

UC San Diego

UC San Diego Electronic Theses and Dissertations

Title

Optimizing Protein Degradation and Improving Energy Regeneration in Escherichia coli Cell-Free Systems and Developing A Simple Method to Dual Site-Specifically Label a Protein Using Tryptophan Auxotrophic Escherichia coli

Permalink

<https://escholarship.org/uc/item/7w5248db>

Author

Wu, Ti

Publication Date

2021

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA SAN DIEGO

Optimizing Protein Degradation and Improving Energy Regeneration in *Escherichia coli* Cell-Free Systems and Developing a Simple Method to Dual Site-Specifically Label a Protein Using Tryptophan Auxotrophic *Escherichia coli*

A dissertation submitted in partial satisfaction of the requirements for the degree Doctor of Philosophy

in

Chemistry

by

Ti Wu

Committee in charge:

Professor Simpson Joseph, Chair

Professor Neal K. Devaraj

Professor Thomas Hermann

Professor Terence Hwa

Professor Ulrich F. Müller

2021

Copyright

Ti Wu, 2021

All rights reserved.

The Dissertation of Ti Wu is approved, and it is acceptable in quality  
and form for publication on microfilm and electronically.

University of California San Diego

2021

## DEDICATION

This work is dedicated to everyone who have supported me on the journey of science and life.

## TABLE OF CONTENTS

<b>Dissertation Approval Page</b> .....	<b>iii</b>
<b>Dedication</b> .....	<b>iv</b>
<b>Table of Contents</b> .....	<b>v</b>
<b>List of Figures</b> .....	<b>viii</b>
<b>List of Tables</b> .....	<b>ix</b>
<b>Acknowledgements</b> .....	<b>x</b>
<b>Vita</b> .....	<b>xii</b>
<b>Abstract of the Dissertation</b> .....	<b>xiii</b>
<b>Part I: Optimization and Improving Protein Synthesis and Degradation in <i>Escherichia coli</i> Cell-Free Systems</b> .....	<b>1</b>
<b>Chapter 1: Introduction to <i>Escherichia coli</i> Cell-Free Systems</b> .....	<b>2</b>
Acknowledgement.....	5
<b>Chapter 2: Optimization of ClpXP activity and protein synthesis in an <i>E. coli</i> extract-based cell-free expression system</b> .....	<b>6</b>
2.1. Abstract.....	6
2.2. Introduction .....	6
2.3. Results .....	8
2.4. Discussion.....	33
2.5. Methods .....	38
2.5.1 Cloning and protein purification.....	38
2.5.2 Fluorescent protein measurement and quantification.....	39
2.5.3 <i>In vitro</i> protein degradation assay .....	40
2.5.4 $k_{cat}$ and $K_M$ measurement.....	40

2.5.5	Macromolecular crowding effect on ClpXP activity.....	41
2.5.6	Optimization of ClpXP activity in cell extract.....	41
2.5.7	<i>In vitro</i> Transcription-Translation (TX-TL) of deGFP.....	42
2.5.8	Optimization of TX-TL system under ClpXP degradation condition .....	42
2.5.9	ClpXP proteolytic activity under optimized TX-TL condition .....	43
2.5.10	Simultaneous protein synthesis and degradation in TX-TL systems .....	43
2.5.11	ATP consumption assay .....	43
2.5.12	Vesicle preparation.....	44
2.5.13	Degradation assay in vesicles.....	44
2.6.	Acknowledgement.....	45
 <b>Chapter 3: Improving Energy Supplies in <i>E. coli</i> Cell-Free Systems with Class III</b>		
<b>Polyphosphate Kinase 2 .....</b>		
		<b>46</b>
3.1.	Abstract.....	46
3.2.	Introduction .....	46
3.3.	Results .....	50
3.4.	Discussion.....	59
3.5.	Methods .....	60
3.5.1	Cloning and protein purification.....	60
3.5.2	Kinetic measurement of MrPPK2 activity.....	60
3.5.3	ATP consumption assay .....	60
3.5.4	Fluorescent protein measurement and quantification.....	61
3.5.5	Cell-free Transcription-Translation (TX-TL) of deGFP .....	61
3.5.6	<i>In vitro</i> and extract-based cell-free protein degradation assay .....	62
3.6.	Acknowledgement.....	62

<b>Part II: Dual Site-Specifically Label a Protein Using Tryptophan Auxotrophic <i>Escherichia coli</i></b> .....	<b>63</b>
<b>Chapter 4: A Simple Method to Dual Site-Specifically Label a Protein Using Tryptophan Auxotrophic <i>Escherichia coli</i></b> .....	<b>64</b>
4.1. Abstract .....	64
4.2. Introduction .....	65
4.3. Results .....	67
4.3.1. General schema of the protocol .....	67
4.3.2. Design of single cysteine single tryptophan RF1 (scswRF1) .....	68
4.3.3. Expression of 5-hydroxytryptophan-incorporated scswRF1 .....	71
4.3.4. Dual site-specific labeling of 5-hydroxytryptophan-incorporated scswRF1 .....	73
4.3.5. Fluorescence spectroscopic analysis of dye-labeled RF1 .....	75
4.4. Discussions .....	75
4.5. Materials and Methods .....	77
4.5.1. Chemicals, buffers, and bacterial strains .....	77
4.5.2. Mutant RF1 expression and purification .....	77
4.5.3. Expression of RF1 protein with 5-hydroxytryptophan (5-HW) .....	78
4.5.4. Labeling of RF1 mutants .....	79
4.5.5. Staining, imaging, and fluorescence spectroscopy .....	79
4.6. Acknowledgement .....	80
<b>References</b> .....	<b>81</b>

## LIST OF FIGURES

Figure 1.1 The production and components in bacterial cell-free systems .....	3
Figure 2.1 Degradation of sfGFP-ssrA by ClpXP in buffer .....	10
Figure 2.2 Supplementary data for degradation of sfGFP-ssrA .....	12
Figure 2.3 Optimization of ClpXP degradation in the TX-TL system.....	15
Figure 2.4 Degradation of sfGFP-ssrA in vesicles by ClpXP and optimizing both TX-TL and ClpXP activity .....	19
Figure 2.5 Optimization of the ClpXP activity under TX-TL condition.....	21
Figure 2.6 Optimization of protein synthesis under the ClpXP degradation reaction condition by adding magnesium glutamate .....	24
Figure 2.7 Optimization of protein synthesis under ClpXP degradation condition by adding polyamines.....	26
Figure 2.8 Effects of glycerol on TX-TL protein synthesis, and ATP consumption assays .....	28
Figure 2.9 ATP consumption and protein degradation by ClpXP.....	31
Figure 2.10 Effects of AMP and ADP on protein degradation by ClpXP.....	34
Figure 3.1 Outline of the challenge of AMP accumulation in Cell-Free Systems .....	49
Figure 3.2 Energy consumption profiles of degradation by ClpXP .....	51
Figure 3.3 Aminoacylation contributes a significant part of AMP accumulation.....	54
Figure 3.4 Phosphorylation reaction catalyzed by MrPPK2 .....	55
Figure 3.5 Energy consumption profile in CFS with MrPPK2 and polyP .....	57
Figure 3.6 Effects of adding MrPPK2 into CFS.....	58
Figure 4.1 Schema of incorporation of tryptophan analogs into recombinant protein using Trp auxotrophic <i>E. coli</i> for dual site-specific labeling.....	69
Figure 4.2. Structural schema of <i>E. coli</i> release factor 1 (RF1) in open and closed conformations .....	70
Figure 4.3 Incorporation of 5-hydroxytryptophan into single-cysteine single-tryptophan RF1 ...	72
Figure 4.4 Site-specific dual-labeling of 5HW-incorporated RF1 and spectroscopic analysis.....	74

## LIST OF TABLES

Table 1.1 History of key developments of cell-free systems .....	4
Table 2.1 Final concentration of the energy regeneration system and $Mg^{2+}$ .....	22

## ACKNOWLEDGEMENTS

First and foremost, I would like to thank my advisor, Professor Simpson Joseph. Without his guidance and support, this dissertation would not have been possible. I would also like to acknowledge my collaborators from Devaraj lab, Hasty lab, and Tsimring lab in UC San Diego, Bowman lab in CU Boulder, and Manoharan lab in Harvard University, and colleagues from Müller lab, Toor lab, Zid lab and McHugh lab, for all the discussions and ideas exchanged during our physical and virtual meetings.

I would also like to thank my lab mates: Xinying Shi, for her up-lifting laughter in the lab and the guidance and help in everything related to experimental biochemistry; Dr. Krista Trappl, without her I would not be able to start the tryptophan project, and we would probably miss most birthdays of our lab members; Dr. Youssi Athar and Aaron Ward, my classmates and first American friends I made in this country, with whom I am so glad to share the time and space in Joseph lab during our PhD journey; and all other Joseph lab members and alumni for the coffee breaks and happy hours we spent together.

Last but not least, I want to give huge thanks to my family and friends. Thank all the people I have met in the past years in the United States, especially in San Diego, for all the memories we have created and shared. Huge thanks to my parents for raising me up and nurturing me as an independent and open-minded person, and selflessly supporting all the decisions I made throughout my life. Finally, I would also like to thank my dear friend and life partner Dr. Ching-Hsin Huang for her love and support.

Chapter 1, in part, contains material currently being prepared for submission for publication. Ti Wu, and Simpson Joseph. The dissertation author was the primary author of this paper.

Chapter 2, in full, is a reprint of published material: Xinying Shi, Ti Wu, Christian Cole, Neal K. Devaraj, Simpson Joseph. Optimization of ClpXP activity and protein synthesis in an *E. coli* extract-based cell-free expression system. *Sci. Rep.* 8, 587 (2018). The dissertation author was one of the primary authors of this paper.

Chapter 3, in part, contains material currently being prepared for submission for publication. Ti Wu, and Simpson Joseph. Improving Energy Supplies in *E. coli* Cell-Free Systems with Class III Polyphosphate Kinase 2. The dissertation author was the primary author of this paper.

Chapter 4, in full, contains material currently in submission to PLOS ONE as in November 2021. Ti Wu and Simpson Joseph. A Simple Method to Dual Site-Specifically Label a Protein Using Tryptophan Auxotrophic *Escherichia coli*. The dissertation author was the primary author of this paper.

## VITA

2011 Bachelor of Science, Chemistry, National Taiwan University, Taiwan

2011-2012 Republic of China (Taiwan) Air Force (Conscription)

2012-2013 Research Assistant, National Taiwan University, Taiwan

2013-2014 Research Assistant, Academia Sinica, Taiwan

2014-2015 Teaching Assistant, University of California San Diego, U.S.A.

2015-2018 Research Assistant, University of California San Diego, U.S.A.

2018-2021 Teaching Assistant, University of California San Diego, U.S.A.

2021 Doctor of Philosophy, University of California San Diego, U.S.A.

## PUBLICATIONS

\* These authors contribute equally to the works.

1. G.-M. Ho\*, C.-J. Huang\*, E. Y.-Z. Li, S.-K. Hsu, T. Wu, M. M. L. Zulueta, K. B. Wu, S.-C. Hung. Unconventional exo selectivity in thermal normal-electron-demand Diels-Alder reactions. *Sci. Rep.* 6, 35147 (2016).
2. X. Shi\*, T. Wu\*, C. Cole, N. K. Devaraj, S. Joseph. Optimization of ClpXP activity and protein synthesis in an *E. coli* extract-based cell-free expression system. *Sci. Rep.* 8, 587 (2018).
3. T. Wu, S. Joseph. A Simple Method to Dual Site-Specifically Label a Protein Using Tryptophan Auxotrophic *Escherichia coli*. *PLOS ONE*. (In submission, 2021)

## FIELDS OF STUDY

Major Field: Physical Science

Studies in Chemistry and Biochemistry

Professors Simpson Joseph

## ABSTRACT OF THE DISSERTATION

Optimizing Protein Degradation and Improving Energy Regeneration in *Escherichia coli* Cell-Free Systems and Developing a Simple Method to Dual Site-Specifically Label a Protein Using Tryptophan Auxotrophic *Escherichia coli*

by

Ti Wu

Doctor of Philosophy in Chemistry

University of California San Diego, 2021

Professor Simpson Joseph, Chair

This dissertation consists of two parts: the first is about *E. coli* cell-free systems, describing the works on optimization and improving protein synthesis and degradation and a novel method to improve energy supplies in extract-based cell-free systems; the second part is about a method for incorporating 5-hydroxytryptophan into *E. coli* release factor 1 protein (RF1) for dual site-specific labeling.

Protein synthesis and degradation are fundamental processes in all living cells—the former creates all the protein required for biochemical and cellular functions, and the latter is essential to remove both damaged proteins and intact proteins that are no longer needed by the cell. Moreover,

tuning protein synthesis and degradation is an important strategy for cell to regulate enzymatic activities and more complex events such as cell cycle. Here we are interested in creating synthetic genetic circuits that function in a cell-free expression system, which will require not only an efficient protein expression platform but also a robust protein degradation system in cell extract. Therefore, we purified and tested the activity of *E. coli* ClpXP protease in cell-free transcription-translation (TX-TL) systems that used *E. coli* S30 cell extract. Surprisingly, our studies showed that purified ClpXP added to the TX-TL system has very low proteolytic activity. The low activity of ClpXP was correlated with the rapid consumption of adenosine triphosphate (ATP) in cell extract.

We further showed that adenosine monophosphate (AMP) accumulated in the cell-free systems, reducing the availability of the other two adenosine nucleotides, ATP, which is the energy molecule, and adenosine diphosphate (ADP), which is used for regenerating ATP by the energy regeneration system. We used class III polyphosphate kinase 2 from *Meiothermus ruber* (MrPPK2) to mitigate AMP accumulation by transforming AMP back to ADP and hence increase the concentration of ADP and ATP as well as improve the performance of protein synthesis and degradation in the cell-free systems.

Site-specifically labeling proteins with multiple dyes or molecular moieties is an important yet not-trivial task for biochemical and biophysical research, such as when using Förster resonance energy transfer (FRET) to study dynamics of protein conformational change. Many strategies have been devised, but usually done on a case-by-case scenario. We developed an easy-to-use protocol for site-specific dual labeling by incorporating 5-hydroxytryptophan using a tryptophan auxotroph strain of *E. coli*, which can provide an orthogonal bioconjugation handle additional to conventional amine- or thiol-based labeling method. As demonstration, we incorporated 5-

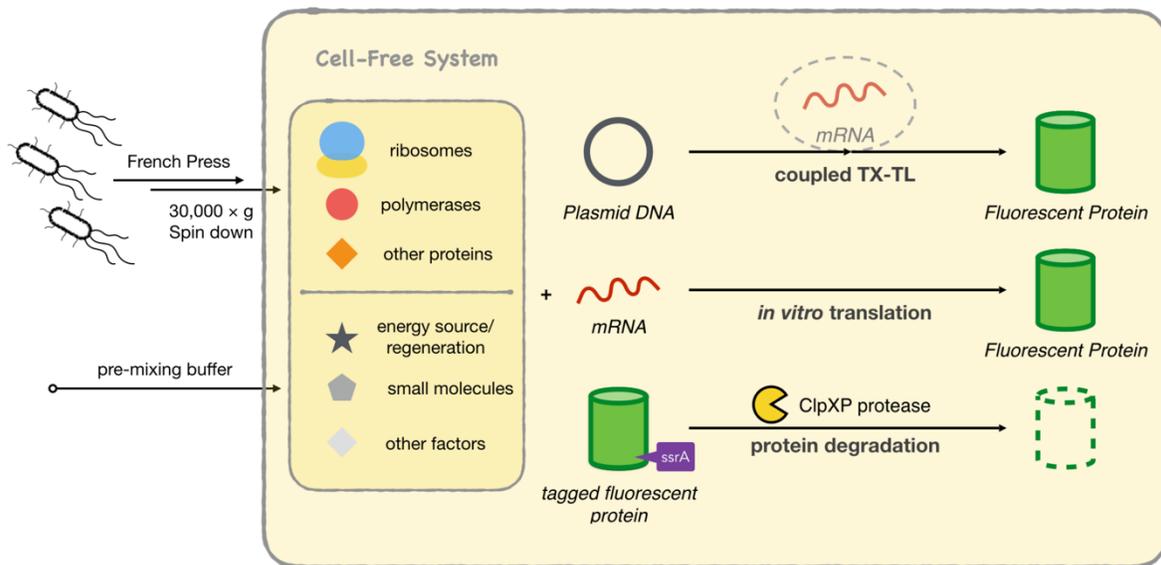
hydroxytryptophan into *E. coli* release factor 1 (RF1), a protein known to possess two different conformations, and labeled two different fluorophores specifically on the 5-hydroxytryptophan and cysteine sites. This method provides an easier way to achieve dual- or multi-labeling of protein that will be useful for biochemical or biophysical experiments like FRET.

**Part I: Optimization and Improving Protein Synthesis  
and Degradation in *Escherichia coli* Cell-Free Systems**

# Chapter 1: Introduction to *Escherichia coli* Cell-Free Systems

Cell-free systems are aqueous mixtures containing various essential biochemical components which make it capable to recapitulate intracellular activities without the constraint of cell membrane (**Figure 1.1**)<sup>1,2</sup>. It has a long history that could be dated back to the early days of modern biochemistry and molecular biology (**Table 1.1**), when Eduard Buchner showed that fermentation processes could be performed outside the cells of living organisms by the turn of the 20<sup>th</sup> century<sup>3</sup>, and Marshall W. Nirenberg and J. Heinrich Matthaei used cell-free peptide synthesis with customized messenger RNA to decipher the first genetic code in 1961<sup>4,5</sup>. Recently cell-free systems have re-gaining popularity along with the blooming field of synthetic biology, and serve as a foundation for many different applications, e.g. metabolic engineering and prototyping genetic circuits, creating novel biosensor or diagnostical tools, and powering artificial cells and cell-mimics<sup>2,6-9</sup>.

Traditionally, cell-free systems are made of cell extracts from different organisms, such as *Escherichia coli*<sup>10,11</sup>, *Saccharomyces cerevisiae*<sup>12</sup>, wheat embryos<sup>10,13</sup>, rabbit reticulocyte<sup>10</sup>, and even human cell lines<sup>14,15</sup>. Cell extract that contains essential molecular machineries, including ribosomes, polymerases, proteases, and many other enzymes or protein factors will then be mixed with additional molecules, such as amino acids, nucleotides, nucleic acids, salts, small molecular factors, or even macromolecular crowding agents or additional purified proteins, to create a fully functional cell-free system. Another way to build cell-free systems is to assemble them using purified protein components, and hence also called PURE system<sup>16,17</sup>. The advantage of PURE



**Figure 1.1 The production and components in bacterial cell-free systems**

An illustration showing the production, components, and commonly performed biochemical reactions in the bacterial cell-free systems.

**Table 1.1 History of key developments of cell-free systems**

<b>Year</b>	<b>Researchers</b>	<b>Milestones</b>
1907	Eduard Buchner	Cell-free fermentation of yeast <sup>3</sup>
1954	Paul Zamecnik	First protein synthesis in cell-free system <sup>18</sup>
1961	Marshall Nirenberg and J. Heinrich Matthaei	Using cell-free system to decipher the genetic codes <sup>4,5</sup>
1967	JoAnne DeVries and Geoffrey Zubay	Optimizing <i>E. coli</i> cell-free system for circular DNA <sup>19</sup>
1988	Alexander Spirin	Using continuous flow to supply cell-free system <sup>20</sup>
2001	Yoshihiro Shimizu and Takuya Ueda	Inventing <i>E. coli</i> PURE system <sup>16</sup>

system is that it has a well-defined composition, so it's easier to customize by adding additional purified components for different experiments and applications, but the downside is the cost and relatively poor performance. For *E. coli* PURE system, some progresses have been made recently to improve protein synthesis yield<sup>21</sup> and reduce the production cost<sup>22</sup>, yet extract-based systems are still general more efficient (in batch mode it produces milligrams per milliliter reaction comparing to a few hundred micrograms in PURE system) and economical<sup>21-24</sup>.

Recently, the applications of cell-free systems have gone far beyond protein synthesis<sup>1,2,8</sup>. Cell-free technology is widely used in different kinds of applications, including but not limited to designing genetic circuits<sup>25-27</sup>, prototyping engineering metabolic pathways<sup>28</sup>, biomanufacturing<sup>29,30</sup>, creating novel biosensor or diagnostic tools<sup>6,31,32</sup>, and creating artificial cells or cell-mimics<sup>7,9,25,33,34</sup>. It could also be deployed to different materials, for example, paper<sup>6</sup>, hydrogel<sup>35</sup>, and clay hydrogel<sup>7,36</sup>.

Despite all the exciting progress, some fundamental problems of cell-free systems still need to be explored, such as compatibility to different biochemical reactions, energy regeneration, macromolecular crowding effects, more efficient and economical production methods, and how to develop cell-free systems from more natural or engineered cell types. In the following two chapters, we are presenting our works on optimizing of protein degradation and synthesis and developing an enzymatic method to improve energy supply in *E. coli* extract-based cell-free expression system.

## **Acknowledgement**

Chapter 1, in part, contains material currently being prepared for submission for publication. Ti Wu, and Simpson Joseph. The dissertation author was the primary author of this paper.

# Chapter 2: Optimization of ClpXP activity and protein synthesis in an *E. coli* extract-based cell-free expression system

## 2.1. Abstract

Protein degradation is a fundamental process in all living cells and is essential to remove both damaged proteins and intact proteins that are no longer needed by the cell. We are interested in creating synthetic genetic circuits that function in a cell-free expression system. This will require not only an efficient protein expression platform but also a robust protein degradation system in cell extract. Therefore, we purified and tested the activity of *E. coli* ClpXP protease in cell-free transcription-translation (TX-TL) systems that used *E. coli* S30 cell extract. Surprisingly, our studies showed that purified ClpXP added to the TX-TL system has very low proteolytic activity. The low activity of ClpXP was correlated with the rapid consumption of adenosine triphosphate (ATP) in cell extract. We improved the activity of ClpXP in cell extract by adding exogenous ATP and an energy regeneration system. We then established conditions for both the protein synthesis and protein degradation by ClpXP to occur simultaneously in the TX-TL systems. The optimized conditions for ClpXP activity will be useful for creating tunable synthetic genetic circuits and *in vitro* synthetic biology.

## 2.2. Introduction

Protein degradation is an essential process in all living cells that is used to remove damaged or misfolded proteins and to tune gene expression by temporally controlling the concentration of regulatory proteins. In bacteria, targeted protein degradation is carried out by the AAA+ family of

proteases, which include ClpAP, ClpXP, FtsH, HslUV (also known as ClpYQ), or Lon<sup>37,38</sup>. The mechanism of ClpXP is well understood and therefore ClpXP protein degradation is widely used in in vitro synthetic biology<sup>33,39,40</sup>. The ClpXP protease is an oligomer formed by 6 identical ClpX proteins, that form a hexameric ring, and 14 identical ClpP proteins, which form a double-ring structure<sup>41,42</sup>. The overall structure of the ClpXP protease resembles a barrel with a central axial pore. The ClpX protein uses ATP binding and hydrolysis to unfold protein substrates and translocate the denatured polypeptide through the central pore to the proteolytic chamber formed by the ClpP protease<sup>37,38</sup>. Most protein substrates that are targeted for degradation have a tag sequence that is recognized by a specific protease. In the case of ClpXP, protein substrates that have five distinct motifs (three located at the N-terminus and two at the C-terminus) are targeted for degradation<sup>43</sup>. Additionally, the *E. coli* ClpXP protease degrades proteins that have an 11-amino acid ssrA tag at the C terminus<sup>44</sup>. The ssrA tag is recognized by SspB, an adaptor protein that also binds to ClpX<sup>45-47</sup>. The interaction of SspB both with the ssrA tag and with ClpX enhances substrate recognition and improves degradation at lower substrate concentrations<sup>45</sup>.

We are interested in establishing dynamic genetic circuits using a cell-free transcription-translation (TX-TL) system. To create dynamic genetic circuits, it is essential to have robust protein degradation in the TX-TL system to provide a constant flux of proteins. Although previous studies have used ClpXP to target specific regulatory and reporter proteins for degradation in vitro, these studies used simple buffers to test the activity of ClpXP or used the endogenous ClpXP in the TX-TL system<sup>33,39</sup>. In these latter studies, the amount of active ClpXP in cell extract is low, resulting in a low rate of protein degradation<sup>39</sup>. To overcome this problem, one study added the genes coding for ClpX and ClpP to the TX-TL system to increase the concentration of the ClpXP protease in cell extract, which resulted in improved degradation of the target proteins<sup>40</sup>.

Nevertheless, adding known concentration of highly active, purified ClpXP to the TX-TL system would offer greater flexibility in the design and modeling of dynamic gene circuits. Indeed a previous report showed that adding a purified, linked-hexameric version of ClpX to a TX-TL system increased the degradation of *ssrA*-tagged reporter proteins, indicating that this approach may be useful for creating dynamic gene circuits<sup>48</sup>.

Here we report the optimization of protein degradation in *E. coli* S30 cell extract-based TX-TL system by the ClpXP protease. To establish a protein degradation system that is highly active in the TX-TL system, we decided to use purified *E. coli* ClpXP protease. Our studies revealed that ClpXP activity is poor in the extract-based system because the concentration of ATP is low in the standard S30 cell extract used for TX-TL. Fortunately, adding ATP and an energy regeneration system significantly increased the activity of ClpXP in cell extract. We also demonstrate that ClpXP is active in cell extract enclosed inside phospholipid vesicles, suggesting that our work could have future application in the development of artificial cells. By using radioactive ATP, we analyzed how ATP was consumed and converted to other adenine nucleotides in either reaction buffers or in the TX-TL system under different reaction conditions. Finally, we optimized the TX-TL system so that both protein synthesis and protein degradation could occur simultaneously and efficiently. Our results suggest that TX-TL systems supplemented with known concentrations of ClpXP and an energy regeneration system could be used to create synthetic genetic circuits that function in vesicles, enabling robust and predictable protein degradation in artificial cells.

