12855. https://doi.org/10.1021/ja0471727

[38] Zepeck, F., Gräwert, T., Kaiser, J., Schramek, N., Eisenreich, W., Bacher, A., & Rohdich, F. (2005). Biosynthesis of Isoprenoids. Purification and Properties of IspG Protein from *Escherichia coli*. *The Journal of Organic Chemistry*, *70*(23), 9168–9174. https://doi.org/10.1021/jo0510787

[39] Schwartz, C. J., Giel, J. L., Patschkowski, T., Luther, C., Ruzicka, F. J., Beinert, H., & Kiley, P. J. (2001). IscR, an Fe-S cluster-containing transcription factor, represses expression of *Escherichia coli* genes encoding Fe-S cluster assembly proteins. *Proceedings of the National Academy of Sciences*, *98*(26), 14895–14900. https://doi.org/10.1073/pnas.251550898

[40] Akhtar, M. K., & Jones, P. R. (2008). Deletion of iscR stimulates recombinant clostridial Fe-Fe hydrogenase activity and H2-accumulation in Escherichia coli BL21(DE3). *Applied Microbiology and Biotechnology*, *78*(5), 853–862. https://doi.org/10.1007/s00253-008-1377-6

[41] Young, T. S., & Schultz, P. G. (2010). Beyond the canonical 20 amino acids: expanding the genetic lexicon. *The Journal of Biological Chemistry*, *285*(15), 11039–11044. https://doi.org/10.1074/jbc.R109.091306

[42] Hong, S. H., Ntai, I., Haimovich, A. D., Kelleher, N. L., Isaacs, F. J., & Jewett, M. C. (2014). Cell-free Protein Synthesis from a Release Factor 1 Deficient Escherichia coli Activates Efficient and Multiple Site-specific Nonstandard Amino Acid Incorporation. *ACS Synthetic Biology*, *3*(6), 398–409. https://doi.org/10.1021/sb400140t

[43] Lu, Y., Welsh, J. P., Chan, W., & Swartz, J. R. (2013). Escherichia coli-based cell free production of flagellin and ordered flagellin display on virus-like particles. *Biotechnology and Bioengineering*, *110*(8), 2073–2085. https://doi.org/10.1002/bit.24903

[44] Young, T. S., Ahmad, I., Yin, J. A., & Schultz, P. G. (2010). An enhanced system for unnatural amino acid mutagenesis in E. coli. *Journal of Molecular Biology*, *395*(2), 361–374. https://doi.org/10.1016/j.jmb.2009.10.030

[45] Ohtake, S., Kita, Y., & Arakawa, T. (2011). Interactions of formulation excipients with proteins in solution and in the dried state. *Advanced Drug Delivery Reviews*, *63*(13), 1053–1073. https://doi.org/10.1016/j.addr.2011.06.011

[46] *cOmplete Protease Inhibitor Cocktail Tablets provided in a glass vial Version 14,* Product Information Sheet*,* Merck, https://www.sigmaaldrich.com/TR/en/product/roche/11836153001.

[47] Albayrak, C., & Swartz, J. R. (2013). Cell-free co-production of an orthogonal transfer RNA activates efficient site-specific non-natural amino acid incorporation. *Nucleic Acids Research*, *41*(11), 5949–5963. https://doi.org/10.1093/nar/gkt226

[48] Fan, C., Xiong, H., Reynolds, N. M., & Söll, D. (2015). Rationally evolving tRNAPylfor efficient incorporation of noncanonical amino

acids. *Nucleic Acids Research*, *43*(22), e156–e156. https://doi.org/10.1093/nar/gkv800

[49] Imburgio, D., Rong, M., Ma, K., & McAllister, W. T. (2000). Studies of Promoter Recognition and Start Site Selection by T7 RNA Polymerase Using a Comprehensive Collection of Promoter Variants†. *Biochemistry*, *39*(34), 10419–10430. https://doi.org/10.1021/bi000365w

[50] Sun, Q.-Y., Ding, L.-W., He, L.-L., Sun, Y.-B., Shao, J.-L., Luo, M., & Xu, Z.-F. (2009). Culture of Escherichia coli in SOC medium improves the cloning efficiency of toxic protein genes. *Analytical Biochemistry*, *394*(1), 144–146. https://doi.org/10.1016/j.ab.2009.07.023