### **2.3. Results**

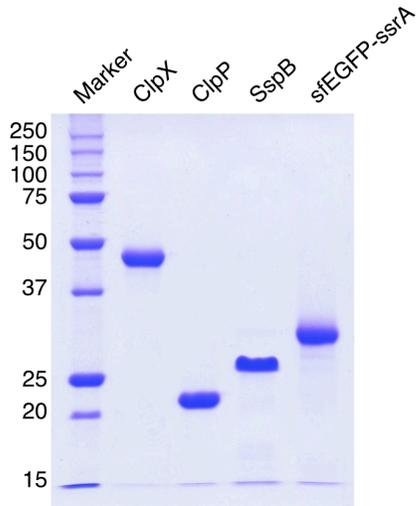
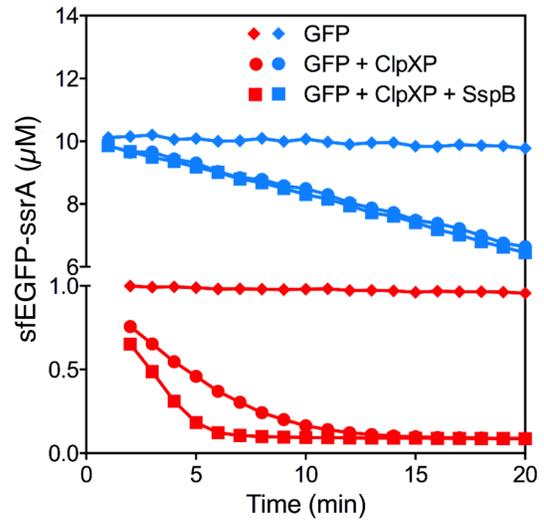
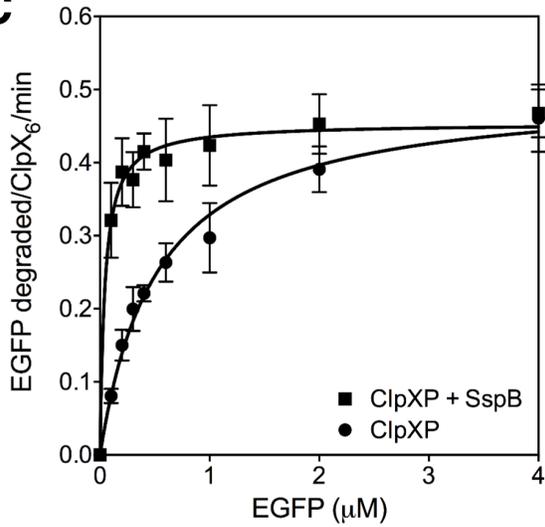
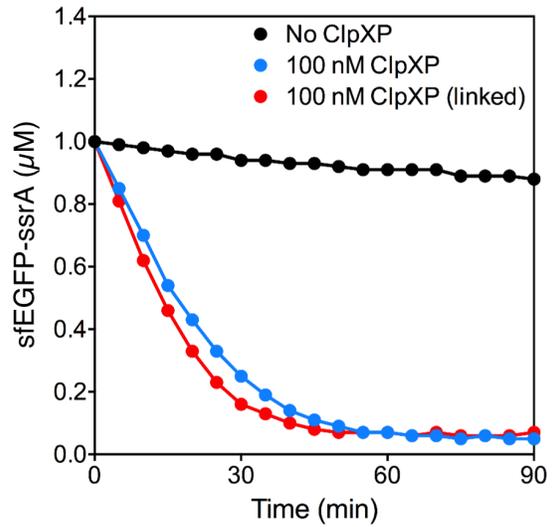
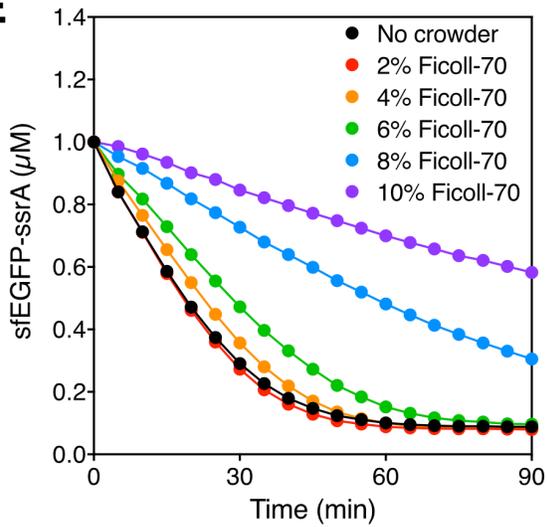
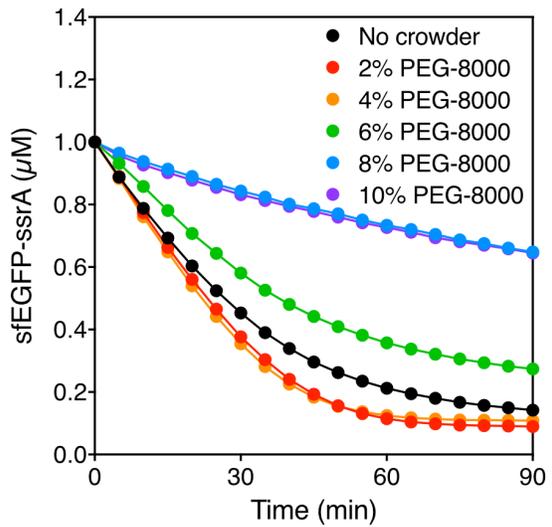
To establish the ClpXP protein degradation system, we overexpressed and purified C-terminal His-tagged ClpX, ClpP and SspB using *E. coli* (**Figure 2.1A**). We also constructed a

vector to overexpress and purify the N-terminal His-tagged superfolder GFP that has the *ssrA* tag sequence at the C-terminus (sfGFP-*ssrA*). We used sfGFP-*ssrA* as the substrate for the ClpXP protein degradation system. The sfGFP-*ssrA* protein was highly expressed in *E. coli*, and the yield after purification was about 30 mg per liter culture. We analyzed the activity of the purified ClpXP system to degrade sfGFP-*ssrA* by measuring the fluorescence intensity (**Figure 2.1B**). Additionally, we verified that the decrease in fluorescence intensity was due to protein degradation by ClpXP using SDS-PAGE (**Figure 2.2A**). Steady-state kinetic parameters ( $k_{\text{cat}}$  and  $K_M$ ) were determined in the PD buffer. The  $k_{\text{cat}}$  for the degradation of sfGFP-*ssrA* by ClpXP was  $0.46 \pm 0.01 \text{ min}^{-1}$  in the presence of SspB and  $0.50 \pm 0.02 \text{ min}^{-1}$  in the absence of SspB (**Figure 2.1C**). The  $K_M$  was  $0.043 \pm 0.007 \text{ }\mu\text{M}$  in the presence of SspB and  $0.52 \pm 0.06 \text{ }\mu\text{M}$  in the absence of SspB. Thus, SspB is important for protein degradation by ClpXP when the concentration of the protein substrate (sfGFP-*ssrA*) is low. These results are consistent with previously published data<sup>45</sup>.

Previous studies showed that a linked hexameric version of ClpX is more active in protein degradation *in vitro* than the monomeric wild type ClpX<sup>48</sup>. We purified the linked hexameric version of ClpX and compared its activity to the monomeric ClpX. The degradation of sfGFP-*ssrA* by the wild type ClpXP and the linked ClpXP was monitored. The linked ClpXP was slightly more active (initial rates are  $38 \text{ nM min}^{-1}$  and  $30 \text{ nM min}^{-1}$ , respectively) than the wild type ClpXP (**Figure 2.1D**). However, we decided to optimize the activity of the wild type ClpXP because the yield of the purified linked ClpXP was very low.

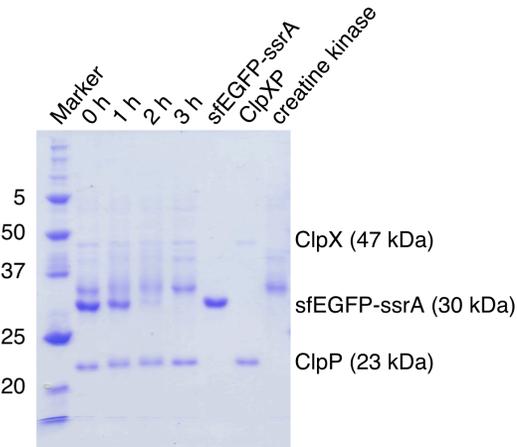
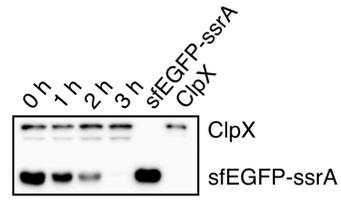
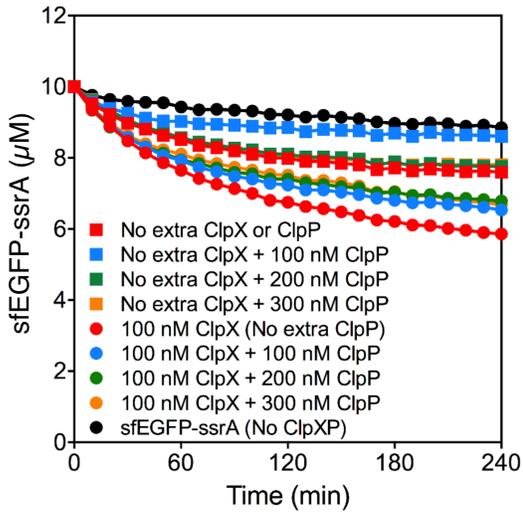
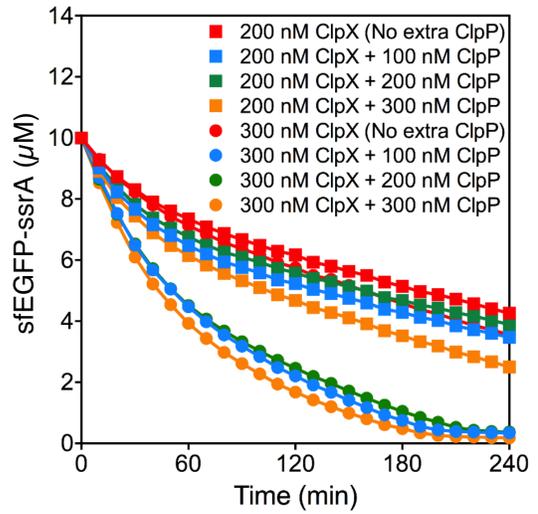
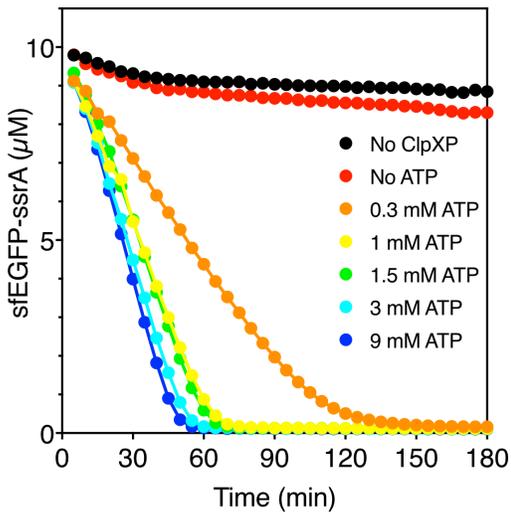
## Figure 2.1 Degradation of sfGFP-ssrA by ClpXP in buffer

(A) SDS-PAGE showing the purified *E. coli* ClpX, ClpP, SspB, and sfGFP-ssrA proteins. (B) Degradation of sfGFP-ssrA by ClpXP in the presence (squares) or in the absence (circles) of SspB performed in the PD buffer. The concentration of sfGFP-ssrA was either 1  $\mu\text{M}$  (red) or 10  $\mu\text{M}$  (blue), and the concentration of ClpXP was 300 nM. No ClpXP was added to control reactions done in parallel (diamonds). (C) Measurement of steady-state kinetic parameters ( $k_{cat}$  and  $K_M$ ). The initial rate of protein degradation per ClpXP molecule at different sfGFP-ssrA concentration was used to determine the kinetic parameters. The  $k_{cat}$  for ClpXP degradation of sfGFP-ssrA is  $0.46 \pm 0.01 \text{ min}^{-1}$  in the presence of SspB (square) and  $0.50 \pm 0.02 \text{ min}^{-1}$  in the absence of SspB (circle). The  $K_M$  is  $0.04 \pm 0.007 \mu\text{M}$  in the presence of SspB and  $0.52 \pm 0.05 \mu\text{M}$  in the absence of SspB. (D) Protein degradation by the wild type ClpX (blue) and the linked-hexameric version (red) of ClpX<sub>6</sub>. The degradation of 1  $\mu\text{M}$  sfGFP-ssrA by 100 nM of each version of ClpXP. No ClpXP was added to the control reaction done in parallel (black). (E) Crowding effect of Ficoll-70 on ClpXP degradation. 1  $\mu\text{M}$  sfGFP-ssrA were degraded by 100 nM ClpXP in the presence of 0 to 10% w/w of crowding agent. (F) Crowding effect of PEG-8000 on ClpXP degradation. 1  $\mu\text{M}$  sfGFP-ssrA were degraded by 100 nM ClpXP in the presence of 0 to 10% w/w of crowding agent. The error bars show the standard deviation from at least three independent experiments. Plots without error bars are representative of at least three repeated experiments.

**A****B****C****D****E****F**

## Figure 2.2 Supplementary data for degradation of sfGFP-ssrA

(A) SDS-PAGE showing the degradation of 10  $\mu$ M sfGFP-ssrA by 100 nM ClpXP in the PD buffer system. Samples were collected at 0-, 1-, 2-, and 3-hour time points. (B) Western Blot showing degradation of 10  $\mu$ M sfGFP-ssrA degradation by 300 nM ClpXP in the cell extracts with premix 4. Samples were collected at 0-, 1-, 2-, and 3-hour time points. (C) and (D) The degradation of 10  $\mu$ M sfGFP-ssrA at different concentrations of ClpX and ClpP in the TX-TL system with premix 4. (C) sfGFP-ssrA were degraded in the presence of no extra ClpX with addition of no extra ClpP (red, square), 100 nM ClpP (blue, square), 200 nM ClpP (green, square) and 300 nM ClpP (orange, square) or 100 nM ClpX with addition of no extra ClpP (red, circle), 100 nM ClpP (blue, circle), 200 nM ClpP (green, circle) and 300 nM ClpP (orange, circle). A control reaction (black) with no extra ClpX and ClpP was performed in the TX-TL system. (D) sfGFP-ssrA were degraded in the presence of 200 nM ClpX with addition of no extra ClpP (red, square), 100 nM ClpP (blue, square), 200 nM ClpP (green, square) and 300 nM ClpP (orange, square), or 300 nM ClpX with addition of no extra ClpP (red, circle) 100 nM ClpP (blue, circle), 200 nM ClpP (green, circle) and 300 nM ClpP (orange, circle). (E) Titration of ATP concentrations in protein degradation assay. 10  $\mu$ M sfGFP-ssrA was degraded by 300 nM ClpXP in premix buffers supplying 0 mM (red), 0.3 mM (orange), 1 mM (yellow), 1.5 mM (green), 3 mM (cyan), and 9 mM (blue) of ATP. Data are representative of three repeated experiments.

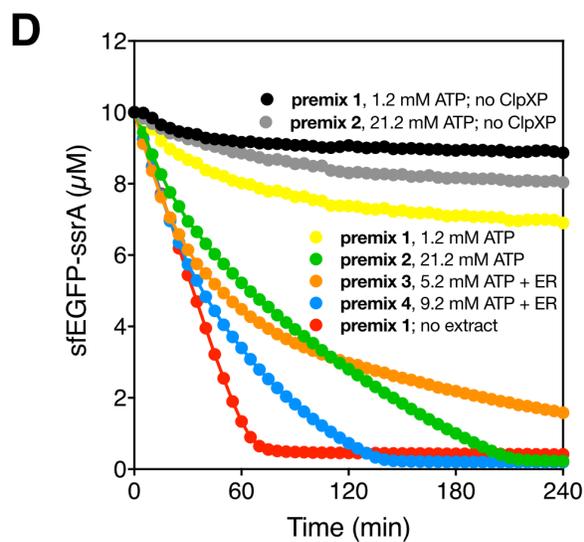
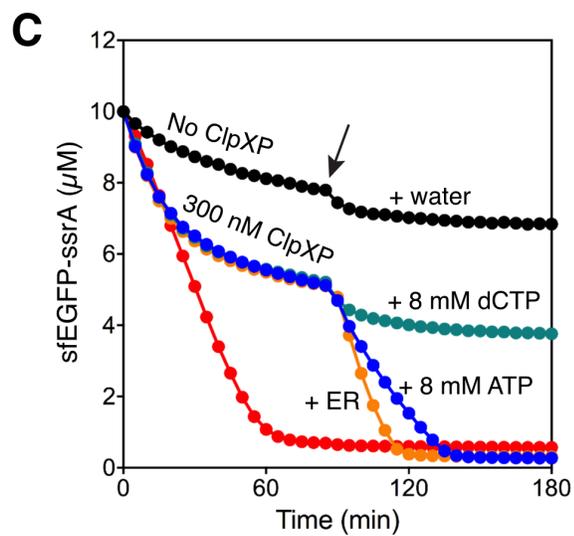
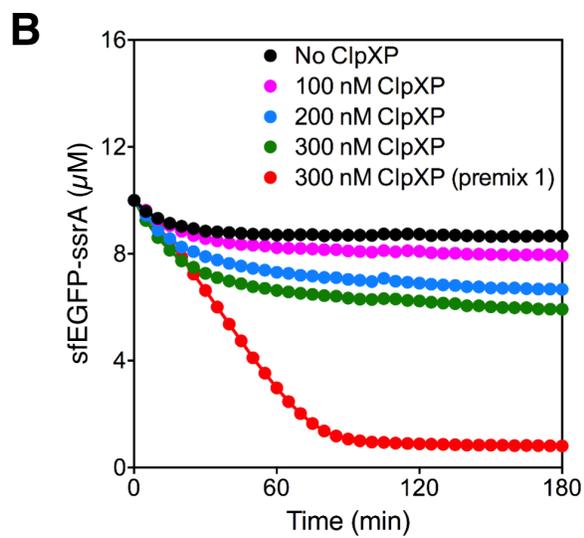
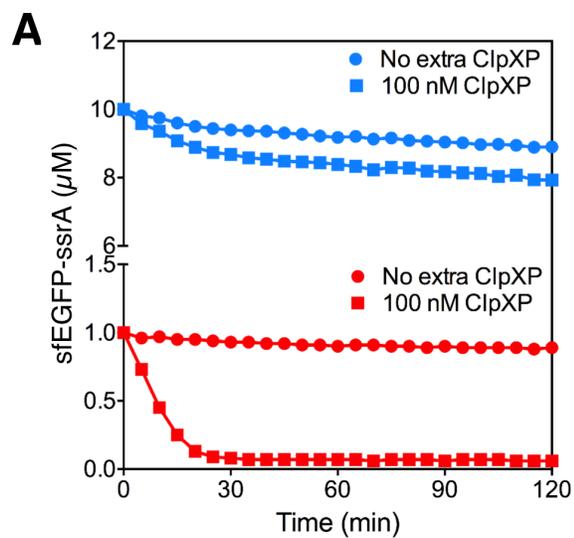
**A****B****C****D****E**

We also examined the macromolecular crowding effects on the proteolytic activity of ClpXP. Studies have shown that macromolecular crowders generally stimulate biochemical reactions<sup>49-51</sup>, and hence 2% PEG-8000 is a standard component in TX-TL system. Here we added two commonly used macromolecular crowders, Ficoll-70 and PEG-8000 up to 10% w/w to ClpXP reactions. Protein degradation by ClpXP was not affected by a low concentration of crowders but higher concentrations (larger than about 4%) decreased ClpXP activity (**Figure 2.1E and 2.1F**). It is possible that the crowded environment stabilizes sfGFP-ssrA protein and slows down the unfolding process, which is the ATP-consuming step in ClpXP-dependent protein degradation. This may explain the inhibition of ClpXP activity at high concentrations of macromolecular crowders. Alternatively, high concentrations of crowders may cause protein precipitation or non-specific binding of proteins to the crowders resulting in the lowered ClpXP activity<sup>16</sup>. Nevertheless, we confirmed that 2% PEG-8000 present in TX-TL system will not affect the performance of ClpXP.

One of our long-term goals is to develop artificial cells that can execute dynamic genetic circuits. An essential feature for a dynamic genetic circuit is to have a sink to degrade the newly synthesized proteins without damaging the proteins required for TX-TL<sup>40</sup>. Targeted protein degradation could be achieved using the ClpXP protease. However, the *E. coli* S30 extract that we used for *in vitro* TX-TL has very low proteolytic activity from endogenous ClpXP and other proteases. (**Figure 2.3A**). Increasing the concentration of purified ClpXP protease in the *E. coli* S30 extract led to only a modest increase in the degradation of sfGFP-ssrA (**Figure 2.3B**). We hypothesized that the energy in the system, 1.2 mM ATP and an energy regeneration system containing 30 mM phosphoenolpyruvate (PEP) and 48 U/mL pyruvate kinase (PK), was insufficient and limited the proteolytic activity of ClpXP in the TX-TL system. To verify this, we

### Figure 2.3 Optimization of ClpXP degradation in the TX-TL system

(A) ClpXP activity in the TX-TL system. The degradation of 1  $\mu$ M (red) and 10  $\mu$ M (blue) sfGFP-ssrA by 100 nM extra ClpXP added to the complete TX-TL system (square) or by the endogenous ClpXP present in the TX-TL system (circle). (B)(C)(D) A positive control reaction was performed with 300 nM ClpXP and 10  $\mu$ M sfGFP-ssrA in premix 1 (red). (B) The degradation of sfGFP-ssrA at different concentrations of ClpXP in the TX-TL system. 10  $\mu$ M sfGFP-ssrA were degraded by 0 nM (black), 100 nM (magenta), 200 nM (blue), or 300 nM (green) of ClpXP. (C) Analyzing the effect of increasing the ATP concentration on protein degradation by the wild type ClpXP in the TX-TL system. The degradation of 10  $\mu$ M sfGFP-ssrA by 0 nM (black), or 300 nM of ClpXP. After the degradation reactions were incubated for 1.5 hour, either 1  $\mu$ L of water (magenta), 8 mM dCTP (green), 8 mM ATP (blue) or 8 mM ATP, 16 mM PEP, and 16 U/mL pyruvate kinase (indicated as ER) (orange) solution was added at the time point shown by the arrow. 1  $\mu$ L of water was also added in negative control reaction (black) at the same time point. (D) Optimization of protein degradation by the wild type ClpXP in the TX-TL system. The degradation of 10  $\mu$ M sfGFP-ssrA by 300 nM ClpXP at different concentrations of ATP and in the presence of energy regeneration system. The standard TX-TL system with premix1 (yellow), premix 2 (orange), premix 3 (green), or premix 4 (blue). Control reactions were performed without ClpXP with premix 1 (black) and premix 2 (grey). Positive control reaction (red) was the degradation of 10  $\mu$ M sfGFP-ssrA with 300 nM ClpXP in premix 1 without cell extract. The final concentration of ATP is 1.2 mM in premix 1, 21.2 mM in premix 2, 5.2 mM ATP in premix 3, and 9.2 mM in premix 4. Data are representative of at least three repeated experiments.



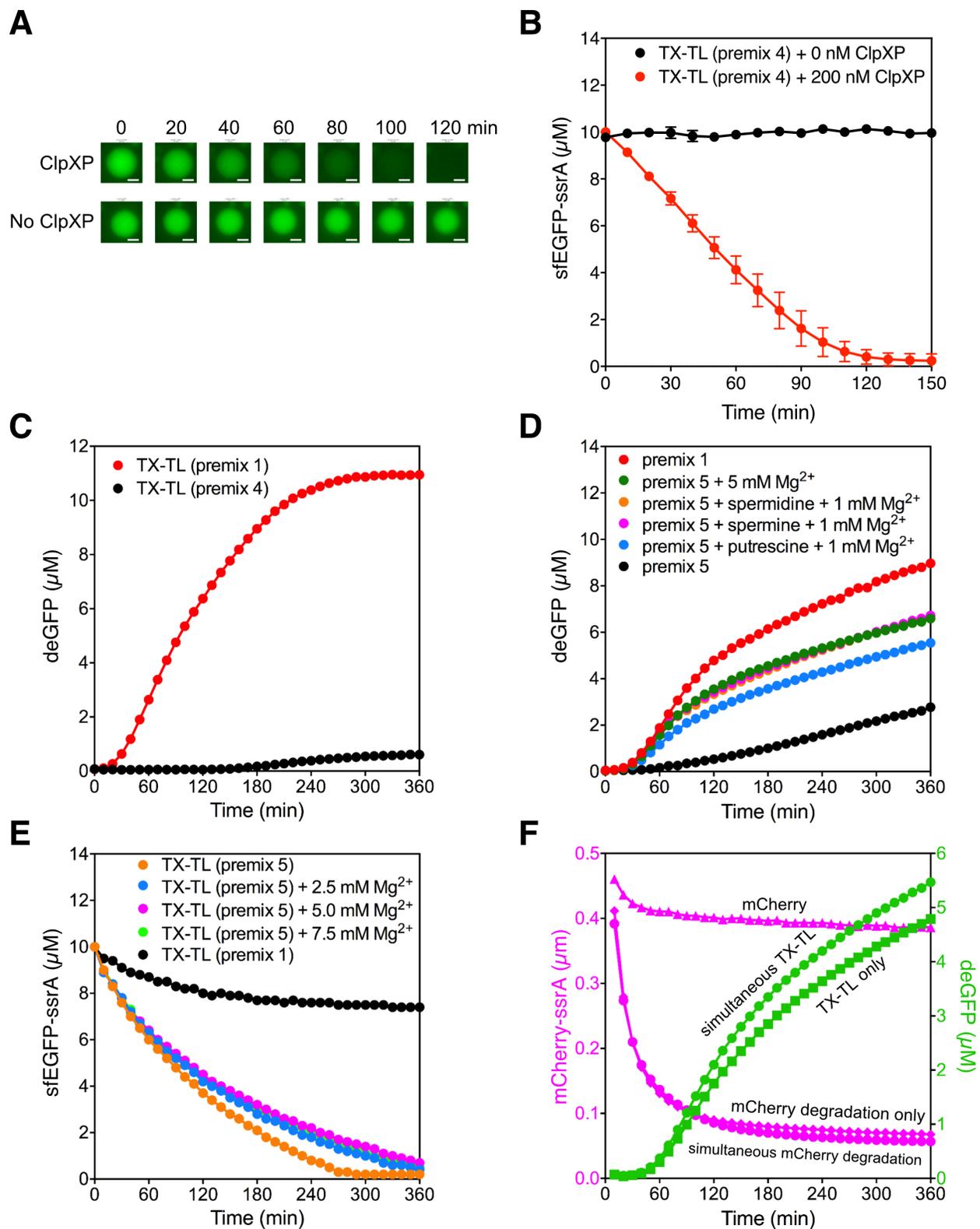
added either (1) 8 mM ATP or (2) 8 mM ATP plus extra energy regeneration components (16 mM PEP and 16 U/mL PK), 90 minutes after degradation reactions were started. As a control, we added water or dCTP (to account for Mg<sup>2+</sup> chelation by ATP) to identical reactions performed in parallel. Addition of ATP or ATP plus energy regeneration components greatly stimulated the degradation of sfGFP-ssrA and the reactions went nearly to completion (**Figure 2.3C**). These results showed that the concentration of ATP in the S30 cell extract must be increased for optimal ClpXP activity. *E. coli* cell extract contains some endogenous ClpX and ClpP<sup>48</sup>, so we next examined the amounts of additional ClpX and ClpP needed for optimal activity. Our results showed that degradation of 10 μM sfGFP-ssrA was efficient with extra 300 nM ClpX and extra 100-300 nM ClpP in the presence of premix 4 (final concentration of ATP, PEP, and PK was 9.2 mM, 41 mM, and 58 U/ml respectively) in the TX-TL system. The proteolytic activity was mainly limited by ClpX concentration but additional ClpP was required to speed up protein degradation when 300 nM ClpX was used<sup>52</sup>, suggesting that there is more endogenous ClpP than ClpX (**Figure 2.2C** and **2.2D**). To further improve the activity of ClpXP in the cell-free expression systems, we added the energy regeneration system and titrated the concentration of ATP in the degradation reaction. Our studies showed that the S30 extract with 9.2 mM ATP, 41 mM PEP and 58 U/mL PK resulted in efficient degradation of sfGFP-ssrA by ClpXP (**Figure 2.3D**). We also verified that the decrease in fluorescence intensity was due to protein degradation by ClpXP using Western blot (**Figure 2.2B**). It is also worth noting that although degradation was efficient and almost complete, it was still slower than in buffer control. Finally, all premixes were tested without adding purified ClpXP to serve as negative controls and see if they were activating the endogenous ClpXP. The results with different premixes but no ClpXP added were similar, which indicated that additional energy components will not activate endogenous ClpXP significantly.

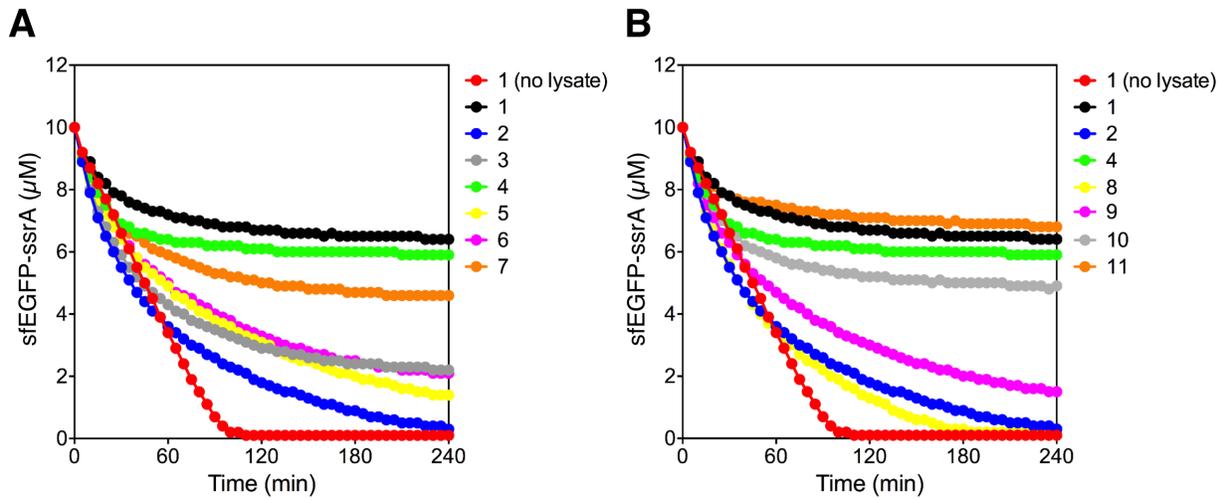
Since one of our long-term goals is to reconstitute genetic circuits in artificial cells, we tested the degradation of sfGFP-ssrA by ClpXP in lipid bound vesicles. The lipid used to form vesicles was 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine (POPC) and vesicles were formed by the inverse-emulsion method<sup>53-55</sup>. The vesicles contained S30 extract with 10  $\mu$ M sfGFP-ssrA and 200 nM ClpXP. Degradation of sfGFP-ssrA was monitored by fluorescence imaging using a spinning disk confocal microscope. ClpXP completely degraded sfGFP-ssrA in about 2 hours, demonstrating that ClpXP was still very active in the vesicles (**Figure 2.4A and 2.4B**). This is an important first step for building dynamic genetic circuits in vesicles.

Another essential feature to enable a dynamic genetic circuit is for the synthesis of target proteins simultaneous to protein degradation. This requires that conditions must be optimized for both protein degradation and protein synthesis. Therefore, we tested the efficiency of TX-TL under the optimized ClpXP degradation condition. The reporter protein synthesized is deGFP, an N-terminal truncated, C-terminal modified version of enhanced green fluorescent protein (eGFP), which has been shown to have better expression yield in TX-TL<sup>56</sup>. Unfortunately, TX-TL of deGFP was almost fully inhibited in the optimized condition for ClpXP degradation (**Figure 2.4C**). We tested whether degradation by ClpXP could occur at conditions that were more similar to the TX-TL reaction by titrating the concentration of ATP, and presence or absence of PEP and PK (**Figure 2.5A and 2.5B**). These experiments showed that efficient degradation of 10  $\mu$ M sfGFP-ssrA by ClpXP occurs when an additional 4-8 mM ATP is added (**Figure 2.5A**). Additionally, it is not essential to add extra PEP or PK to the ClpXP degradation reaction because they are present in sufficient concentrations in the premix buffer (**Figure 2.5B**). Thus, protein degradation by ClpXP in the S30 extract was efficient if the concentration of ATP is 7.2-9.2 mM (premix 1 has 1.2 mM ATP and we added extra 6-8 mM ATP).

## Figure 2.4 Degradation of sfGFP-ssrA in vesicles by ClpXP and optimizing both TX-TL and ClpXP activity

(A) Fluorescence images of sfGFP-ssrA degradation in a representative vesicle at 0, 20, 40, 60, 80, 100, 120 minutes. The degradation of 10  $\mu$ M sfGFP-ssrA in the TX-TL cell extract with 200 nM ClpXP added (top), or no ClpXP added (bottom) are shown. The scale bars represent 10  $\mu$ m. (B) Time course of ClpXP degradation in vesicles. The degradation kinetics of sfGFP-ssrA in the presence (red) or absence of ClpXP (black) in vesicles. The degradation kinetics of at least 10 vesicles were measured to calculate the standard deviations. (C) Protein synthesis in the optimized condition for ClpXP activity. Reactions were performed in TX-TL systems containing premix 1 (red) or premix 4 (black). (D) Effects of  $Mg^{2+}$  and polyamines on TX-TL reactions. Reactions were performed under standard condition with premix 5 and additional Mg-glutamate and/or polyamines: 5 mM Mg-glutamate (green); 2 mM spermidine and 1mM Mg-glutamate (orange); 0.5 mM spermine and 1 mM Mg-glutamate (magenta); 8 mM putrescine and 1 mM Mg-glutamate (blue). Control reactions were performed with premix 1 (red) and premix 5 (black). (E) Effect of additional  $Mg^{2+}$  on ClpXP activity in TX-TL systems. Control reaction (black) was performed in premix 1. Premix 5 was used for all the other reactions. The extra Mg-glutamate added to the reactions were: 0 mM (orange), 2.5 mM (blue), 5 mM (red), and 7.5 mM (green). (F) Simultaneous protein synthesis and degradation in TX-TL systems. All the reactions were performed under standard condition with premix 5 and extra 5 mM Mg-glutamate. Control reactions for protein degradation were performed with 500 nM mCherry-ssrA without ClpXP (pink triangles) or with 300 nM ClpXP (pink diamonds); TX-TL reaction control for deGFP synthesis was performed in the absence of mCherry-ssrA and ClpXP (green squares); Combined reactions included degradation of 500 nM mCherry-ssrA by 300 nM ClpXP (pink circles) and deGFP synthesis (green circles). Data are representative of at least three repeated experiments.





**Figure 2.5 Optimization of the ClpXP activity under TX-TL condition**

The concentration of ClpXP was 300 nM and the concentration of sfGFP-ssrA was 10 μM in all reactions. (A) and (B) Degradation of sfGFP-ssrA by ClpXP in the presence of different concentrations of ATP, PEP and PK. Control reactions in premix 1 with and without lysate are shown in red and black colors, respectively. The reactions are labeled 1 to 11 and the final concentrations of ATP, PEP, PK, and  $Mg^{2+}$  are shown in the Table 2.1. Data are representative of three repeated experiments.

**Table 2.1 Final concentration of the energy regeneration system and Mg<sup>2+</sup>**

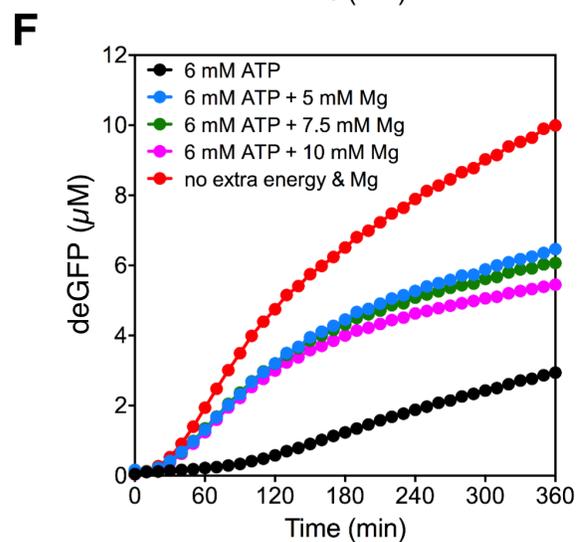
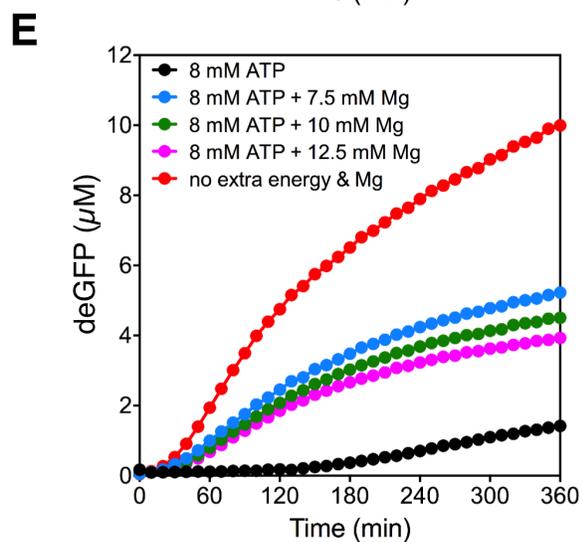
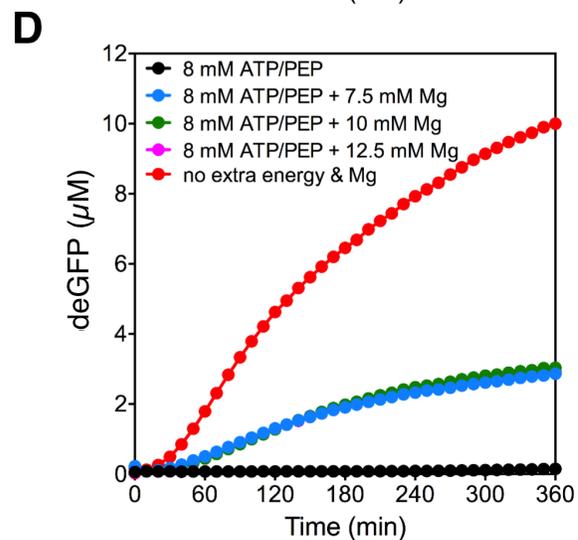
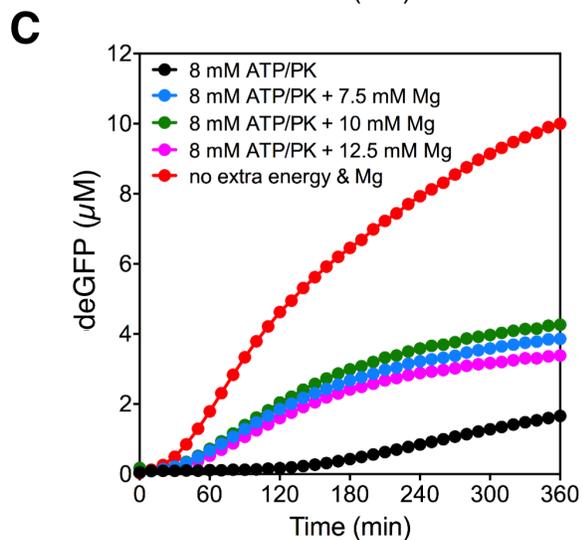
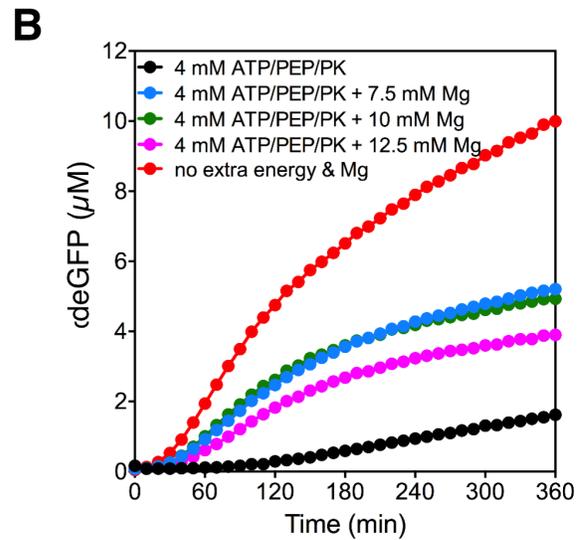
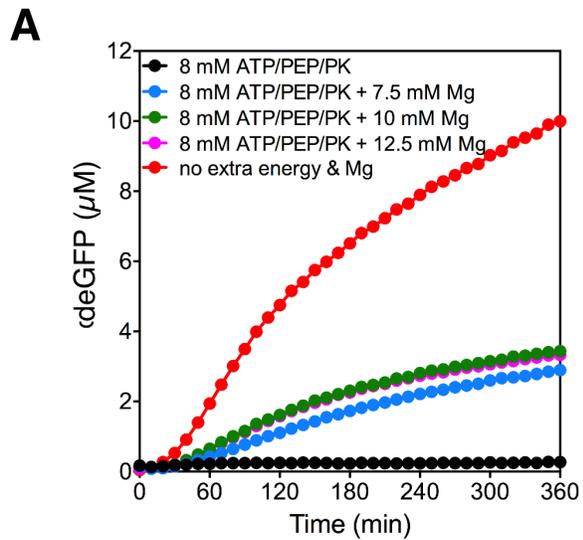
Reactions	ATP (mM)	PEP (mM)	PK (U/mL)	Mg <sup>2+</sup> (mM)
1	1.2	25	42	14.2
2	<b>9.2</b>	<b>41</b>	<b>58</b>	14.2
3	<b>5.2</b>	<b>41</b>	<b>58</b>	14.2
4	1.2	<b>41</b>	<b>58</b>	14.2
5	<b>9.2</b>	25	42	14.2
6	<b>7.2</b>	25	42	14.2
7	<b>5.2</b>	25	42	14.2
8	<b>9.2</b>	<b>41</b>	42	14.2
9	<b>9.2</b>	25	<b>58</b>	14.2
10	1.2	<b>41</b>	42	14.2
11	1.2	25	<b>58</b>	14.2

Changes from reaction 1 are indicated in bold

In contrast, the TX-TL reaction normally has 1.2 mM ATP, and 7.2-9.2 mM ATP inhibits the reaction. We reasoned that the extra ATP that we added to improve the activity of ClpXP was chelating  $Mg^{2+}$  thereby lowering the concentration of free  $Mg^{2+}$  causing the inhibition of TX-TL. Indeed, a recent study showed that nearly half of the  $Mg^{2+}$  in the cell is chelated by nucleotides <sup>57</sup>. We tested whether TX-TL activity could be restored by adding extra  $Mg^{2+}$ . Increasing the concentration of  $Mg^{2+}$  resulted only in a modest 30% increase in TX-TL compared to the control reaction without extra  $Mg^{2+}$  (**Figure 2.6A**). We titrated the concentration of ATP,  $Mg^{2+}$ , PEP, and PK to further optimize the TX-TL reaction (**Figure 2.6B to 2.6F**). We also showed that the polyamines, such as spermine, putrescine, and spermidine, could partly substitute for the  $Mg^{2+}$  in the TX-TL reaction (**Figure 2.7**). The addition of an extra 5 mM Mg-glutamate (final  $Mg^{2+}$  concentration is 19.2 mM), or 0.5 mM spermine and 1 mM Mg-glutamate, or 2 mM spermidine and 1 mM Mg-glutamate increased the activity of the TX-TL reaction to ~70% of the maximum. (**Figure 2.4D**). ClpXP was stored in buffer supplemented with 10% glycerol. Therefore, the optimized TX-TL reaction has 2.5% final glycerol concentration, which could be responsible for the ~30% inhibition in gene expression (**Figure 2.8A and 2.8B**). The optimized conditions (with extra 6 mM ATP and 5 mM Mg-glutamate) were then used to test the activity of ClpXP again. ClpXP activity was slightly slowed by the addition of extra 5 mM Mg-Glutamate (**Figure 2.4E**). Thus, we established new optimized conditions (7.2 mM ATP and 19.2 mM  $Mg^{2+}$  final concentrations) in which both the TX-TL reaction and protein degradation by ClpXP can occur simultaneously, although it is not the most efficient condition for each individual reaction. The optimized conditions were used to test whether the synthesis of deGFP and the degradation of mCherry-ssrA by ClpXP could occur simultaneously. As shown in **Figure 2.4F**, we observed the

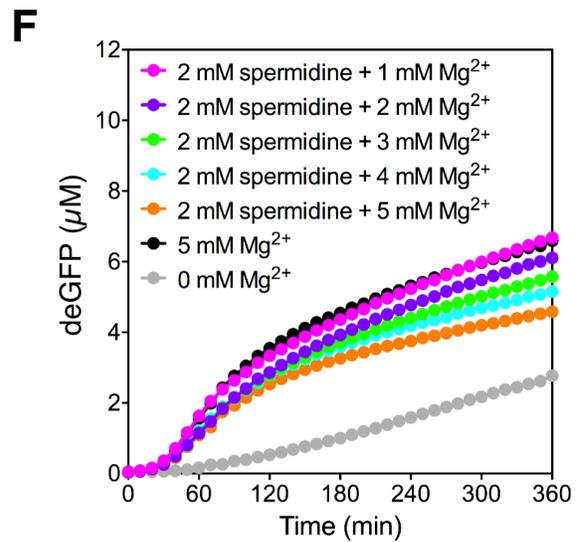
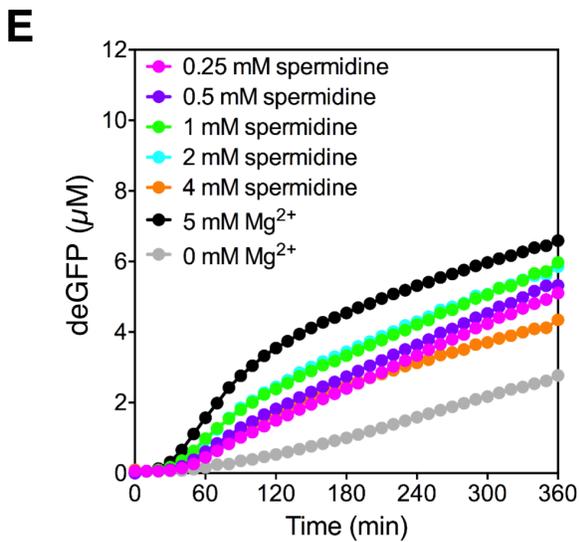
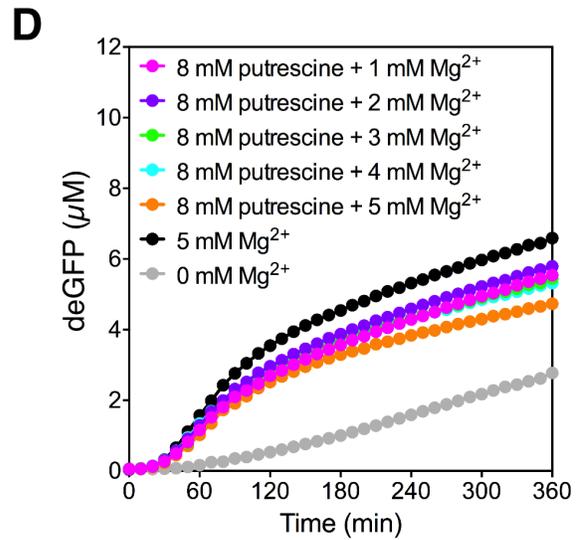
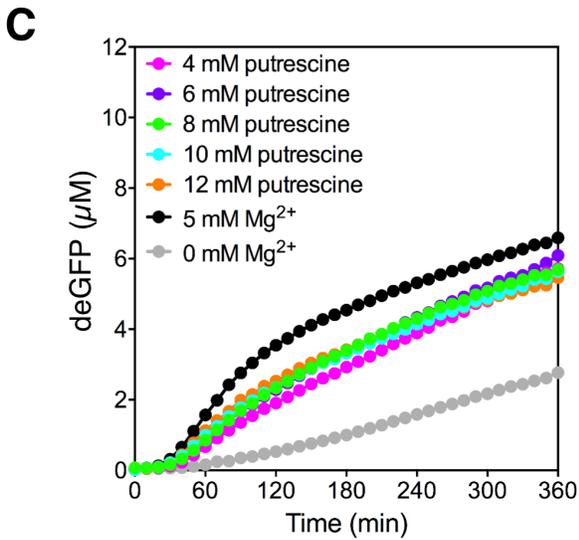
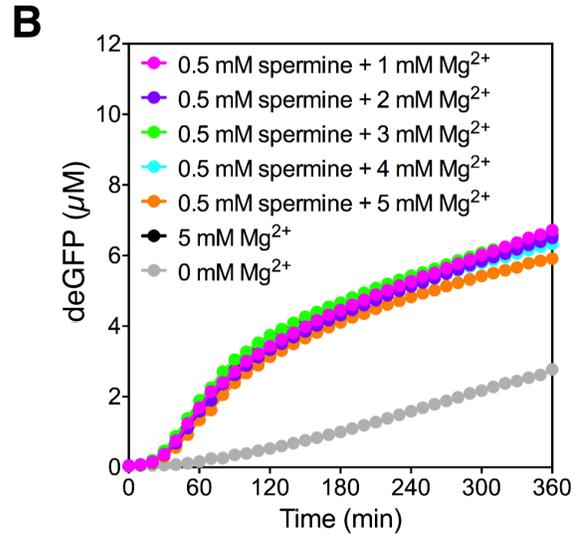
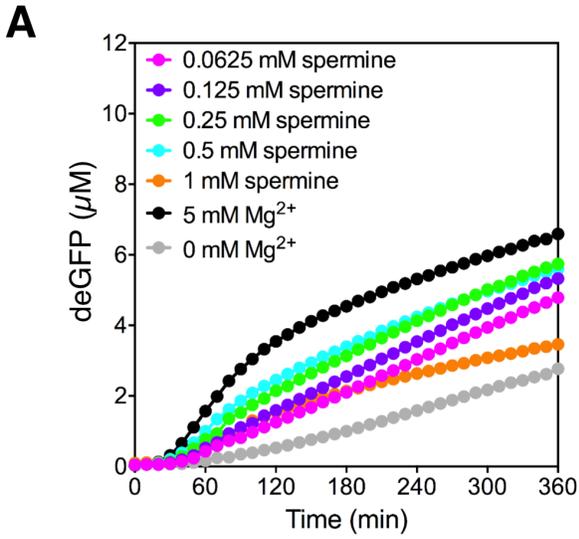
**Figure 2.6 Optimization of protein synthesis under the ClpXP degradation reaction condition by adding magnesium glutamate**

All reactions were carried out in premix 1 with different combinations of extra ATP, PEP, PK, and Mg-glutamate. Control reaction (red) were performed in the standard condition with premix 1 (final concentrations are 1.2 mM ATP, 25 mM PEP, 42 U/mL PK, and 14.2 mM total  $Mg^{2+}$ ). Optimization of protein synthesis were based on the extra ATP, PEP and PK added to premix 1. (A) TX-TL reactions with extra 8 mM ATP, 16 mM PEP, and 16 U/mL PK (final concentrations are 9.2 mM ATP, 41 mM PEP, and 58 U/mL PK). The concentrations of the extra Mg-glutamate added are: 0 mM (black), 7.5 mM (blue), 10 mM (green), and 12.5 mM (magenta). (B) TX-TL reactions with extra 4 mM ATP, 16 mM PEP, and 16 U/mL PK (final concentrations are 5.2 mM ATP, 41 mM PEP, and 58 U/mL PK). The concentrations of the extra Mg-glutamate added are: 0 mM (black), 7.5 mM (blue), 10 mM (green), and 12.5 mM (magenta). (C) TX-TL reactions with extra 8 mM ATP and 16 U/mL PK (final concentrations are 9.2 mM ATP, and 58 U/mL PK). The concentrations of the extra Mg-glutamate added to the reactions are: 0 mM (black), 7.5 mM (blue), 10 mM (green), and 12.5 mM (magenta). (D) TX-TL reactions with extra 8 mM ATP and 16 mM PEP (final concentrations are 9.2 mM ATP, and 41 mM PEP). The concentrations of the extra Mg-glutamate added to the TX-TL reactions are: 0 mM (black), 7.5 mM (blue), 10 mM (green), and 12.5 mM (magenta). (E) TX-TL reactions with extra 8 mM ATP (final concentration is 9.2 mM ATP). The concentrations of the extra Mg-glutamate added are: 0 mM (black), 7.5 mM (blue), 10 mM (green), and 12.5 mM (magenta). (F) TX-TL reactions with extra 6 mM ATP (final concentration is 7.2 mM ATP). The concentrations of the extra Mg-glutamate added are: 0 mM (black), 7.5 mM (blue), 10 mM (green), and 12.5 mM (magenta). Data are representative of three repeated experiments.



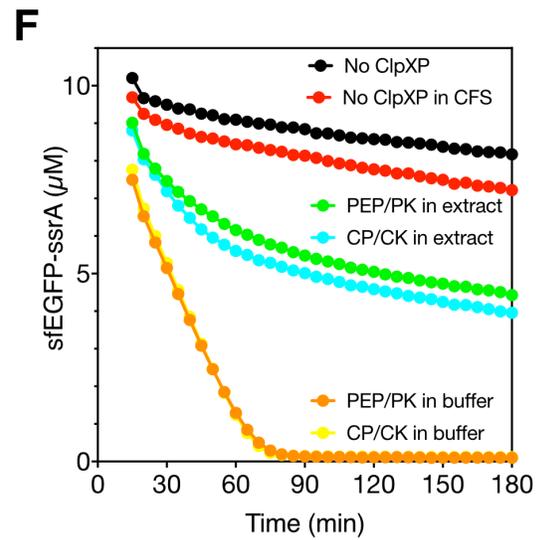
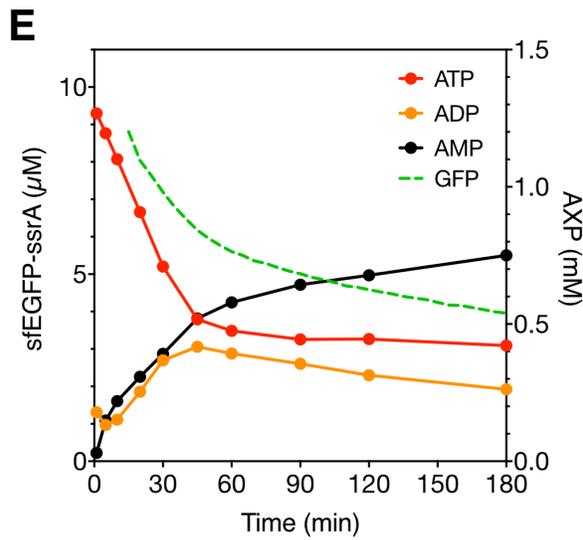
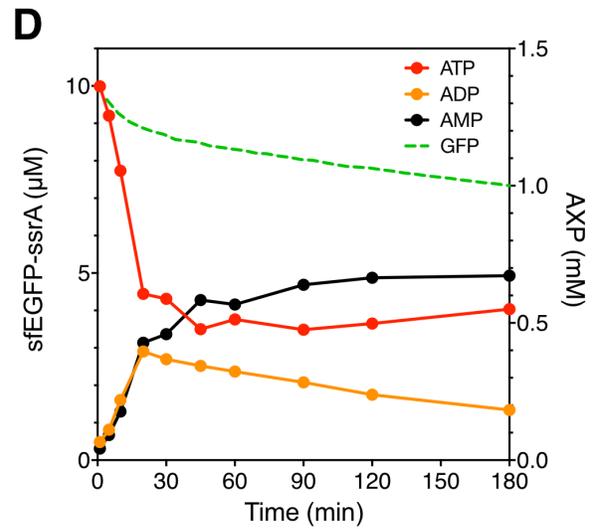
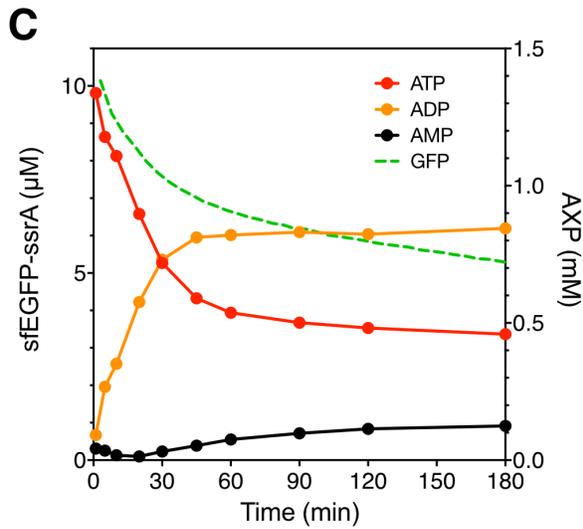
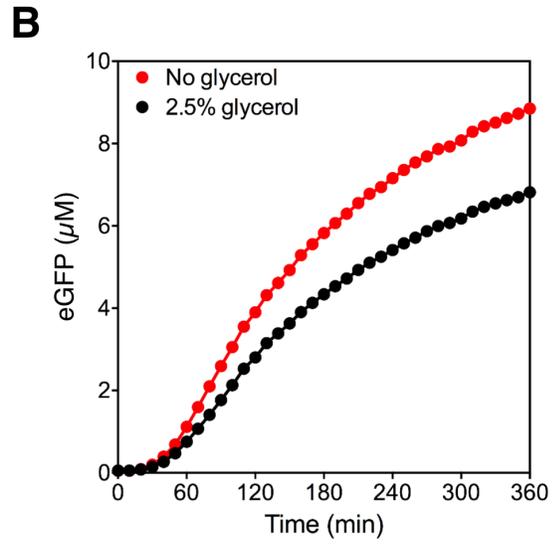
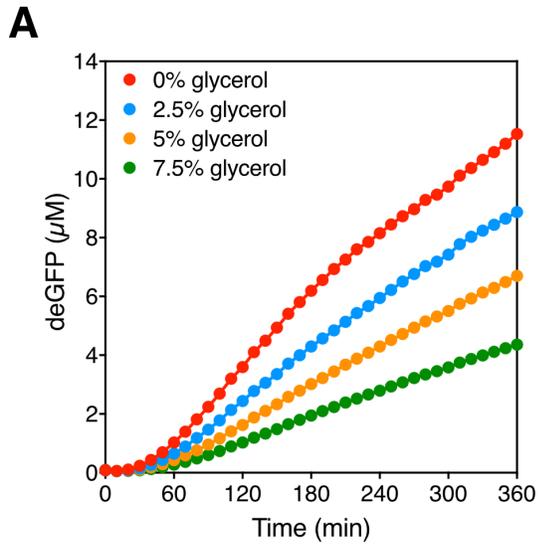
## Figure 2.7 Optimization of protein synthesis under ClpXP degradation condition by adding polyamines

All reactions were carried out in the standard condition with premix 5 (final concentrations are 7.2 mM ATP, and 14.2 mM total  $\text{Mg}^{2+}$ ) with different concentrations of extra polyamines or combinations of extra polyamine and Mg-glutamate. Control TX-TL reactions were performed in premix 5 with extra 5 mM (black) or 0 mM (grey) Mg-glutamate added. (A) Optimization with spermine. The concentration of spermine in the TX-TL reactions are: 0.0625 mM (magenta), 0.125 mM (purple), 0.25 mM (green), 0.5 mM (blue), and 1 mM (orange). (B) Optimization with 0.5 mM spermine and extra Mg-glutamate. The concentrations of the extra Mg-glutamate added to the TX-TL reactions are: 1 mM (magenta), 2 mM (purple), 3 mM (green), 4 mM (blue), and 5 mM (orange). (C) Optimization with putrescine. The concentration of putrescine in the TX-TL reaction are: 4 mM (magenta), 6 mM (purple), 8 mM (green), 10 mM (blue), and 12 mM (orange). (D) Optimization with 8 mM putrescine and extra Mg-Glutamate. The concentrations of the extra Mg-glutamate added to the TX-TL reactions are: 1 mM (magenta), 2 mM (purple), 3 mM (green), 4 mM (blue), and 5 mM (orange). (E) Optimization with spermidine. The concentration of spermidine in the TX-TL reactions are: 0.25 mM (magenta), 0.5 mM (purple), 1 mM (green), 2 mM (blue), and 4 mM (orange). (F) Optimization with 2 mM spermidine and extra Mg-Glutamate. The concentrations of the extra Mg-glutamate added to the TX-TL reactions are: 1 mM (magenta), 2 mM (purple), 3 mM (green), 4 mM (blue), and 5 mM (orange). Data are representative of three repeated experiments.



## **Figure 2.8 Effects of glycerol on TX-TL protein synthesis, and ATP consumption assays**

(A) Effect of glycerol on gene expression in the TX-TL system. The expression of deGFP in the TX-TL system with premix 1 in the presence of 0% (red); 2.5% (blue); 5% (orange); and 7.5% (green) glycerol. (B) Glycerol inhibits deGFP expression in the TX-TL system with premix 5. TX-TL reactions without glycerol (red), and with 2.5% glycerol (black). (C) (D) (E) are graphs showing protein degradation and ATP consumption assays. All reactions were carried out with 300 nM ClpXP and 10  $\mu$ M sfGFP-ssrA in different conditions for 3 hours. (C) The reaction was carried out in premix 1 without any pyruvate kinase. (D) The reaction was carried out without ClpXP in the TX-TL system without additional pyruvate kinase in the premix 1. (E) The reaction was carried out in the TX-TL system with premix 1, and CP/CK replaced PEP/PK for energy regeneration. (F) Comparison of protein degradation with PEP/PK and CP/CK energy regeneration systems in both buffer and in the TX-TL system. Data are representative of three repeated experiments.



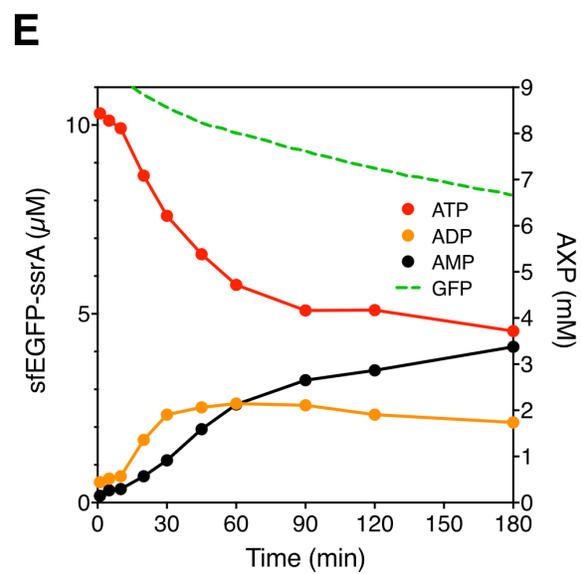
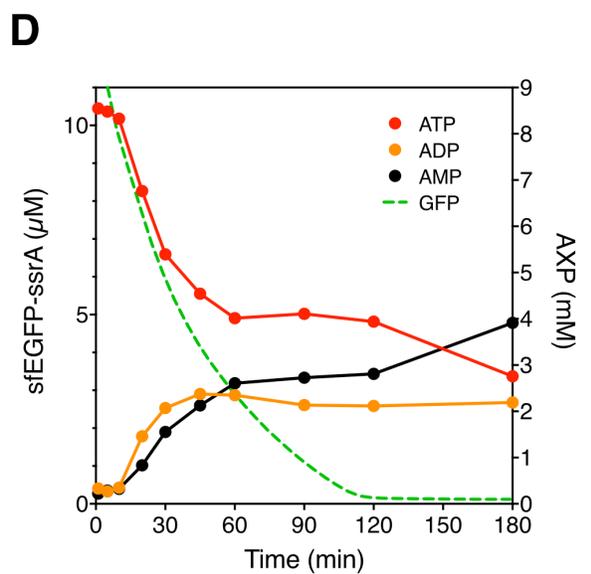
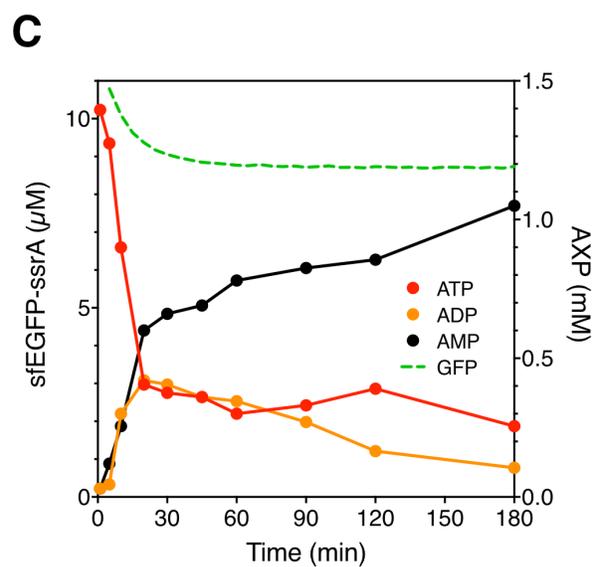
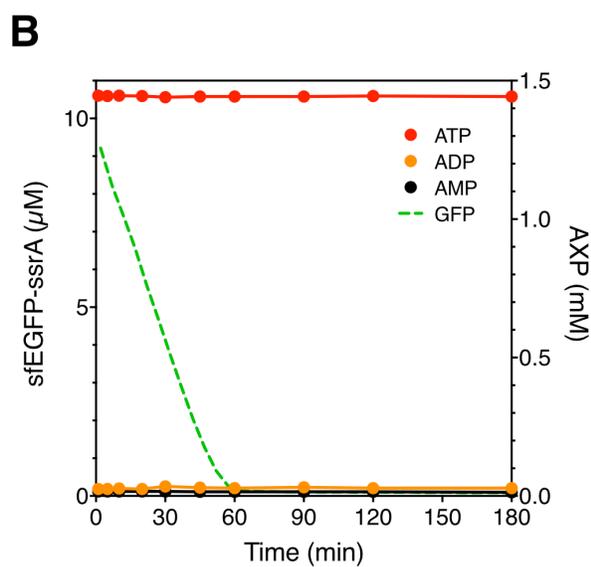
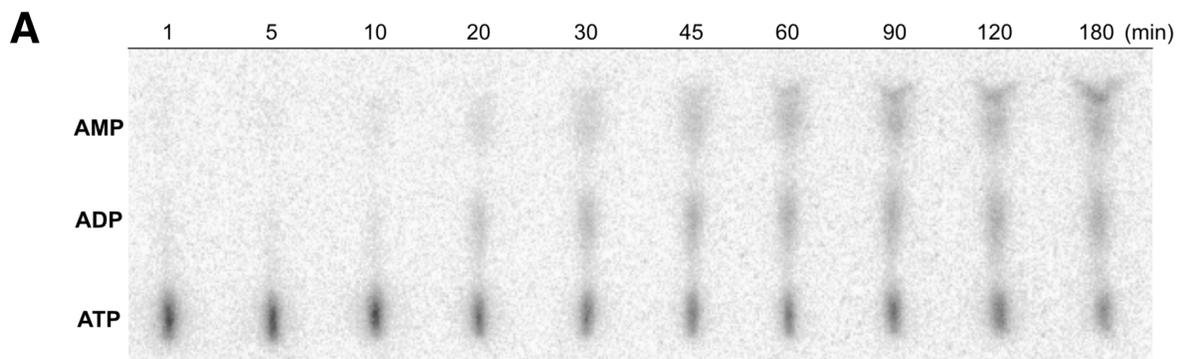
efficient synthesis of deGFP and the degradation of mCherry-ssrA in the same reaction confirming that the optimized conditions are suitable for the design and testing of gene circuits *in vitro*.

To better understand why the same energy regeneration system and ATP concentration, which was enough to saturate 300 nM ClpXP (**Figure 2.2E**), gave us very different ClpXP activity in buffers and the TX-TL systems, we established an ATP consumption assay using PEI-cellulose TLC to monitor the concentrations of [ $\alpha$ - $^{32}$ P]-ATP and its hydrolyzed products (**Figure 2.9A**). The degradation of sfGFP-ssrA by ClpXP was performed in the presence of [ $\alpha$ - $^{32}$ P]-ATP and the degradation reaction was continuously monitored by the decrease in the fluorescence intensity of sfGFP-ssrA. In parallel, an aliquot of the degradation reaction was used to monitor ATP consumption at fixed time points. As a control reaction, we used ClpXP with premix 1, which shows zero-order-like reaction kinetics for protein degradation, and a stable level of ATP concentration (**Figure 2.9B**), demonstrating that the energy regeneration system can maintain the ATP concentration. When we used premix 1 without any PK, the enzyme responsible for regenerating ATP by transferring a phosphate group from PEP to ADP, protein degradation slowed down quickly, and ATP concentration decreased as ADP concentration increased. This result shows that the stable level of ATP in the reaction with premix 1 is maintained by the ATP regeneration mechanism of PK (**Figure 2.8C**). Thus, ClpXP is very active in premix 1 which has 1.2 mM ATP and an efficient energy regeneration system with 25 mM PEP and 42 U/mL PK.

However, one of the problems we encountered originally is that the activity of ClpXP decreases drastically in the TX-TL system. We discovered that ATP was consumed rapidly in the TX-TL system, and simultaneously high levels of AMP accumulated (**Figure 2.9C and 2.8D**). This behavior was not observed in the buffer systems (**Figure 2.9B and 2.8C**). Even with

## Figure 2.9 ATP consumption and protein degradation by ClpXP

(A) A representative phosphor image of ATP consumption assay by ClpXP in the TX-TL system with additional 8 mM ATP, 16 mM phosphoenolpyruvate, and 16 U/mL pyruvate kinase. Samples were collected and quenched at 1, 5, 10, 20, 30, 45, 60, 90, 120, and 180 minutes after the reaction started. ATP, ADP, and AMP were separated on a PEI-cellulose TLC plate and quantified with a phosphor imager. The dark spots correspond to radioactive ATP, ADP, and AMP, respectively, from bottom to top. Spots were quantified and normalized by each lane and the ratio of each adenosine phosphate species were plotted. (B) (C) (D) (E) are graphs showing protein degradation and ATP consumption. All reactions were performed with 300 nM ClpXP (except in (E)) and 10  $\mu$ M sfGFP-ssrA in different conditions for 3 hours. (B) The reaction was performed in premix 1. (C) The reaction was performed in the TX-TL system with premix 1. (D) The reaction was performed in the TX-TL system with premix 4. (E) The reaction was performed without ClpXP in the TX-TL system with premix 4. Data are representative of at least three repeated experiments.



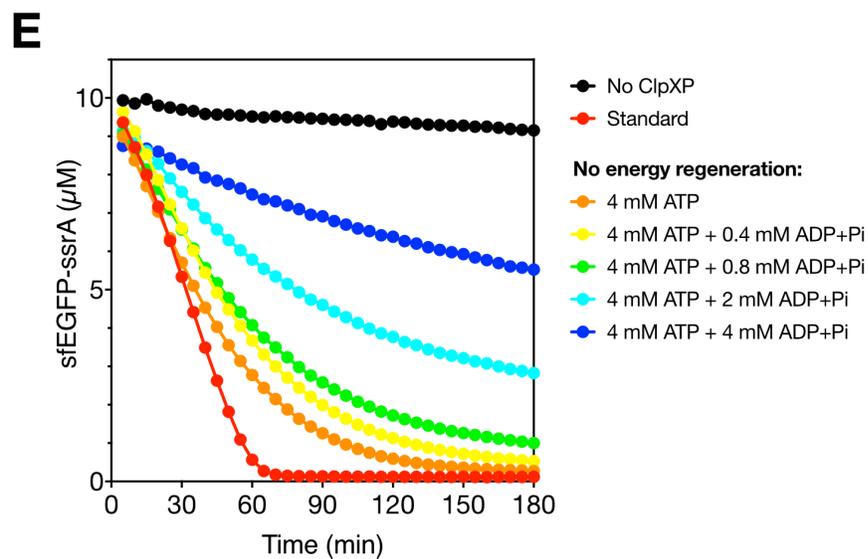
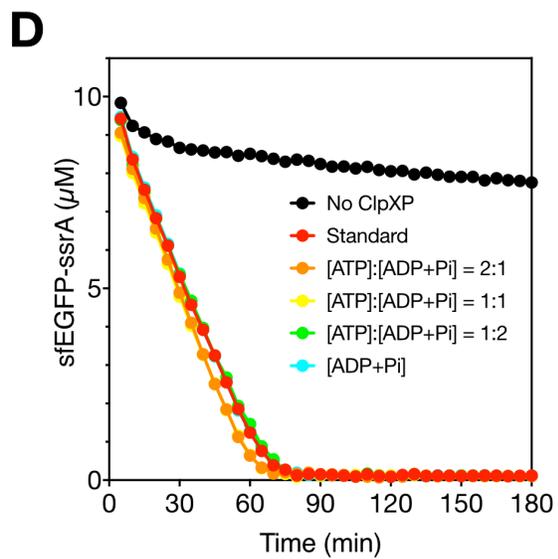
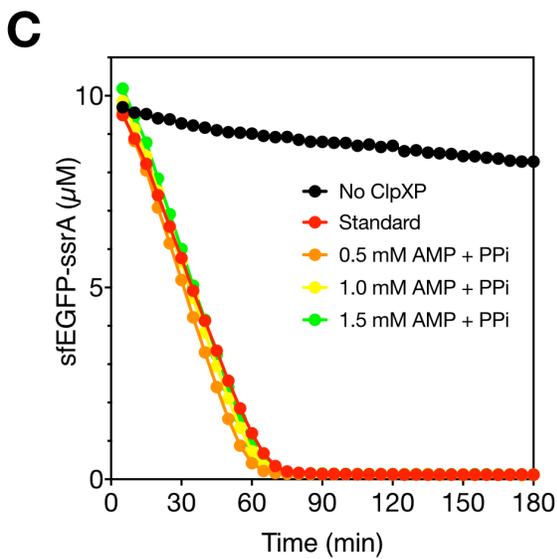
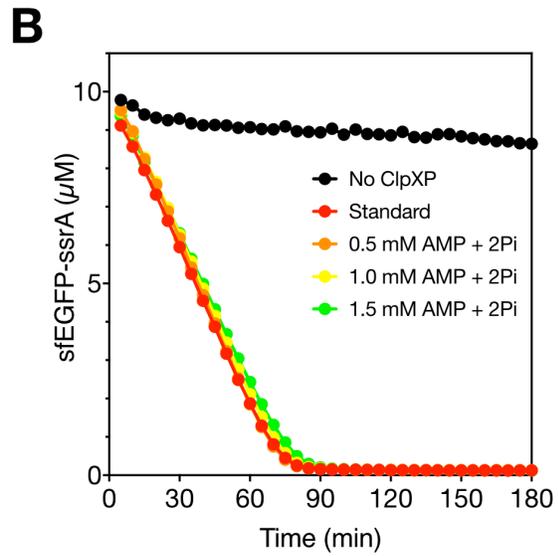
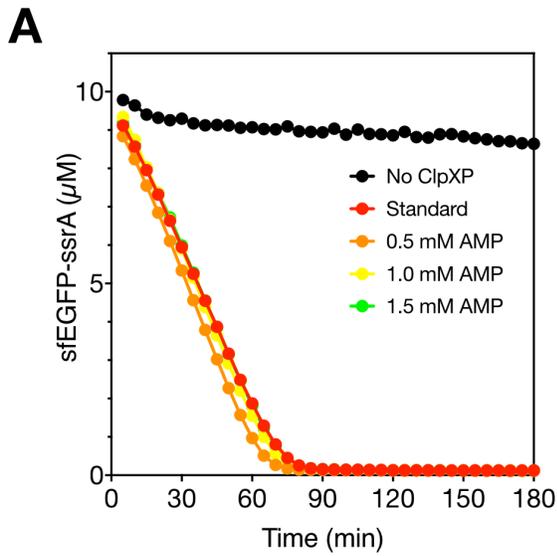
the creatine kinase-phosphocreatine energy regeneration system, ATP was consumed quickly, while AMP accumulated to high levels (**Figure 2.8E**). In the improved system with premix 4 (extra 8 mM ATP, 16 mM PEP, 16 U/mL PK were added, and final concentrations are 9.2 mM ATP, 41 mM PEP, and 58 U/mL PK), we can see that ATP concentration went down in the first hour and then reached a steady-state level concentration of about 4 mM (**Figure 2.9D**) while ADP concentration also reached a steady-state level, and sfGFP-ssrA is degraded at a reasonably rapid rate. ATP concentration went down again after two hours as AMP concentration went up. Interestingly, byproducts from ATP hydrolysis, such as AMP, inorganic phosphate, and pyrophosphate, do not inhibit ClpXP (**Figure 2.10A, 2.10B, and 2.10C**), but the concentration of ADP or the [ATP]/[ADP] ratio are more important for optimal activity in the buffer system (**2.10E**). Finally, we found that more than 50% of the ATP was consumed in the first hour in cell extract even without the addition of exogenous ClpXP (**Figure 2.9E**). Our results demonstrate that ATP is mostly consumed by enzymes present in the *E. coli* S30 cell extract. Therefore, for optimal degradation of proteins by ClpXP in the TX-TL system, the steady-state concentration of ATP must be maintained high enough by adding more ATP or a more efficient energy regeneration system.

## 2.4. Discussion

Protein degradation plays a key role for controlling gene expression by removing regulatory proteins that are no longer needed by the cell. Studies have shown that protein degradation is also important for spatially and temporally coupling synthetic genetic circuits in cells<sup>58,59</sup>. However, the activity of the endogenous proteases in cell extract is low and insufficient to build dynamic *in vitro* genetic circuits.

### Figure 2.10 Effects of AMP and ADP on protein degradation by ClpXP

All reactions were carried out in premix 1 supplying different concentrations of ATP and extra AMP or ADP. Control reactions were performed in premix 1 with 0 nM (black) or 300 nM ClpXP (red). (A) (B) (C) Effects of AMP on the activity of ClpXP. 0.5 mM (orange), 1 mM (yellow), and 1.5 mM (green) of AMP, AMP with two equivalents of phosphoric acid (Pi), or AMP with one equivalent of disodium pyrophosphate (PPi) were added to the reactions. (D) Effects of ADP on the activity of ClpXP. Total concentration of [ATP] + [ADP+Pi] was fixed at 1.5 mM, but the ratio of [ATP] to [ADP+Pi] was varied as 1.5 mM ATP, [ATP]:[ADP+Pi] = 2:1, [ATP]:[ADP+Pi] = 1:1, [ATP]:[ADP+Pi] = 1:2, 1.5 mM ADP+Pi. (E) Effects of [ATP]/[ADP] ratio on the activity of ClpXP. 4 mM ATP and 0.4 mM to 4 mM ADP+Pi were added to reactions in the absence of any energy regeneration systems. Initial rates were the linear regression slopes of the data in the first 30 min. Data are representative of three repeated experiments.



In a previous study, purified ClpXP was used to degrade ssrA-tagged eGFP in bulk and in vesicles, but the reaction was performed in a simple buffer<sup>33</sup>. In another study, the degradation rate by the endogenous ClpXP present in the *E. coli* S30 extract was analyzed with 1  $\mu$ M of deGFP-ssrA substrate<sup>39</sup>. The rate of degradation by the endogenous ClpXP was slow ( $v_o = 10 \text{ nM min}^{-1}$ ) and inadequate to create dynamic genetic circuits. To overcome this shortcoming, a subsequent study cloned the *clpP-clpX* genes from *E. coli* into a vector under the P70a promoter<sup>40</sup>. The vector was added at different concentrations to the TX-TL cell extract and preincubated to produce ClpX and ClpP proteins. Then eGFP-ssrA protein was added as a substrate and the reaction was monitored. A high rate of protein degradation was observed with 6 nM plasmid and 5  $\mu$ M eGFP-ssrA. Thus, the low activity of the endogenous ClpXP was overcome by synthesizing ClpXP in the TX-TL system, and this approach could be used to create simple gene circuits *in vitro*<sup>40</sup>.

However, synthesizing ClpXP from vectors by the TX-TL system has some disadvantages. First, determining the precise concentration of ClpXP in the reaction will require additional experiments. Second, the production of ClpXP in parallel by the TX-TL system may interfere with the function of complex genetic circuits. Therefore, we decided to purify the *E. coli* ClpXP protease and test its activity in a TX-TL cell extract. Consistent with previous studies, the activity of the purified ClpXP was high in the PD buffer but was very low when added to the standard S30 cell extract used for TX-TL<sup>33,39</sup>. Our results show that the low activity of the purified ClpXP is due to the low steady-state concentration of ATP in the *E. coli* S30 cell extract used for TX-TL, which has 1.2 mM ATP and an energy regeneration system (25 mM PEP and 42 U/mL PK) for supporting efficient protein synthesis. Importantly, increasing the final concentration of ATP to 9.2 mM, PEP to 41 mM, and PK to 58 U/mL in the TX-TL significantly improved the activity of the purified ClpXP. The degradation of sfGFP-ssrA in the S30 cell extract went to completion

under the optimized reaction conditions. This indicates that protein degradation by ClpXP consumes lots of ATP and the steady-state level of ATP in the cell extract is insufficient to support robust protein degradation. Indeed, previous studies have shown that ClpXP hydrolyzes about 600 ATPs to degrade one molecule of the I27 domain of human titin (121 residues)<sup>60</sup>. We also show that ClpXP is active in cell extract inside vesicles. The rate of sfGFP-ssrA degradation by ClpXP in vesicles is similar to the rate observed in bulk indicating that vesicles containing cell extract optimized for gene expression and protein degradation could be used to create artificial cells that execute dynamic genetic circuits.

We analyzed how energy consumption and regeneration affect the activity of ClpXP. First, we showed that our energy regeneration system using PEP/PK is efficient that we can even use ADP as the initial energy molecule (**Figure 2.10D**). We also confirmed that AMP and other byproducts do not directly inhibit ClpXP. So the correlation between the accumulation of AMP and the lower activity of ClpXP could be accounted by the decreasing total amount of energy molecules, including both ATP (for reaction) and ADP (for regenerating ATP). The accumulation of AMP is probably due to some enzymes in the cell extract, e.g. adenylate kinase, ADP-dependent NAD(P)HX dehydratase, and aminoacyl tRNA synthetases. Inhibiting these enzymes or finding a way to regenerate ATP or ADP from AMP could potentially help ClpXP to work better in the extract-based cell-free systems<sup>61,62</sup>.

Moreover, we noticed that the steady-state level of ATP concentration in the optimized conditions (**Figure 2.9D**) was about 4 mM in the TX-TL system; however, the rates of protein degradation by ClpXP were slower compared to the reaction with 1.2 mM ATP in the premix buffer (**Figure 2.9B**). This suggests that ATP concentration is not the only factor to affect the activity of ClpXP in our cell-free systems. Additional experiments showed that the rates of protein

degradation by ClpXP decreased when we added higher concentration of ADP to achieve lower [ATP]/[ADP] ratio in the buffer without energy regeneration system (**Figure 2.10E**). This is because protein degradation by ClpXP is mainly powered by the ATPase domains on the ClpX subunit and changes in the [ATP]/[ADP] ratio is expected to affect the activity of ClpXP significantly.

Our studies showed that protein degradation by ClpXP in the S30 extract needs a higher concentration of ATP than is optimal for the TX-TL reaction. We systematically optimized the concentration of ATP,  $Mg^{2+}$ , PK, and PEP to identify a new condition that is optimal for both ClpXP activity and the TX-TL reaction. The new optimized condition is 7.2 mM ATP, 19.2 mM  $Mg^{2+}$ , 25 mM PEP, and 42 U/mL PK. We used the optimized condition to successfully carry out simultaneous synthesis of deGFP and the degradation of mCherry-ssrA by ClpXP. Although the new conditions are not optimal for the individual reactions, we have identified conditions where the activity of ClpXP and the TX-TL reaction can be altered easily by changing the steady-state concentration of ATP or  $Mg^{2+}$ . Our results implicate that to establish dynamic genetic circuits containing multiple energy consuming biochemical reactions, such as protein synthesis or protein degradation by AAA<sup>+</sup> proteases, the supply and turnover of energy molecules should be accounted with care. We showed that it is possible to achieve high efficiency in both protein synthesis and protein degradation in extract-based cell-free systems. More complicated dynamic genetic circuits could be optimized based on the same principle, and we believe that the ability to modulate the rates of protein synthesis and protein degradation will be very useful for fine tuning the performance of *in vitro* genetic circuits.

## **2.5. Methods**

### **2.5.1 Cloning and protein purification**

*E. coli clpX*, *clpP*, and *sspB* genes were subcloned in pMCSG26 vector with C-terminal His-tag. sfGFP-ssrA and mCherry-ssrA were subcloned in pMCSG7 vector with N-terminal His-tag. Plasmid pACYC-FLAG-dN6-His encoding ClpX linked hexamer version was from Addgene (# 22143). *E. coli* BL21 (DE3) cells were used to overexpress and purify all proteins. Cells were grown at 37 °C to  $A_{600}$  0.5-0.8, chilled to 18 °C for ClpX, chilled to 30 °C for ClpP and SspB, chilled to 37 °C for sfGFP-ssrA and mCherry-ssrA. Protein overexpression was induced with 1 mM IPTG and the cells were grown for another 3-4 hours. The cells were lysed using a French press at 18000 psi and the proteins purified by Ni-NTA affinity chromatography as previously described<sup>63,64</sup>. Purified ClpP, and SspB proteins were dialyzed against storage buffer (50 mM Tris-HCl, pH 7.5, 200 mM KCl, 25 mM MgCl<sub>2</sub>, 1 mM DTT, 10% Glycerol). ClpX was purified further using a Superdex 75 16/60 column equilibrated in the above storage buffer. The concentrations of ClpX<sub>6</sub>, ClpP<sub>14</sub>, and SspB<sub>2</sub> were 1.4 μM (hexamer), 12 μM(tetradecamer), 57 μM(dimer), respectively. sfGFP-ssrA was dialyzed against buffer (20 mM K-Hepes, pH 7.5). mCherry-ssrA were stored in buffer containing 50 mM Tris-HCl, pH 7.5, 150 mM NaCl. The concentrations of sfGFP-ssrA and mCherry-ssrA were 120 μM and 125 μM, respectively. ClpX linked hexamer protein was purified by Ni-NTA affinity chromatography and further purified using a Superdex 200 16/60 size exclusion column (GE Healthcare) and stored in the buffer containing 50 mM Tris-HCl, pH7.6, 300 mM NaCl, 1 mM DTT, 10% glycerol. The concentration of the linked ClpX<sub>6</sub> was 3 μM. Protein concentrations were measured using Bradford method. All proteins were stored at -80 °C.

### **2.5.2 Fluorescent protein measurement and quantification**

The fluorescence of sfGFP-ssrA or deGFP was measured by using Tecan GENios microplate reader (excitation filter wavelength (band width): excitation filter 485 (20) nm;

emission filter wavelength (band width): 535 (25) nm; gain 45) unless stated otherwise. Fluorescence intensity was transformed into concentration unit using a standard curve made with our purified sfGFP-ssrA. Data was analyzed and graphs were created by Graphpad Prism software. All TX-TL and degradation reactions were performed at 30°C.

### **2.5.3 *In vitro* protein degradation assay**

Degradation assay of sfGFP-ssrA by ClpXP was performed with 10  $\mu$ L reaction volume in PD buffer (25 mM K-Hepes, pH 7.6, 200 mM KCl, 5 mM MgCl<sub>2</sub>, 0.032% Nonidet P-40, and 10% glycerol) with an energy regeneration system (4 mM ATP, 16U/mL creatine kinase, and 16 mM creatine phosphate). If SspB was used, the concentration of SspB was the same as sfGFP-ssrA and SspB was pre-incubated with sfGFP-ssrA for the degradation assay. 100 nM ClpXP (ClpX<sub>6</sub>:ClpP<sub>14</sub>=1:3) and 1  $\mu$ M sfGFP-ssrA was used for most degradation assay in PD buffer because we could analyze the function of SspB in the reaction and easily follow the time course of the reaction. Adenosine nucleotides, phosphates, and pyrophosphates were prepared as 100 mM stock solution at pH 7.0 and added separately to desired concentration. Adenosine 5'-monophosphate sodium salt (AMP, prod# A1752), adenosine 5'-diphosphate monopotassium salt dehydrate (ADP, prod# A5285), and sodium pyrophosphate decahydrate (PPi, prod# 221368) were purchased from Sigma Aldrich, and 85% phosphoric acid was purchased from Macron Chemicals (prod# 2796-05). The effect of ADP and AMP on ClpXP activity were analyzed in premix 1. Fluorescence of sfGFP-ssrA was monitored by microplate reader. 10  $\mu$ M sfGFP-ssrA degradation by 300 nM ClpXP was analyzed using SDS-PAGE.

### **2.5.4 $k_{cat}$ and $K_M$ measurement**

$k_{cat}$  and  $K_M$  measurement was performed in PD buffer by combining 50 nM ClpXP with varied concentration of sfGFP-ssrA. Degradation rates were calculated from the initial linear loss

of fluorescence. Curve fitting was performed using the Michaelis-Menten equation built in Graphpad Prism software. The experiment was repeated two times.

### **2.5.5 Macromolecular crowding effect on ClpXP activity**

Degradation was performed in PD buffer system with 1  $\mu$ M sfGFP-ssrA, 100 nM ClpXP, and various crowders, including up to 20% w/w PEG-400, PEG-8000, PEG-20000, Ficoll-70, and Ficoll-400. Reactions were performed in 10  $\mu$ L of volume on a 384-well plate and fluorescence of sfGFP-ssrA was monitored by using microplate reader.

### **2.5.6 Optimization of ClpXP activity in cell extract**

*E. coli* Rosetta 2 (DE3) cell extract was prepared as described previously<sup>30</sup> with minor modifications. Briefly, 1 g of cells were resuspended in 1.1 mL of S30A buffer. The cells were lysed by passing it twice through a French Press at 18000 psi. The extract was clarified by centrifugation at 12000g for 10 min. The clear supernatant was preincubated at 37°C for 80 min and the extract was clarified again at 12000g for 10 min. The clear supernatant was dialyzed against 1 L of S30B buffer containing 60 mM K-glutamate and 14 mM Mg-glutamate for 3 hours. The extract was clarified at 12000g for 10 min and was stored as small aliquots at -80°C. Degradation reactions were performed in 12  $\mu$ L final volume consisting of 4  $\mu$ L cell extract (final protein concentration is 7.5-8 mg/mL), 3  $\mu$ L ClpXP (final concentration 300 nM) or buffer contains 25 mM MgCl<sub>2</sub> and 200 mM KCl, 1  $\mu$ L sfGFP-ssrA (at different concentrations), 4  $\mu$ L premix 1 [42 mM K-Hepes, pH 8.2, 0.7 mM CTP and UTP, 1.2 mM GTP and ATP, 25 mM phosphoenol pyruvate (PEP), 42 U/mL pyruvate kinase (PK), 2.5 mM of 20 amino acid, 1.7% PEG-8000, 0.057 mM folinic acid, 0.17 mg/mL tRNA, 0.8 mM putrescine, 8 mM Magnesium glutamate, 120 mM potassium glutamate, all at final concentrations]<sup>65</sup>. The ClpX<sub>6</sub> to ClpP<sub>14</sub> ratio was 1:1 in the reactions with cell extract and 1:1.3 in the reactions without cell extract.

Additionally, the ClpXP degradation in cell extract was optimized by increasing the concentration of ATP, PEP and PK as follows: premix 2 (extra 20 mM ATP was added, and final ATP concentration is 21.2 mM), premix 3 (extra 4 mM ATP, 16 mM PEP, 16 U/mL PK were added, and final concentrations are 5.2 mM ATP, 41 mM PEP and 58 U/mL PK), and premix 4 (extra 8 mM ATP, 16 mM PEP, 16 U/mL PK were added, and final concentrations are 9.2 mM ATP, 41 mM PEP and 58 U/mL PK). Premix 5 (extra 6 mM ATP was added to premix 1. The final ATP concentration is 7.2 mM). The differences in the concentrations of ATP, PEP, PK, and  $Mg^{2+}$  in the premixes are shown in Table S1. 10  $\mu$ M sfGFP-ssrA degradation by 300 nM ClpXP was analyzed using Western blot with anti-His tag antibody purchased from Thermo Scientific (Cat# MA1-21315).

### **2.5.7 *In vitro* Transcription-Translation (TX-TL) of deGFP**

Standard TX-TL reactions were composed of 4  $\mu$ L S30 extract (final protein concentration is 7.5-8 mg/mL), 1  $\mu$ L pBEST-deGFP plasmid (Addgene plasmid #40019, final concentration is 8.3 nM), 4  $\mu$ L premix 1 (final concentration of ATP is 1.2 mM), and 3  $\mu$ L of ClpXP storage buffer (which contributes 50 mM KCl, 6.2 mM  $MgCl_2$  and 2.5% glycerol to the final reaction). Reactions were performed in 12  $\mu$ L of volume on a 384-well plate and fluorescence of deGFP was monitored by microplate reader.

### **2.5.8 Optimization of TX-TL system under ClpXP degradation condition**

Optimization of TX-TL system under the ClpXP degradation condition were performed in the presence of premix 5 (final concentration of ATP is 7.2 mM). Extra Mg-Glutamate or polyamines were added to restore the activity. Reactions were performed in 12 $\mu$ L volume containing of 4  $\mu$ L cell extract (final concentration of protein is 7.5-8 mg/mL), 4  $\mu$ L premix 5 (final concentration of ATP is 7.2 mM), and 1  $\mu$ L pBEST-deGFP plasmid (final concentration is

8.3 nM), 3  $\mu$ L of ClpXP storage buffer (which contributes 50 mM KCl and 6.2 mM MgCl<sub>2</sub> to the final reaction), and extra Mg<sup>2+</sup> and/or polyamines.

### **2.5.9 ClpXP proteolytic activity under optimized TX-TL condition**

ClpXP reactions were performed in 12  $\mu$ L final volume consisting of 4  $\mu$ L cell extract (final protein concentration is 7.5-8 mg/mL), 4  $\mu$ L premix 5 (final concentration of ATP is 7.2 mM), 3  $\mu$ L ClpXP (final concentration 300 nM) and magnesium glutamate (at different concentration), 1  $\mu$ L sfGFP-ssrA (final concentration is 10  $\mu$ M).

### **2.5.10 Simultaneous protein synthesis and degradation in TX-TL systems**

Protein synthesis and ClpXP reactions were performed in 12  $\mu$ L final volume consisting of 4  $\mu$ L cell extract (final protein concentration is 7.5-8 mg/mL), 4  $\mu$ L premix 5 with pBEST-deGFP plasmid (final concentrations are: 7.2 mM ATP, 19.2 mM Mg<sup>2+</sup> and 5 nM plasmid), 3  $\mu$ L ClpXP (final concentration 300 nM), and 1  $\mu$ L mCherry-ssrA (final concentration is 0.5  $\mu$ M). Reactions were performed at 30 °C on a 384-well plate and the fluorescence of deGFP (excitation filter wavelength (band width): 485 (20) nm, emission filter wavelength (band width): 535 (25) nm) and mCherry-ssrA (excitation filter wavelength (band width): 560 (10) nm and emission filter wavelength (band width): 612 (20) nm) were simultaneously monitored by using Tecan Infinite F200 microplate reader.

### **2.5.11 ATP consumption assay**

This method was modified from Rajagopal and Lorsch's protocol of ATP and GTP Hydrolysis Assay (TLC)<sup>66</sup>. PEI-cellulose TLC plate was cut to the appropriate size, pre-developed completely with water, and then dried. After setting up reactions with [ $\alpha$ -<sup>32</sup>P]-ATP, 3  $\mu$ L reaction solution was taken out and mixed with 3  $\mu$ L quenching solution (1 M HCOOH) at different time

points. In addition, 9  $\mu\text{L}$  reaction solution of each sample was taken out right after mixing and transferred to a 384-well plate for parallel measurement of protein degradation. When samples from all time points were collected, 1  $\mu\text{L}$  of each quenched sample was spotted on the marked places on the plate, and air dried. The TLC plate was developed with 0.5 M  $\text{K}_2\text{HPO}_4$ , pH 3.5, for 25 min, then dried. The TLC plate was wrapped in a clear plastic wrap and exposed to a storage phosphor screen (Amersham Bioscience). The screen was then scanned with Bio-Rad Personal Molecular Imager™ (PMI™) System or GE Healthcare Life Sciences Typhoon FLA 9500, and the data was analyzed with the Bio-Rad Quantity One® or ImageJ software.

#### **2.5.12 Vesicle preparation**

Vesicles were prepared with POPC (1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine) purchased from Avanti Inc. (cat# 850457P) as described previously<sup>53–55</sup>. To prepare the lipid suspension for the inner leaflet, POPC was first dissolved in chloroform, and the chloroform was subsequently evaporated with dry nitrogen in the hood affording a lipid film. The POPC film was then re-dissolved in heavy mineral oil (Fisher Scientific cat# O122-1) at 5 mg/mL, heated at 50°C, and sonicated for one hour in a water bath to evenly distribute the lipid. 10  $\mu\text{L}$  of degradation reaction (described in next section) was added to 100  $\mu\text{L}$  of POPC oil and vigorously mixed until it was a cloudy emulsion formed. The emulsion was then layered over 100  $\mu\text{L}$  of 0.5 M Tris-HCl, pH 7.5, centrifuged for 10 min at 9000 rcf at room temperature. The bilayer vesicles were formed as they passed through the water/oil phase and pelleted at the bottom. Oil and some buffers were discarded, and the vesicles were resuspended in the rest of buffer.

#### **2.5.13 Degradation assay in vesicles**

Degradation reaction was prepared by mixing 200 nM ClpXP and 10  $\mu\text{M}$  sfGFP-ssrA in optimized TX-TL system (plus extra 8 mM ATP, 16 mM phosphoenolpyruvate, and 16 U/mL

pyruvate kinase). Vesicles were then prepared via the inverse emulsion method using the degradation reaction mixture as previously described. 10  $\mu$ L of the resulting vesicle solution was added to a glass microscope slide and covered with a glass coverslip supported with vacuum grease. Fluorescence microscope images were acquired on a Yokagawa spinning disk system (Yokagawa, Japan) built around an Axio Observer Z1 motorized inverted microscope (Carl Zeiss Microscopy GmbH, Germany) with a 63x, 1.40 NA oil immersion objective to an ORCA-Flash 4.0 camera managed using Micromanager software. 10-20 vesicles were selected and pictures were taken every 10 min using a 488 nm, 100 mW OPSL laser, to excite the sfEGFP. Fluorescence intensity was quantified by using ImageJ software.

## **2.6. Acknowledgement**

Chapter 2, in full, is a reprint of published material: Xinying Shi, Ti Wu, Christian Cole, Neal K. Devaraj, Simpson Joseph. Optimization of ClpXP activity and protein synthesis in an *E. coli* extract-based cell-free expression system. *Sci. Rep.* 8, 587 (2018). The dissertation author was one of the primary authors of this paper.

# Chapter 3: Improving Energy Supplies in *E. coli* Cell-Free Systems with Class III Polyphosphate Kinase 2

## 3.1. Abstract

Energy is the key limiting factor in cell-free systems, which are useful tools to perform simple or complex biological reactions without the constraint of cell membrane. Typical energy regeneration machineries are built-in to the cell-free systems and they regenerate ATP by phosphorylating ADP using a specific pair of enzyme and high-energy phosphate-containing molecule. However, we showed that in extract-based cell-free systems AMP accumulates and reduces the overall availability of the energy currency, ADP and ATP. Here we present a strategy using a recombinant class III polyphosphate kinase 2 enzyme, which can catalyze the phosphorylation of both AMP and ADP with inorganic polyphosphate, to mitigate AMP accumulation and improve energy supply in *Escherichia coli* extract-based cell-free systems.

## 3.2. Introduction

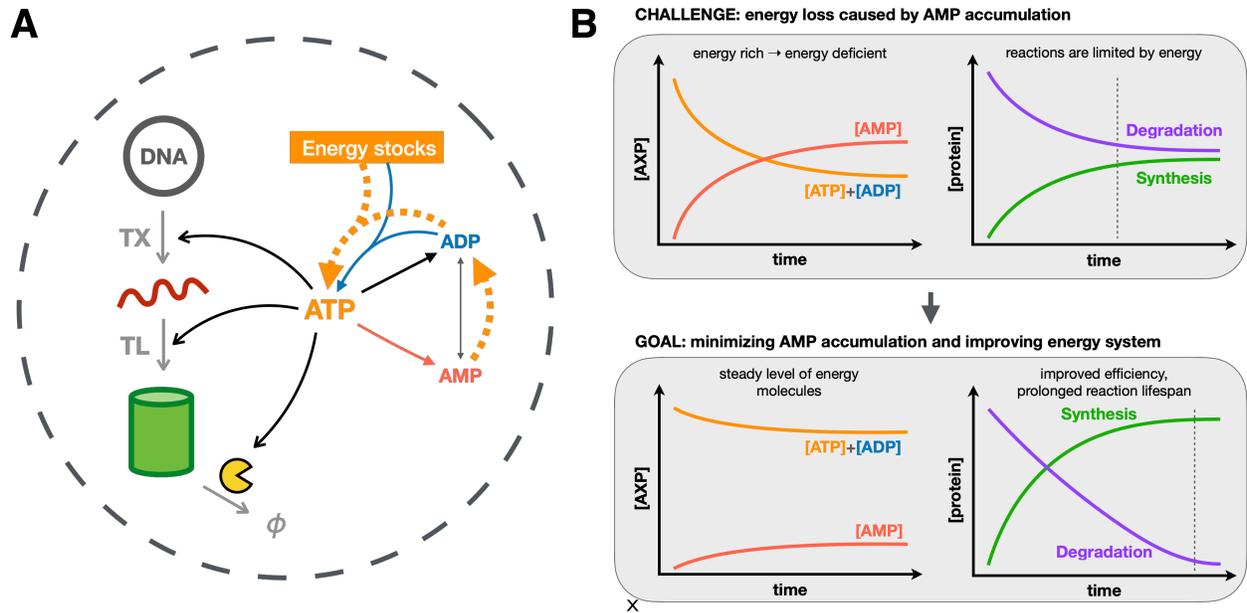
Throughout the history of the cell-free systems, the main focus in the field is to increase the efficiency of cell-free protein synthesis by improving the energy regeneration systems<sup>11,67</sup>. Unlike living organisms, cell-free systems cannot actively intake carbon sources from external environments to produce the energy molecules, adenosine triphosphate (ATP); instead, energy stock molecules (usually molecules with high energy phosphate bonds) and a corresponding energy regeneration mechanism (usually an enzyme or a metabolic pathway that can transfer the phosphate group from energy stock molecule onto adenosine diphosphate, ADP) must be added to regenerate ATP.

In our previous work<sup>34</sup>, we established an *Escherichia coli* extract-based cell-free systems that can perform efficient protein synthesis as well as degradation. The proteolytic machinery of choice is ClpXP, a bacterial ATPase-associated protease complex, which uses ATP to degrade proteins with specific polypeptide tag. ClpXP is energy-demanding and leads us to investigate the details of energy consumption profiles to optimize the proteolytic activity in the systems. We found that a significant amount of adenosine monophosphate (AMP) accumulated in the extract-based cell-free systems (**Figure 2.8 and 2.9**), which implies a reduction in the overall availability of energy currency, namely, ADP and ATP, and the commonly used energy regeneration systems based on creatine kinase or pyruvate kinase are not powerful enough to compete with the side reactions.

We investigated whether enzymatically transforming AMP back to ADP or ATP could reduce AMP accumulation and thus improve the energy supply in extract-based cell-free systems. As shown in **Figure 3.1**, cell-free systems commonly faced the challenge that adenosine nucleotides are transformed into AMP and accumulate in the system, which is often neglected in the traditional energy regeneration strategies. As a result, the amount of overall usable energy molecules ADP and ATP decreases as AMP accumulates, and hence energy-consuming biochemical processes such as protein synthesis and degradation slow down or even stop. To solve the problem, our goal is to incorporate phosphorylation pathways to regenerate ADP and ATP from AMP, and in theory, to improve energy supplies in the cell-free systems for better reaction efficiency and possibly longer lifespan.

Polyphosphate kinase 2 (PPK2) is a family of enzymes which catalyze the phosphorylation of adenosine nucleotides using polyphosphate (polyP) as substrate. Class I PPK2 can transfer one phosphate from polyP to an ADP and form an ATP; class II PPK2 can transfer one phosphate from

polyP to an AMP and form an ADP; the relatively new class III PPK2 was discovered in 2014 and it can catalyze the phosphorylation of both ADP and AMP<sup>68</sup>, which can potentially solve AMP accumulation and at the same time improve the energy regeneration in cell-free systems. It has been shown that polyphosphate and enzymes like polyphosphate:AMP phosphotransferase (PAP) can be used for energy regeneration for *in vitro* enzymatic reactions<sup>62,69</sup>. Itoh et al. have used PAP to enhance protein synthesis in commercial *E. coli* cell-free system in 2006, but in the following decade there was not much progress on applying polyphosphate and related enzymes to improve cell-free systems. Recently, the structural models from X-ray crystallography for class III PPK2 were reported, showing mechanistic insights on how this special class of enzyme works<sup>70,71</sup>. After that, several studies reported using class III PPK2 to regenerate energy for *in vitro* enzymatic reactions<sup>72-74</sup> and protein synthesis in PURE system<sup>75</sup>, yet reactions in extract-based cell-free systems remained unexplored. Here, we present an enzymatic strategy based on class III PPK2 enzyme from microorganism *Meiothermus ruber* (MrPPK2) to improve energy supply in *E. coli* extract-based cell-free systems and enhance cell-free protein synthesis and degradation.



**Figure 3.1 Outline of the challenge of AMP accumulation in Cell-Free Systems**

(A) Diagram of gene expression and energy regeneration in cell-free environments. (B) The challenge and goal are to reduce AMP accumulation and increase effective energy supplies in CFES, and hopefully it will improve the reaction efficiency and lifespan.

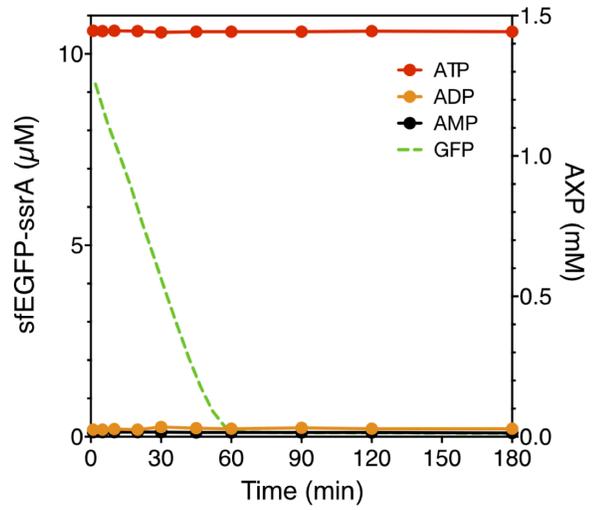
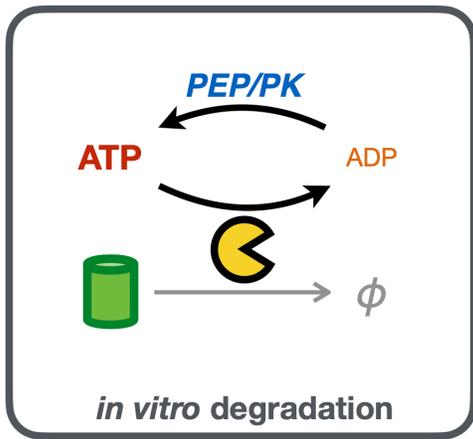
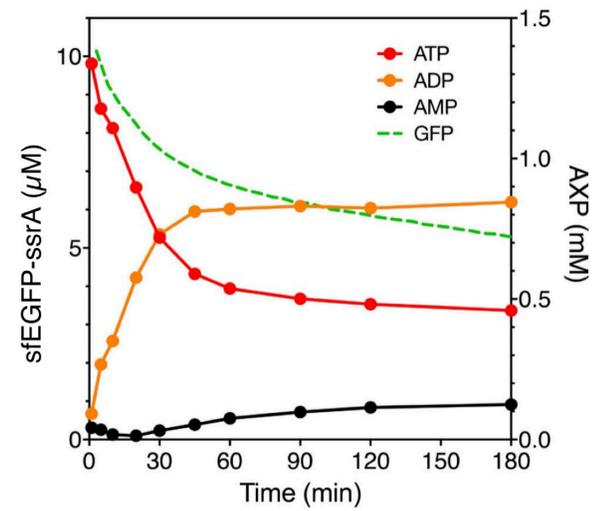
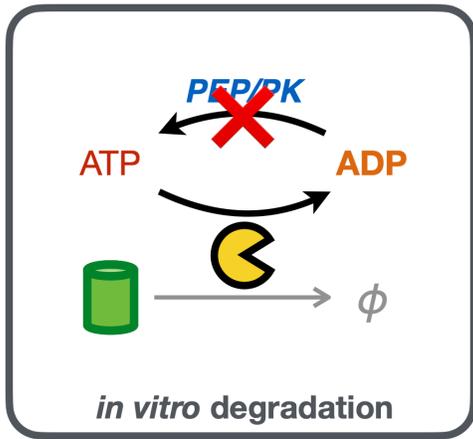
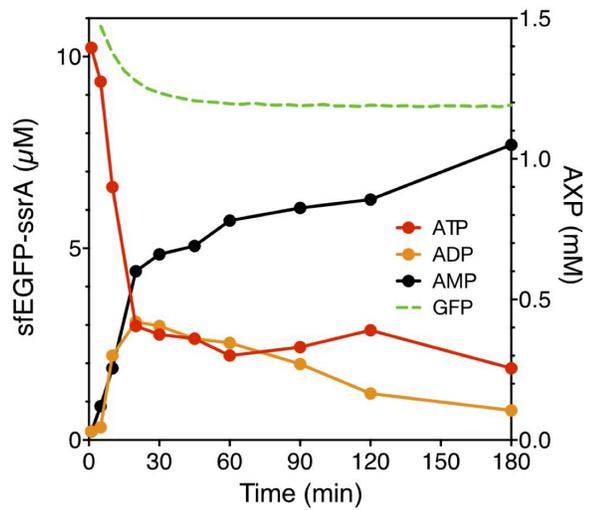
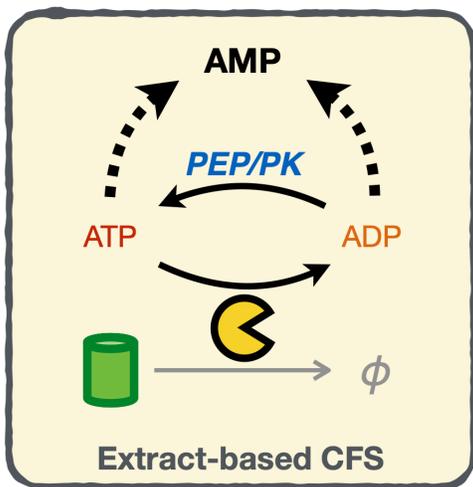
### 3.3. Results

We first found AMP accumulation problematic when incorporating ClpXP-based protein degradation into extract-based cell-free system<sup>34</sup>. In an extract-free buffer supplied with energy regeneration system, purified ClpXP protease could efficiently degrade ssrA-tagged superfolder GFP (sfGFP-ssrA) at constant rate until completion, and the energy consumption assay showed that ATP concentration was maintained at a high steady-state level (**Figure 3.2A**), showing the energy regeneration system using phosphoenol pyruvate (PEP) and pyruvate kinase (PK) to convert ADP to ATP was efficient. Energy regeneration was essential for ClpXP, because if it was removed, protein degradation was slowed down along with the increase in ADP concentration and decrease in ATP concentration (**Figure 3.2B**). In the extract-based cell-free system, ClpXP performed poorly, and it was consistent with the fact that ATP was consumed drastically in a short time, but in addition to the increased ADP concentration, AMP accumulation was more significant and in the end accounted for more than two third of all adenosine nucleotides (**Figure 3.2C**), suggesting that even though an efficient ADP-to-ATP conversion was included, AMP accumulation still diminished energy supply in the extract-based cell-free system.

We suspected that aminoacylation of transfer RNA (tRNA), or so-called tRNA “charging”, might be a major contributor to AMP accumulation because it’s a critical step in protein synthesis. It is a series of reactions that ligates a specific pair of amino acid and tRNA by their corresponding aminoacyl tRNA synthetase, which requires ATP and produces AMP and pyrophosphate (PP<sub>i</sub>) as byproducts. Using energy consumption assay to monitor the change of adenosine nucleotides<sup>34</sup>—ATP, ADP, and AMP—in the cell-free systems under standard condition, which contains a pyruvate kinase-based energy regeneration system but no expression plasmid DNA nor ClpXP, our studies showed that ATP is actively consumed and transformed into ADP and AMP until reaching a steady

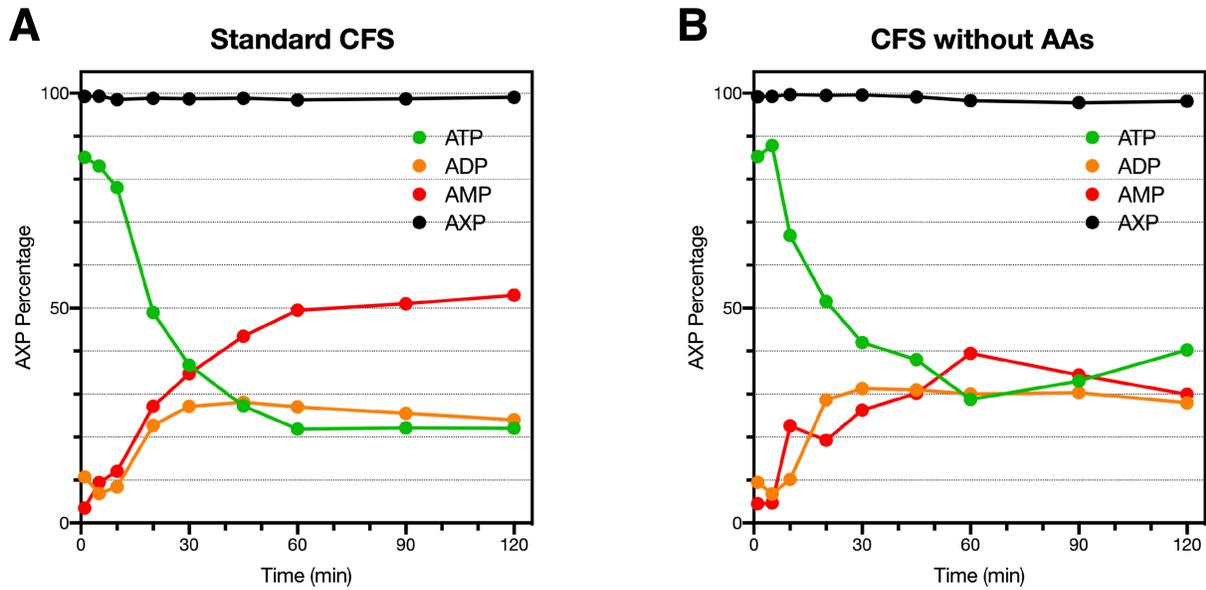
### **Figure 3.2 Energy consumption profiles of degradation by ClpXP**

Degradation of sfGFP-ssrA by ClpXP was monitored kinetically by measuring fluorescence and converted into concentration unit (green dashed line, left y-axis). The curve was overlaid with energy consumption profiles, showing concentration (right y-axis) of ATP (red solid line), ADP (orange solid line), and AMP (black solid line). Three reactions conditions were shown here: (A) *in vitro* protein degradation supplied with PEP/PK energy regeneration system. (B) *in vitro* protein degradation supplied without PEP/PK energy regeneration system. (C) Extract-based cell-free protein degradation, also supplied with PEP/PK energy regeneration system.

**A****B****C**

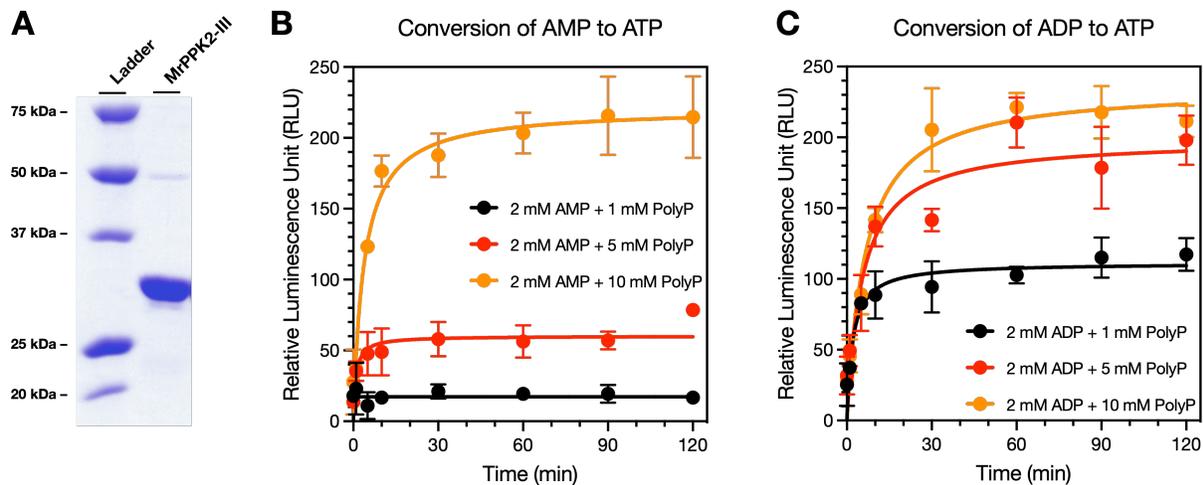
state after one hour (**Figure 3.3A**). This shows that even in the background there are strong ATP consuming activities in the cell-free system. By removing amino acids in the cell-free systems, which presumably removes tRNA aminoacylation reactions from the system, there is roughly a 40% reduction of AMP accumulation and increased amount of both ADP and ATP (**Figure 3.3B**). This confirms that tRNA aminoacylation does play a huge role in contributing to AMP accumulation. However, tRNA aminoacylation is essential for protein synthesis and cannot be removed from cell-free systems. Another strategy to combat AMP accumulation and improve the energy supply in cell-free systems is needed.

We chose to use a novel class of polyphosphate kinase 2 from microorganism *Meiothermus ruber* (MrPPK2) because it is one of the few class III PPK2 that has been characterized biochemically and structurally<sup>70</sup>. It consists of 267 amino acids with a molecular mass of 31.6 kDa, and forms higher multimers (8-, 12-, or 16-mer) in the presence of polyP in solution. Its activity of converting AMP to ADP is about 30 times higher than of ADP to ATP, which is a desirable feature for our purpose—mainly to bring AMP back to ADP so it could be recharged to ATP by either conventional energy regeneration systems or MrPPK2 itself. We cloned and expressed the recombinant MrPPK2 protein in *E. coli* (**Figure 3.4A**) and performed enzymatic assays to test its activity and optimize the reaction conditions (**Figure 3.4B and 3.4C**). When starting with 2 mM AMP, 2.5 times more of polyphosphate (concentration in terms of phosphate unit is estimated by average length reported by the manufacturer) is not enough to transform even a third of AMP to ATP, yet with 10 mM, namely 5 times more polyphosphate, the conversion from AMP to ATP is as fast as ADP to ATP.



**Figure 3.3 Aminoacylation contributes a significant part of AMP accumulation.**

ATP consumption profile of CFS without DNA or ClpXP. (A) Standard condition. (B) Removing amino acids from standard condition. Concentration is normalized in percentage to the initial combined concentration of all three adenosine nucleotides (AXP, black line). ATP percentage is shown in green line. ADP percentage is shown in orange line. AMP percentage is shown in red line.

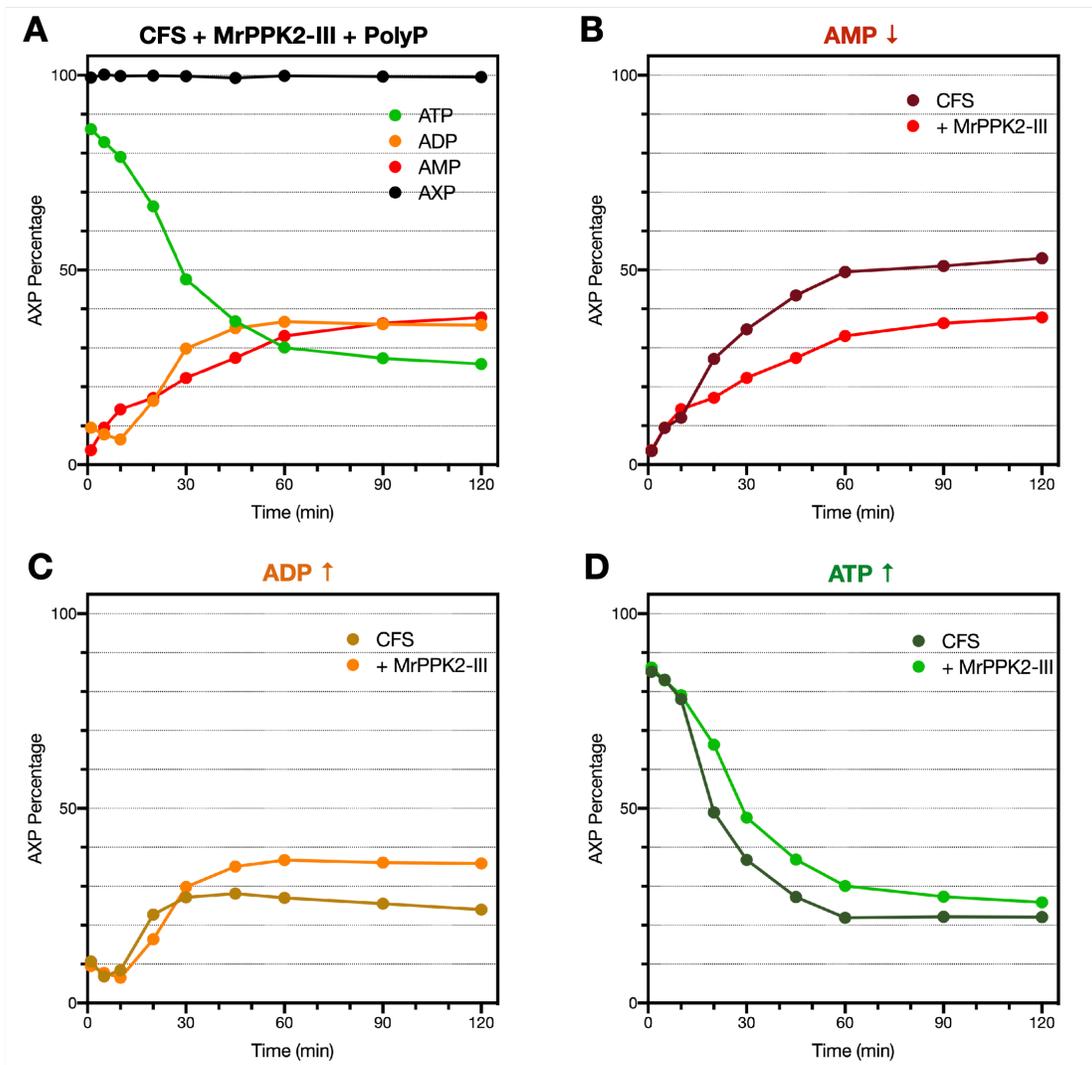


**Figure 3.4 Phosphorylation reaction catalyzed by MrPPK2**

(A) Purified MrPPK2 on SDS-PAGE stained with Coomassie Blue. (B) and (C) Enzymatic activity of MrPPK2 was tested with 2 mM of AMP or ADP and different concentration of polyP and final ATP concentrations were measured using luciferase-based ATP assay. (B) Conversion of AMP to ATP. (C) Conversion of ADP to ATP.

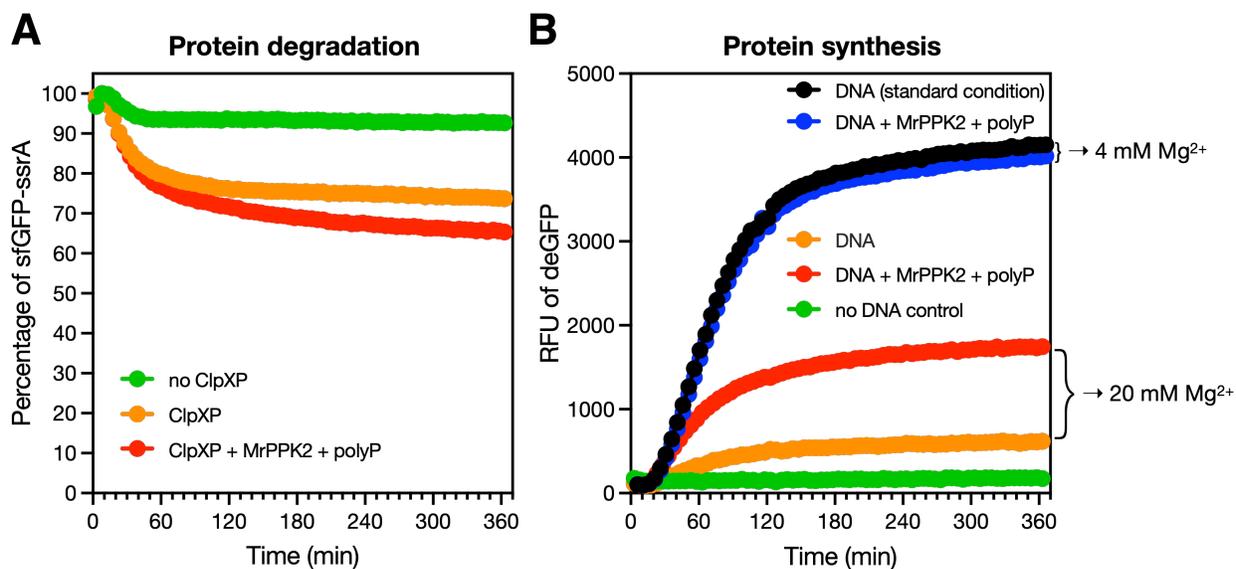
We then added 3  $\mu\text{M}$  of recombinant MrPPK2 and 10 mM polyP into cell-free system and performed energy consumption assay (**Figure 3.5**). We observed a roughly 30% decrease in accumulated AMP concentration and increased steady-state concentration of both ADP and ATP. The results showed that using enzymes like MrPPK2 to establish new energy regeneration pathways to convert AMP back into ADP and ATP could out-compete some unwanted ATP consuming activities and enhance the energy supply in the extract-based cell-free systems.

Finally, we added MrPPK2 and polyP into cell-free protein degradation with ClpXP and protein synthesis reactions (**Figure 3.6**). We observed a minor improvement in protein degradation (**Figure 3.6A**), which is consistent with the energy consumption profiles that ATP and ADP concentration are increased but not brought to a high level (**Figure 3.5**). In protein synthesis (**Figure 3.6B**), we saw a significant improvement of protein synthesis yield when adding MrPPK2 and polyP with 20 mM  $\text{Mg}^{2+}$  final concentration, which is the optimal condition for MrPPK2. However, the absolute yield is much lower than in standard protein synthesis condition reaction (**Figure 3.6B**, black lines) which only contains 4 mM  $\text{Mg}^{2+}$ . It is not surprising because cell-free biochemical reactions are sensitive to salt concentration, especially  $\text{Mg}^{2+}$  and  $\text{K}^+$ , which directly affect the association and dissociation of ribosomal complexes as well as the functionality of many other enzymes<sup>11,76</sup>. These results are consistent with previous reports that high concentration of magnesium may not be preferable for cell-free protein synthesis<sup>34,76</sup>. Nevertheless, adding MrPPK2 and polyP under standard protein synthesis condition did not really help the reaction (**Figure 3.6B**, blue lines).



**Figure 3.5 Energy consumption profile in CFS with MrPPK2 and polyP**

(A) Energy consumption profiles of cell-free systems with added MrPPK2 and polyP. Concentration is normalized in percentage to the initial combined concentration of all three adenosine nucleotides (AXP, black line). ATP percentage is shown in green line. ADP percentage is shown in orange line. AMP percentage is shown in red line. The data are separated into individual graphs in which (B) AMP (C) ADP and (D) ATP concentration profiles (dark color line) are plotted in comparison to standard condition (light color line, see **Figure 3.3A**).



**Figure 3.6 Effects of adding MrPPK2 into CFS**

(A) Cell-free protein degradation by ClpXP. The fluorescence signal of sfGFP-ssrA was kinetically monitored. No protease control is shown in green line. Protein degradation with ClpXP under standard condition was shown in orange line. Protein degradation with ClpXP in a system supplied with MrPPK2 and polyP was shown in red line. (B) Cell-free protein synthesis of deGFP. No DNA control is shown in green line. Protein synthesis under standard condition, which contains 4 mM Mg<sup>2+</sup>, is shown in black line. Protein synthesis under standard condition with added MrPPK2 and polyP is shown in red line. Protein synthesis under optimal condition for MrPPK2, which contains 20 mM Mg<sup>2+</sup>, is shown in orange line. Protein synthesis under high magnesium condition with added MrPPK2 and polyP is shown in red line.

### 3.4. Discussion

We showed that AMP accumulation is a major issue in energy supply in cell-free systems. Even in the PURE system<sup>16</sup>, of which the composition is well-defined, essential reactions such as tRNA aminoacylation will produce a lot of AMP as byproduct. Conventionally, myokinase is added and it transfers a phosphate group from an ATP to an AMP and produces two ADP. Recently, polyphosphate and PPK2 have been deployed to improve energy regeneration in the PURE system<sup>75</sup>. Extract-based cell-free systems are more complex and more side reactions could happen, so energy deficiency is more severe when energy-demanding enzymes were used (e.g. ClpXP, **Figure 2.3C**). Since they are generally more affordable and efficient than PURE systems<sup>23</sup> and widely used in the field, it will be beneficial to further improve the energy supply in the extract-based cell-free system by resolving the problem of AMP accumulation. Some reactions in cell extract which consume ATP or ADP and produce AMP might be avoidable by removing unwanted proteins during preparation processes of cell-free systems, engineering special bacterial strains in which unwanted enzymes are knocked out, or adding inhibitors. In this work we demonstrated a strategy that utilizes PPK2 enzyme to compete with the cellular activity, mitigating AMP accumulation and converting them back into ADP and ATP. It successfully improved the overall energy supply and enhanced cell-free protein degradation and synthesis to a certain degree. While purified PPK2 enzyme was used in this work, it could also be natively expressed in *E. coli* strain used for making cell-free systems, which would ease the preparation work for protein expression and purification.

There is room for improvement for this method. First, the enzyme we chose might not be powerful enough intrinsically. Extensive search and characterization for similar enzymes could provide better candidates. It is also possible to perform protein engineering or directed

evolution<sup>77,78</sup> to obtain novel enzymes which are more active and compatible with applicable reaction conditions (such as 4 mM Mg<sup>2+</sup> used in cell-free protein synthesis).

### **3.5. Methods**

#### **3.5.1 Cloning and protein purification**

Preparation of *E. coli* clpX, clpP, and sfGFP-ssrA proteins were described previously<sup>34</sup>. MrPPK2 gene fragment was purchased from GENEWIZ and subcloned into pMCSG26 vector with C-terminal His-tag. *E. coli* Rosetta (DE3) pLysS cells were used to overexpress proteins and the proteins were purified by Ni-NTA affinity. Protein concentrations were measured using Bradford method. All proteins were stored at -80 °C.

#### **3.5.2 Kinetic measurement of MrPPK2 activity**

The activities of MrPPK2 to convert of AMP to ATP and ADP to ATP were measured by Luciferase-based ATP Bioluminescent Assay (Sigma-Aldrich). A reaction mixture (500 µL) contained 50 mM Tris-HCl, pH 8.0, 20 mM MgCl<sub>2</sub>, 2 mM AMP or ADP (Sigma-Aldrich), 1-10 mM polyP (Acros organics, calculated as single phosphate residues) was initiated by adding MrPPK2 enzyme (1 µM final concentration) and incubated at 37 °C. 40 µL reaction solution was collected at specific time points (1, 5, 10, 30, 60, 90, 120, 180 min) and immediately diluted with 360 µL of 0.1 M EDTA solution to quench reactions. Subsequently, the diluted solution was passed through a 10-kDa MWCO spin column (Amicon Ultra-0.5, EMD Millipore) to remove enzymes, and stored on ice. Collected samples were then mixed with ATP Assay Mix following the manufacturer's instruction, and the luminescence was measured at 595 nm on a 96-well plate using a Tecan GENios microplate reader.

#### **3.5.3 ATP consumption assay**

This method was modified from Rajagopal and Lorsch's protocol of ATP and GTP Hydrolysis Assay (TLC)<sup>66</sup>. PEI-cellulose TLC plate was cut to the appropriate size, pre-developed completely with water, and then dried. After setting up reactions with [ $\alpha$ -<sup>32</sup>P]-ATP, 3  $\mu$ L reaction solution was taken out and mixed with 3  $\mu$ L quenching solution (1 M HCOOH) at different time points. In addition, 9  $\mu$ L reaction solution of each sample was taken out right after mixing and transferred to a 384-well plate for parallel measurement of protein degradation. When samples from all time points were collected, 1  $\mu$ L of each quenched sample was spotted on the marked places on the plate, and air dried. The TLC plate was developed with 0.5 M K<sub>2</sub>HPO<sub>4</sub>, pH 3.5, for 25 min, then dried. The TLC plate was wrapped in a clear plastic wrap and exposed to a storage phosphor screen (Amersham Bioscience). The screen was then scanned with Bio-Rad Personal Molecular Imager™ (PMI™) System or GE Healthcare Life Sciences Typhoon FLA 9500, and the data were analyzed with the Bio-Rad Quantity One® or ImageJ software.

#### **3.5.4 Fluorescent protein measurement and quantification**

The fluorescence of sfGFP-ssrA or deGFP was measured by a Tecan GENios microplate reader (excitation filter wavelength (band width): excitation filter 485 (20) nm; emission filter wavelength (band width): 535 (25) nm; gain 45) unless stated otherwise. Fluorescence intensity was transformed into concentration unit using a standard curve made with our purified sfGFP-ssrA. Data were analyzed and graphs were created by Graphpad Prism software. All TX-TL and degradation reactions were performed at 30°C.

#### **3.5.5 Cell-free Transcription-Translation (TX-TL) of deGFP**

TX-TL reaction was prepared based on a previously described protocol with modification<sup>23</sup>. In short, TX-TL reactions were prepared by mixing 4  $\mu$ L *E. coli* S30 extract (final protein concentration is 7.5-8 mg/mL), 1  $\mu$ L pBEST-deGFP plasmid (Addgene plasmid #40019,

final concentration is 8.3 nM), 4  $\mu$ L premix buffer [42 mM K-Hepes, pH 8.2, 0.7 mM CTP and UTP, 1.2 mM GTP and ATP, 25 mM phosphoenol pyruvate (PEP), 42 U/mL pyruvate kinase (PK), 2.5 mM of 20 amino acid, 1.7% PEG-8000, 0.057 mM folinic acid, 0.17 mg/mL tRNA, 0.8 mM putrescine, 4 mM Magnesium glutamate, 120 mM potassium glutamate, all at final concentrations], and 3  $\mu$ L of protein storage buffer (which contributes 50 mM KCl, 6.2 mM MgCl<sub>2</sub> and 2.5% glycerol to the final reaction). When applied, to the final concentration of 3  $\mu$ M MrPPK2, 10 mM PolyP and 20 mM Mg-glutamate were added. Reactions were performed in 10  $\mu$ L on a 384-well plate at 37°C and fluorescence of deGFP was monitored by a Tecan GENios microplate reader.

### **3.5.6 *In vitro* and extract-based cell-free protein degradation assay**

*In vitro* degradation assay of sfGFP-ssrA by ClpXP was performed with 10  $\mu$ L reaction volume containing premix buffer for TX-TL without adding *E. coli* S30 extract. Cell-free degradation of sfGFP-ssrA by ClpXP was performed in 10  $\mu$ L reaction volume in standard TX-TL condition, in which 10  $\mu$ M sfGFP-ssrA and 300 nM ClpXP were added. When applied, to the final concentration of 3  $\mu$ M MrPPK2, 10 mM PolyP and 20 mM Mg-glutamate were added. Fluorescence of sfGFP-ssrA was monitored by a Tecan GENios microplate reader.

### **3.6. Acknowledgement**

Chapter 3, in full, contains material currently being prepared for submission for publication. Ti Wu, and Simpson Joseph. Improving Energy Supplies in *E. coli* Cell-Free Systems with Class III Polyphosphate Kinase 2. The dissertation author was the primary author of this paper.

**Part II: Dual Site-Specifically Label a Protein Using  
Tryptophan Auxotrophic *Escherichia coli***

# Chapter 4: A Simple Method to Dual Site-Specifically Label a Protein Using Tryptophan Auxotrophic

## *Escherichia coli*

### 4.1. Abstract

Site-specifically labeling proteins with multiple dyes or molecular moieties is an important yet non-trivial task for biochemical and biophysical research, such as when using Förster resonance energy transfer (FRET) to study dynamics of protein conformational change. Many strategies have been devised, but usually done on a case-by-case basis. Expanded genetic codes provided a general platform to incorporate non-canonical amino acids (ncAA), which can also enable multiple site-specific labeling, but it is technically complicated and not suitable for some applications. Here we present a streamlined method that could enable dual site-specific protein labeling by using a tryptophan auxotroph of *Escherichia coli* to incorporate a naturally found tryptophan analog, 5-hydroxytryptophan into a recombinant protein. As a demonstration, we incorporated 5-hydroxytryptophan into *E. coli* release factor 1 (RF1), a protein known to possess two different conformations, and site-specifically attached two different fluorophores, one on 5-hydroxytryptophan and another on a cysteine residue. This method is simple, generally applicable, efficient, and can serve as an alternative way for researchers who want to install an additional labeling site in their proteins.

## 4.2. Introduction

Förster or fluorescence resonance energy transfer (FRET) is one of the most powerful and commonly used technique to understand the dynamics of conformational change in proteins<sup>79–81</sup>. FRET is a phenomenon that an excited “donor” fluorescent molecule transfers energy to an “acceptor” fluorescent molecule via a long-range non-radiative dipole-dipole coupling mechanism. The efficiency of energy transfer ( $E$ ) between two fluorescent molecules depends on the separation distance ( $r$ ) between donor and acceptor molecules with a relationship of inverse 6th-power law ( $E = 1/[1 + ((r/R_o)^6)]$ , where  $R_o$  is the Förster distance of this pair of donor and acceptor). Hence the measurement of the change in FRET efficiency can be used to reveal the dynamic structural information of proteins if two fluorescent dyes are carefully installed so the difference in the separation distance between two dyes can represent the conformational change of the protein of interest.

One of the technical hurdles for implementing FRET experiments is that it is non-trivial and sometimes even tricky to attach two or more different fluorescent dyes in a site-specific manner to a protein<sup>82</sup>. Classic site-specific labeling reactions most commonly utilize thiol-targeting functional groups, such as iodoacetamides and maleimides, amine-targeting functional groups, such as isothiocyanates, activated esters, sulfonyl chlorides, etc., and a few alcohol-targeting reagents. One practical concern is that thiol, amine, and alcohol groups are commonly present in multiple positions in a protein, which means that to achieve multiple site-specific labeling, extensive mutation and engineering may be required. Another strategy is to introduce non-canonical amino acids (ncAA) that possess bioconjugatable side chain into proteins<sup>83–85</sup>, so they are capable to perform a wider arrays of bioconjugation reactions<sup>86</sup>, such as click chemistry, tetrazine ligation, etc. Recent advances in expanded genetic code provide a platform to incorporate

non-canonical amino acids into proteins<sup>87,88</sup>. In short, this technology would assign a codon, usually one of the stop codons, to the ncAA of interest, then find and engineer an orthogonal pair of transfer RNA (tRNA) that can recognize the assigned codon and the corresponding tRNA synthetase that will only catalyze the ligation reaction between that specific ncAA and tRNA. While extremely powerful and versatile, this technology requires some specially engineered organisms and chemical components, and might not be suitable for some research projects, such as monitoring the conformational changes of RF1, which directly compete with the orthogonal tRNA for the stop codon.

Here we present a streamlined method that could enable dual site-specific protein labeling by incorporating a common tryptophan analog, 5-hydroxytryptophan, into a recombinant protein. This method utilizes a tryptophan auxotrophic strain of *E. coli*, which, when supplied with 5-hydroxytryptophan in a minimal growth media, can readily use them for protein synthesis. Combining with 5-hydroxytryptophan targeting bioconjugation chemistry and thiol-targeting maleimide dye, we can achieve dual site-specific labeling in a fascicle manner using the standard recombinant protein expression protocol.

It has been reported that 5-hydroxytryptophan can be specifically labeled with aromatic amines under mild, oxidative condition, such as using ferricyanide as oxidant<sup>89</sup>. This bioconjugation method utilizes the fact that electron-rich 5-hydroxyindole group on 5-hydroxytryptophan is more reactive toward oxidation than other aromatic groups commonly found in proteins, such as the side chain of tyrosine or phenylalanine, therefore under a milder oxidative condition the coupling reaction will preferably target 5-hydroxytryptophan but no other canonical amino acids. While the mechanism of this reaction was not fully elucidated, it was proposed that a C–N or C–C bond could be formed between two coupling partners via radical-generating one-

electron oxidation mechanism or through a further oxidized intermediate followed by nucleophilic addition of amine, as demonstrated in other hydroxyindole-containing molecules<sup>89–92</sup>.

As a model system to test our approach, we used *E. coli* release factor 1 (RF1). Class I release factor proteins, including RF1 and RF2 in bacteria and eRF1 in eukaryotes, are responsible for recognizing stop codon on mRNA at the A site of the ribosome and catalyzing the peptidyl-tRNA hydrolysis and the release of the newly synthesized polypeptide from the ribosome<sup>93–100</sup>. It is known that RF1 has two vastly different conformational states, “open” and “closed”<sup>81,99</sup>. Previously we have used transition metal ion FRET to study the dynamics of this conformation change and the role it plays during translation termination<sup>81</sup>. Here we will use RF1 to demonstrate direct site-specific dual labeling of two different fluorescent dyes, which may open up new opportunities for studying the structural dynamics of RF1 during stop codon recognition. More importantly, this simplified method can be used to dual site-specifically label any protein with fluorescent dyes, biotin, or other moieties.

## **4.3. Results**

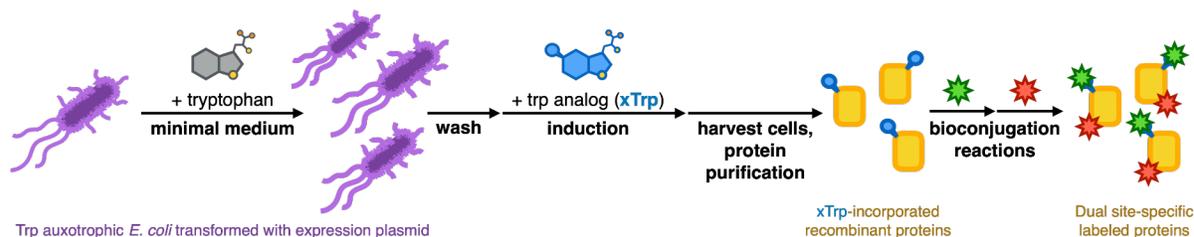
### **4.3.1. General schema of the protocol**

Our goal is to establish a recombinant protein expression protocol to introduce additional bioconjugation reaction sites via incorporation of non-canonical amino acids that can be easily implemented by laboratories equipped with standard molecular biology setup. To minimize technical complexity, residue-specific incorporation of non-canonical amino acids into proteins using amino acid auxotrophs was chosen<sup>84</sup>. To be specific, we chose to use the tryptophan auxotroph *Escherichia coli* strain BL21( $\lambda$ DE3)/NK7402 to incorporate tryptophan analogs<sup>101</sup>, because the general low occurrence of tryptophan in proteins could make this method more feasible, and there are several well-known analogs and corresponding bioconjugation tools

available. Briefly, the method is as follows: the Trp auxotrophic *E. coli* is transformed with an expression plasmid with the gene of interest tightly regulated by the pBAD promoter. The transformed cells are first grown in minimal media supplied with tryptophan. Once the cells grow to the desired density, the cells are spun down and washed to remove free tryptophan molecules in the media, and then resuspended with fresh media supplied with a tryptophan analog of choice and arabinose as the inducer. Finally, the cells are harvested, and the recombinant protein purified via fractionation and/or chromatographic methods. The tryptophan analog-incorporated protein could then be labeled with dyes using various bioconjugation methods (**Figure 4.1**).

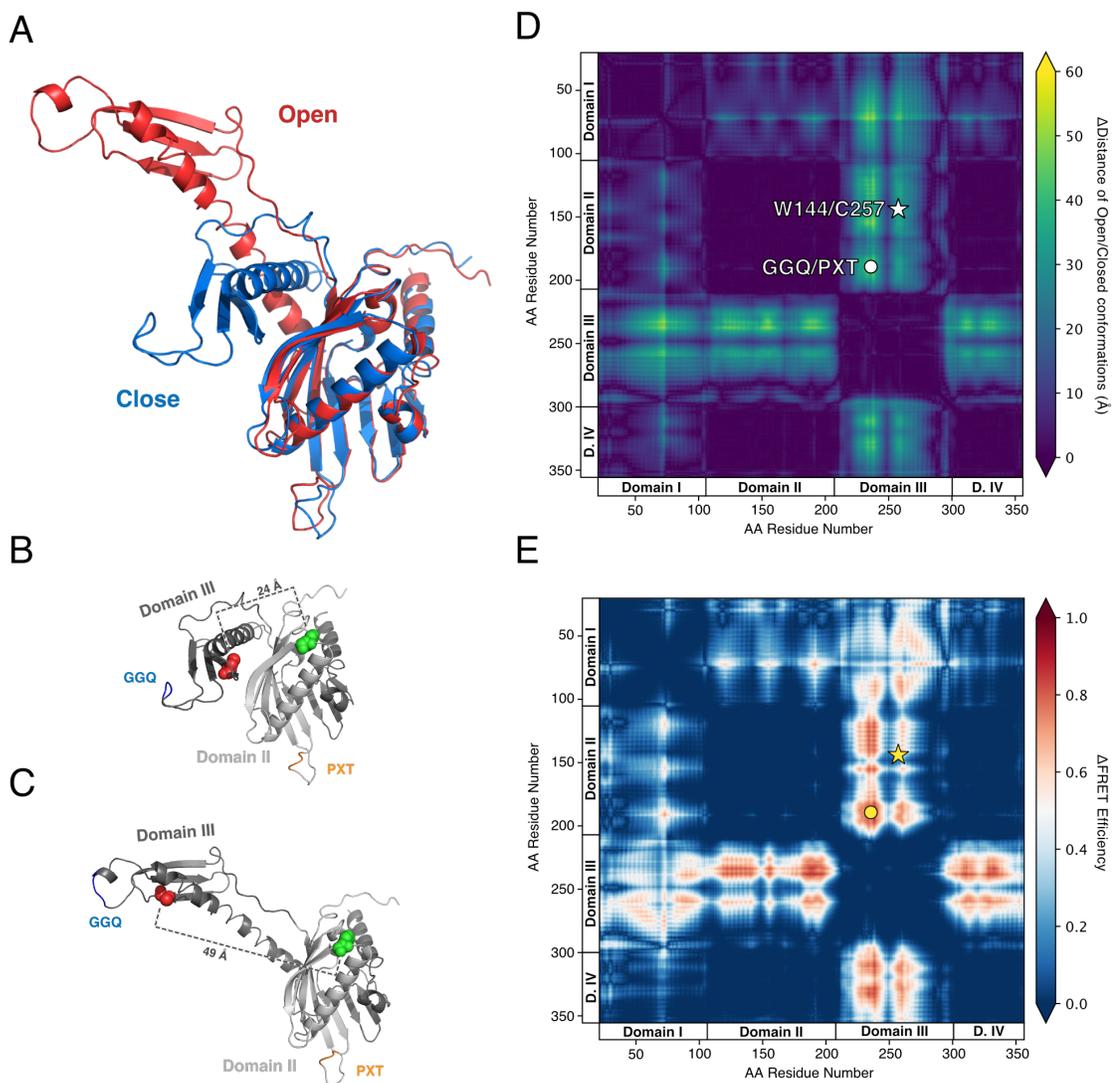
#### **4.3.2. Design of single cysteine single tryptophan RF1 (scswRF1)**

Our model protein is *E. coli* RF1, which shows two distinctive conformations—open and closed (**Figure 4.2A**). A crystal structure of *E. coli* RF1 complexed with PrmC methyltransferase (PDB code: 2B3T) served as the template for the closed conformation<sup>102</sup>, and a cryo-EM structure of *E. coli* RF1 in the translation termination complex (PDB code: 6OSK) was the template for the open conformation<sup>99</sup>. To find two labeling sites whose separation distance can reflect the conformational change, it is natural to use GGQ motif and anti-codon PXT motif as reference points and look for potential sites in the domains they are located in, namely, Domain III and Domain II, respectively. Fortunately, out of three cysteine sites and two tryptophan sites present in wild-type *E. coli* RF1, one of the cysteines (C257) is located in Domain III and one of the tryptophan (W144) is located in Domain II. Based on the model, the estimated separation distances of these two residues are 24Å and 49Å in closed and open states, respectively (**Figure 4.2B** and **4.2C**).



**Figure 4.1 Schema of incorporation of tryptophan analogs into recombinant protein using Trp auxotrophic *E. coli* for dual site-specific labeling**

Transformed cells are first cultured in minimal medium supplied with regular tryptophan until the desired cell density. After a few rounds of washing to remove free tryptophan, cells are resuspended in a new growth medium supplied with the tryptophan analog along with the inducer molecule for protein over-expression. Recombinant protein can then be harvested, analyzed, and site-specifically labeled with compatible bioconjugation reactions.



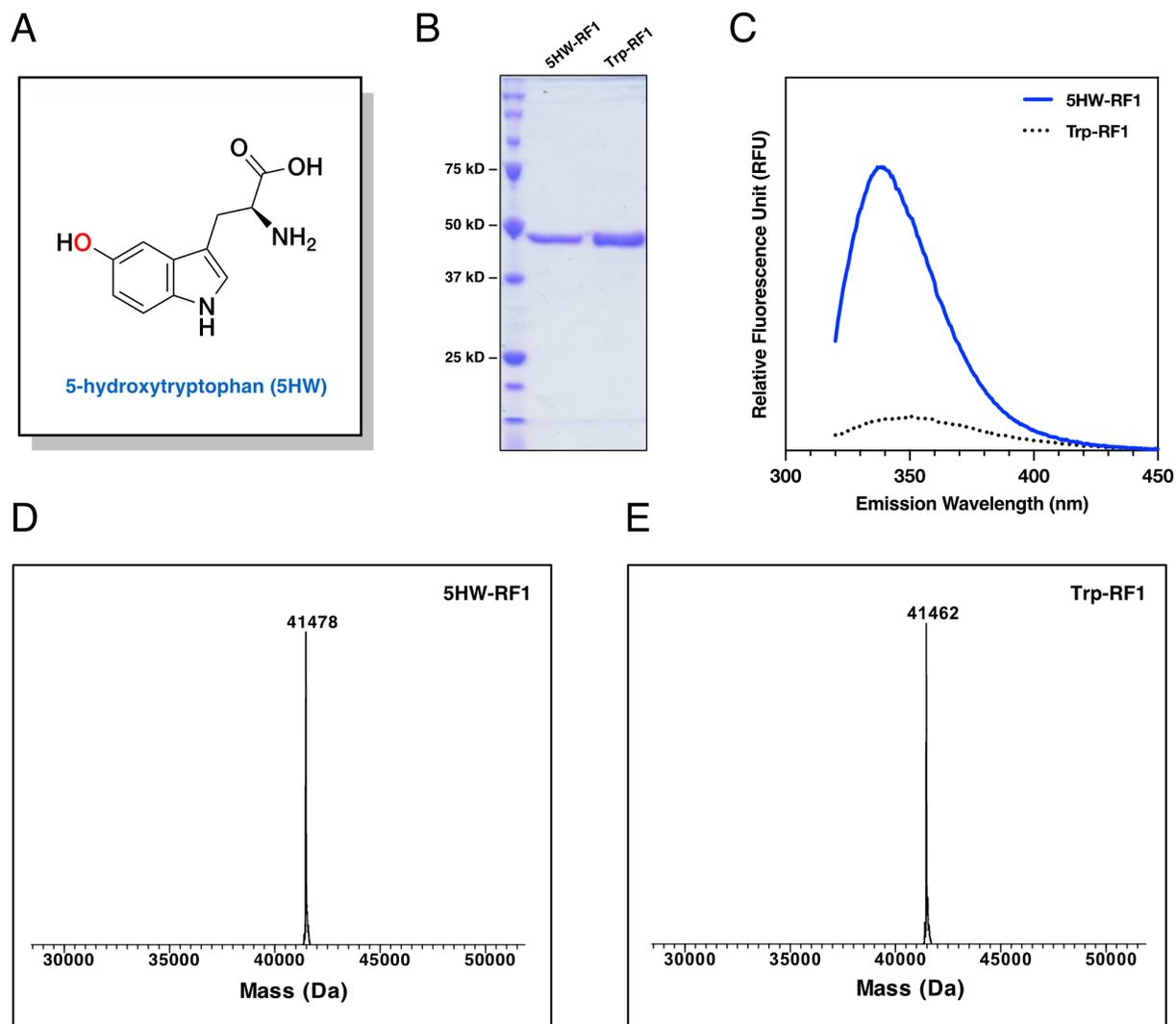
**Figure 4.2. Structural schema of *E. coli* release factor 1 (RF1) in open and closed conformations**

Structures are modified from published models for open (6osk) and close (2b3t) conformation. (A) Alignment of RF1 in open and close conformations. (B)(C) Green spheres indicate the tryptophan residue (W144) and red spheres indicate the cysteine residue (C257) that will be labeled with fluorescent dyes. (D)(E) Residue-to-residue distance map of the open and closed conformations of *E. coli* RF1 models. (D) The color code shows the residue-to-residue difference in separation distance between open/closed states of RF1. (E) The color code shows the residue-to-residue difference in FRET efficiency between open/closed states of RF1.

The optimality of the labeling sites was further examined by using two-dimensional residue-to-residue maps<sup>81,103</sup>. The color-coded distance map shows the distance change of any residue pair in closed and open states, which is calculated based on the relative location of  $\alpha$ -carbons of any two residues in the structural model (**Figure 4.2D**). The dark color region are those residual pairs whose separation distance will not change a lot when the protein changes its conformation, while the light color region shows the residual pairs which have significantly different separation distances in closed and open states. The map clearly shows four folded domains, and the C257/W144 pair lies close to the GGQ/PXT pair yet not on the same vertical and horizontal lines, which means they will not be too close and likely affect the biological functions of GGQ and PXT motifs. The distance map was further transformed into another 2-dimensional map showing the estimated change of FRET efficiency based on fluorescein and tetramethylrhodamine (**Figure 4.2E**). The FRET efficiency map shows that C257/W144 pair may show a significant change in FRET efficiency when labeled with fluorescein and tetramethylrhodamine.

#### **4.3.3. Expression of 5-hydroxytryptophan-incorporated scswRF1**

After trying a few tryptophan analogs, 5-hydroxytryptophan (5HW, **Figure 4.3A**) was chosen as the main focus, because it could be readily incorporated by *E. coli* BL21(DE3)/NK7402 strain with moderately good efficiency<sup>101</sup>, has its own distinctive spectral properties<sup>101,104–106</sup>—it and the proteins containing it absorb between 300 to 320 nm, and would emit at higher wavelength region—and there are bioconjugation methods to target it specifically<sup>89,107</sup>. 5HW is actually a natural occurring amino acid, known as the precursor of the neurotransmitter serotonin<sup>108</sup>. Structurally, it differs from regular tryptophan molecule by one oxygen atom at position 5 on the indole ring (**Figure 4.3A**).



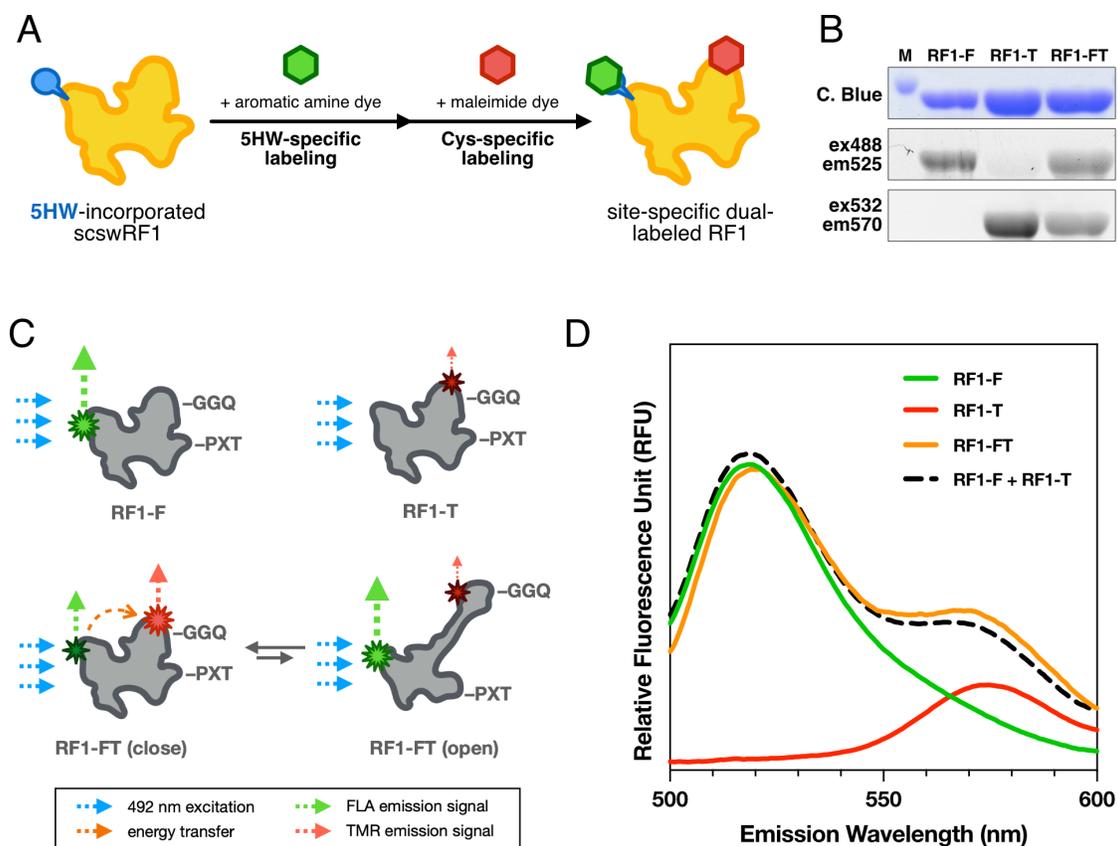
**Figure 4.3 Incorporation of 5-hydroxytryptophan into single-cysteine single-tryptophan RF1**

(A) Chemical structure of 5-hydroxytryptophan. (B) Coomassie-stained SDS-PAGE showing purified 5HW-incorporated scswRF1 (5HW-RF1) and regular single-cysteine single-tryptophan RF1 (Trp-RF1). (C) Fluorescence spectroscopy of 5HW-incorporated RF1 (blue solid) and regular RF1 (black dotted) with excitation wavelength at 310 nm. (D)(E) Deconvoluted LC-ESI-MS spectra of (D) 5HW-RF1 and (E) Trp-RF1, which shows a 16 Da difference caused by regular tryptophan and 5-hydroxytryptophan.

Following the above-mentioned protocol, we could successfully overexpress scswRF1 protein in moderately good yield, a few milligrams per liter liquid culture. The recombinant scswRF1s with 5-hydroxytryptophan (5HW-RF1) or with regular tryptophan (Trp-RF1) were further purified using affinity and ion exchange chromatography (**Figure 4.3B**). With fluorescence spectroscopy 5HW-RF1 shows strong emission at 340 nm when excited at 310 nm while Trp-RF1 does not. LC-ESI-TOF-mass spectrometry shows that the difference in intact protein masses between Trp-RF1 and 5HW-RF1 is exactly 16, the atomic mass of an oxygen atom, indicating the successful incorporation of one and only one 5-hydroxytryptophan into the recombinant scswRF1 protein (**Figure 4.3D-E**).

#### **4.3.4. Dual site-specific labeling of 5-hydroxytryptophan-incorporated scswRF1**

The scswRF1 is designed to be labeled with one fluorophore targeting the cysteine residue and the other targeting the tryptophan analog. For 5-hydroxytryptophan site, 5-aminofluorescein (FLA) could be covalently attached onto the indole ring under a mild oxidative condition<sup>89</sup>, while cysteine was labelled with tetramethylrhodamine (TMR)-maleimide (**Figure 4.4A**). To determine the specificity of the labeling reactions, the scswRF1 was labeled with either FLA (RF1-F) or TMR (RF1-T) and doubly labeled with both FLA and TMR (RF1-FT). The labeled proteins were separated by SDS-polyacrylamide gel electrophoresis (SDS-PAGE) and the gel was scanned for fluorescence signals with a Typhoon scanner (excitation at 488 nm/emission at 525 nm for FLA, excitation at 532 nm/emission at 570 nm for TMR). We observed the FLA-labeled and TMR-labeled proteins only at their corresponding fluorescence channels showing that the proteins were successfully labeled with the individual dyes. The scswRF1 that was reacted with both FLA and TMR dyes was responsive to both fluorescence channels showing that the protein was conjugated to both dyes (**Figure 4.4B**).



**Figure 4.4 Site-specific dual-labeling of 5HW-incorporated RF1 and spectroscopic analysis**

(A) Reaction schema of dual site-specific labeling on 5HW-incorporated RF1 protein. (B) Fluorescence and Coomassie Blue-stained SDS-PAGE gel images of single-labeled (RF1-F with fluorescein, RF1-T with tetramethylrhodamine) and double-labeled (RF1-FT) protein. (C) Fluorescence spectroscopy of regular tryptophan (black dotted) and 5HW-incorporated RF1 (blue solid) with excitation wavelength at 310 nm. (D) Fluorescence spectroscopy of RF1-F (green solid line), RF1-T (red solid line) and RF1-FT (orange solid line) with excitation wavelength at 492 nm. Sum of RF1-F and RF1-T spectra is plotted for comparison (black dashed line). RF1-FT shows slightly lower signal at FLA emission peak (520 nm) and slightly higher signal at TMR emission peak (575 nm) than the combined signal of two single-labeled proteins.

### 4.3.5. Fluorescence spectroscopic analysis of dye-labeled RF1

The dye-labeled scswRF1s also behaved as expected (**Figure 4.4D**). When excited at 492 nm, the RF1 with only FLA conjugated to the 5HW site showed one emission peak at 520 nm (**Figure 4.4D**, green solid line), and the RF1 with only TMR conjugated to the cysteine site showed one emission peak at 575 nm (**Figure 4.4D**, red solid line), while the double-labeled RF1 showed two peaks corresponding to FLA and TMR in the emission spectrum (**Figure 4.4D**, orange solid line), which confirm that the protein is labeled by both dyes.

With both single-labeled and double-labeled proteins in hand, we wished to see if there is FRET in the presumably closed form of RF1 (**Figure 4.4C**). Comparing to the sum of the signals from the two single-labeled proteins (**Figure 4.4D**, black dashed line), the FLA emission peak at 520 nm is slightly lower and TMR emission peak at 575 nm slightly higher in the spectrum of RF1-FT, suggesting there is a small FRET in RF1 protein.

## 4.4. Discussions

Dual or multiple site-specific labeling of protein is useful for various kinds of biochemical and biophysical research, yet it is not a trivial task. To overcome the limitation of direct bioconjugation with amine- or thiol-reactive chemistry, scientists have developed many strategies, such as labeling two fragments of a protein separately and then joining them together into one protein, using technique such as native chemical ligation<sup>109</sup> or intein-mediated ligations<sup>110</sup>. Recent advances in biorthogonal bioconjugation reactions and expanded genetic code enables a more general strategy for site-specific labeling—first incorporate ncAA with a chemical handle on the amino acid side chain, then perform bioconjugation reaction specific to that chemical handle to attach fluorophores or other molecular moieties. Our method is also leveraging the power of ncAA and bioconjugation reactions yet implemented by using auxotrophic strain for ncAA incorporation

for simplicity and efficiency. While not as multi-purpose as expanded genetic code, our method could excel in many scenarios for people who want to install an additional labeling site in their proteins.

Analysis of protein sequences have shown that tryptophan is the rarest amino acid in a protein, on average only one are present in every one hundred amino acids in a protein sequence<sup>111</sup>. This is a key advantage for site-specifically labeling a protein because a single tryptophan at a unique position in a protein can be created with minimal changes to the protein's primary sequence. Additionally, 5-hydroxytryptophan is an economic and commercially available tryptophan analog and can be efficiently incorporated into over-expressed recombinant protein. It has a distinctive fluorescence property compared to regular tryptophan, and can be employed as FRET donor while using AEDANS as acceptor<sup>101</sup>. In the scswRF1 constructed here, it showed a very broad emission range which could even serve as FRET donor to dyes such as fluorescein. Fluorescence dye (e.g., fluorescein amine) or molecular moieties (e.g. 4-carboxydiazonium (4CDZ)-biotin) can be attached onto it under ambient reaction condition, which makes it a useful tool for site-specific labeling when it is incorporated into proteins. In principle, the same strategy can work with other tryptophan analogs, for example 5-azidotryptophan, which can further expand the applicable reactions for site-specific protein labeling.

While we successfully demonstrated that RF1 was site-specifically labeled with two different dyes and have observed some FRET phenomenon, this current construct cannot provide any further insights on the dynamic property of RF1. It may be because the position or the choice of dyes are not optimal. Based on the theoretical calculation according to the residue-to-residue map (**Figure 4.2D and 4.2E**), the separation distance change between open and closed state is 25Å and the expected FRET efficiency change is up to 33%. However, the microenvironments inside

the protein and the equilibrium between open and closed or any other possible intermediate conformations will affect the fluorescence properties of the dye and the observable signals. Further optimization would be required if we want to use it to study the exact conformational state of release factor proteins in solution and further understand the reaction kinetics during translation. For example, screening all the potential labeling sites and their near-by locations, using different fluorophore pairs, and inserting linkers with different lengths could all potentially improve the FRET efficiency.

## **4.5. Materials and Methods**

### **4.5.1. Chemicals, buffers, and bacterial strains**

L-tryptophan, 5-aminofluorescein (FLA), and tetramethylrhodamine-5-maleimide (TMR) were purchased from Sigma-Aldrich. L-5-hydroxytryptophan (Acros Organics) was purchased from Fisher Scientific. M63 minimal media were prepared using premixed M63 Medium Broth powder (VMR). Spectroscopic experiments were carried out in a buffer of 50 mM K-HEPES (pH 7.5) and 300 mM NaCl. The Trp auxotroph *Escherichia coli* BL21 ( $\lambda$ DE3)/NK7402 was a gift by Dr. A. Rod Merrill (University of Guelph, Canada).

### **4.5.2. Mutant RF1 expression and purification**

Mutant RF1 was produced by site-directed mutagenesis (QuikChange, Stratagene). Starting from a cysteine-free RF1 (C51S, C201S, C257S) gene in pPROEx-HTc vector (Invitrogen), we first subcloned the gene into pBAD LIC 8A vector (Addgene #37501), and then introduced single-tryptophan mutation (W55H), followed by the single-cysteine mutation (C257). RF1 mutant proteins were purified by nickel-affinity chromatography and concentrated using a

10-kDa MWCO spin column (Amicon Ultra-15, EMD Millipore). Purified proteins were then quantitated by the Bradford assay, flash-frozen, and stored at -80 °C.

#### **4.5.3. Expression of RF1 protein with 5-hydroxytryptophan (5-HW)**

The 5-hydroxytryptophan-incorporated RF1 protein was expressed and purified using the Trp auxotrophic strain as reported<sup>101</sup> with several adjustments. Electrocompetent *E. coli* BL21 ( $\lambda$ DE3)/NK7402 Trp auxotrophic cells were prepared<sup>112</sup> and stored at -80 °C. The competent cells were transformed with pBAD plasmids containing the desired RF1 mutant gene by electroporation and grown overnight at 37°C on LB/Ampicillin (Amp) plates. Next day, each plate was scraped into 5 mL of Super Optimal Broth (SOB) with ampicillin and 2% glucose (from sterile filtered 20% Glucose solution) and incubated at 37°C for 1 h. The 5 mL culture was then transferred to a 4 L flask containing 1 L M63 minimal medium supplemented with 2.0% glucose, 100  $\mu$ g/mL ampicillin, 0.25M L-Trp, and 0.4% glycerol. This culture was grown to 0.5-0.7 OD600 at 37°C, after which the cells were pelleted by centrifugation. The cell pellet was then washed twice with 500 mL of M63 medium supplemented with 0.2% glycerol to remove all traces of residual L-Trp. The cell pellet was then resuspended into the original volume of M63 media containing 0.6% glycerol and 100  $\mu$ g/mL ampicillin and grown for a further 20 min to deplete any residual tryptophan in the culture. Subsequently, the tryptophan analogues (D, L-forms) (Sigma, St. Louis, MO) were added to the minimal medium at a final concentration of 0.5 mM, and the cells were induced with 1% arabinose (pBAD). The culture was allowed to grow for 3h at 37°C, and the cells were harvested by centrifugation. Proteins were purified by nickel-affinity chromatography and HiTrap Q HP anion exchange chromatography and concentrated using a 10-kDa MWCO spin column (Amicon Ultra-15, EMD Millipore). Purified proteins were then quantitated by the Bradford assay, flash-frozen, and stored at -80 °C.

#### **4.5.4. Labeling of RF1 mutants**

For cysteine labeling, 100  $\mu$ L RF1 mutants (40  $\mu$ M final concentration) in labeling buffer [50 mM K-HEPES (pH 7.5) and 300 mM NaCl] was incubated with 20-fold excess (1 mM final concentration) of tetramethylrhodamine-5-maleimide (TMR) (Invitrogen) at room temperature in the dark for 2–4 h. Bioconjugation of 5-hydroxytryptophan with 5-aminofluorescein (FLA) (Sigma-Aldrich) was carried out as reported<sup>89</sup> with several adjustments. 100  $\mu$ L RF1 mutants (40  $\mu$ M final concentration) in labeling buffer [50 mM K-HEPES (pH 7.5) and 300 mM NaCl] was incubated with 100-fold excess (4 mM final concentration) of FLA and 5 equivalent ferricyanide (0.2 mM final concentration) at room temperature in the dark for 2–4 h.

The excess dye was removed by dialyzing against protein storage buffer [50 mM K-HEPES (pH 7.5) and 100 mM NaCl] in the dark overnight. Proteins were further purified by HiTrap Q HP anion exchange chromatography and concentrated using a 10-kDa MWCO spin column (Amicon Ultra-15, EMD Millipore). Purified proteins were then quantitated by the Bradford assay, flash-frozen, and stored at -80 °C.

#### **4.5.5. Staining, imaging, and fluorescence spectroscopy**

SDS-PAGE were stained with Coomassie Brilliant Blue R-250 following the standard protocol<sup>113</sup>, and stained gels were scanned and digitalized by Epson Perfection 2450 Photo Flatbed Scanner. Gel containing fluorescent protein samples were scanned with Typhoon FLA 9500 imager (GE Healthcare), using 473 nm blue LD laser/LBP (510LP) emission filter for FLA and 532 nm green SHG laser/BPG1 (570DF20) emission filter for TMR. Raw images were analyzed and processed using ImageJ software.

Fluorescence spectroscopy was performed with Jasco FP-8500 Series Fluorometers. 0.1  $\mu\text{M}$  of protein samples were excited at 310 (2.5) nm for 5-hydroxytryptophan, 492 (2.5) nm for FLA, and 544 (2.5) nm for TMR and scanning for a range of emission wavelength at 0.5 nm step. Data were analyzed and plotted using GraphPad Prism software.

#### **4.6. Acknowledgement**

Chapter 4, in full, contains material currently in submission to PLOS ONE as in November 2021. Ti Wu and Simpson Joseph. A Simple Method to Dual Site-Specifically Label a Protein Using Tryptophan Auxotrophic *Escherichia coli*. The dissertation author was the primary author of this paper.

## References

- (1) Hodgman, C. E.; Jewett, M. C. Cell-Free Synthetic Biology: Thinking Outside the Cell. *Metabolic Engineering* **2012**, *14* (3), 261–269. <https://doi.org/10.1016/j.ymben.2011.09.002>.
- (2) Villarreal, F.; Tan, C. Cell-Free Systems in the New Age of Synthetic Biology. *Front. Chem. Sci. Eng.* **2017**, *11* (1), 58–65. <https://doi.org/10.1007/s11705-017-1610-x>.
- (3) Kohler, R. The Background to Eduard Buchner’s Discovery of Cell-Free Fermentation. *Journal of the History of Biology* **1971**, *4* (1), 35–61. <https://doi.org/10.1007/BF00356976>.
- (4) Nirenberg, M. W.; Matthaei, J. H. The Dependence of Cell-Free Protein Synthesis in *E. Coli* upon Naturally Occurring or Synthetic Polyribonucleotides. *Proceedings of the National Academy of Sciences* **1961**, *47* (10), 1588–1602. <https://doi.org/10.1073/pnas.47.10.1588>.
- (5) Matthaei, J. H.; Nirenberg, M. W. Characteristics and Stabilization of DNAase-Sensitive Protein Synthesis in *E. Coli* Extracts. *Proceedings of the National Academy of Sciences* **1961**, *47* (10), 1580–1588. <https://doi.org/10.1073/pnas.47.10.1580>.
- (6) Pardee, K.; Green, A. A.; Ferrante, T.; Cameron, D. E.; DaleyKeyser, A.; Yin, P.; Collins, J. J. Paper-Based Synthetic Gene Networks. *Cell* **2014**, *159* (4), 940–954. <https://doi.org/10.1016/j.cell.2014.10.004>.
- (7) Niederholtmeyer, H.; Chaggan, C.; Devaraj, N. K. Communication and Quorum Sensing in Non-Living Mimics of Eukaryotic Cells. *Nat Commun* **2018**, *9* (1), 5027. <https://doi.org/10.1038/s41467-018-07473-7>.
- (8) Tinagar, A.; Jaenes, K.; Pardee, K. Synthetic Biology Goes Cell-Free. *BMC Biol* **2019**, *17* (1), 64. <https://doi.org/10.1186/s12915-019-0685-x>.
- (9) Bhattacharya, A.; Niederholtmeyer, H.; Podolsky, K. A.; Bhattacharya, R.; Song, J.-J.; Brea, R. J.; Tsai, C.-H.; Sinha, S. K.; Devaraj, N. K. Lipid Sponge Droplets as Programmable Synthetic Organelles. *Proc Natl Acad Sci USA* **2020**, *117* (31), 18206–18215. <https://doi.org/10.1073/pnas.2004408117>.
- (10) Hino, M.; Kataoka, M.; Kajimoto, K.; Yamamoto, T.; Kido, J.-I.; Shinohara, Y.; Baba, Y. Efficiency of Cell-Free Protein Synthesis Based on a Crude Cell Extract from *Escherichia Coli*, Wheat Germ, and Rabbit Reticulocytes. *Journal of Biotechnology* **2008**, *133* (2), 183–189. <https://doi.org/10.1016/j.jbiotec.2007.08.008>.
- (11) Dopp, B. J. L.; Tamiev, D. D.; Reuel, N. F. Cell-Free Supplement Mixtures: Elucidating the History and Biochemical Utility of Additives Used to Support *In Vitro* Protein Synthesis in *E. Coli* Extract. *Biotechnology Advances* **2019**, *37* (1), 246–258. <https://doi.org/10.1016/j.biotechadv.2018.12.006>.
- (12) Wu, C.; Sachs, M. S. Preparation of a *Saccharomyces Cerevisiae* Cell-Free Extract for *In Vitro* Translation. In *Methods in Enzymology*; Elsevier, 2014; Vol. 539, pp 17–28. <https://doi.org/10.1016/B978-0-12-420120-0.00002-5>.

- (13) Sawasaki, T.; Gouda, M. D.; Kawasaki, T.; Tsuboi, T.; Toscana, Y.; Takai, K.; Endo, Y. The Wheat Germ Cell-Free Expression System. In *Chemical Genomics*; Zanders, E. D., Ed.; Walker, J. M., Series Ed.; Methods in Molecular Biology™; Humana Press: Totowa, NJ, 2005; Vol. 310, pp 131–144. [https://doi.org/10.1007/978-1-59259-948-6\\_10](https://doi.org/10.1007/978-1-59259-948-6_10).
- (14) Mikami, S.; Masutani, M.; Sonenberg, N.; Yokoyama, S.; Imataka, H. An Efficient Mammalian Cell-Free Translation System Supplemented with Translation Factors. *Protein Expression and Purification* **2006**, *46* (2), 348–357. <https://doi.org/10.1016/j.pep.2005.09.021>.
- (15) Burgenson, D.; Gurramkonda, C.; Pilli, M.; Ge, X.; Andar, A.; Kostov, Y.; Tolosa, L.; Rao, G. Rapid Recombinant Protein Expression in Cell-Free Extracts from Human Blood. *Sci Rep* **2018**, *8* (1), 9569. <https://doi.org/10.1038/s41598-018-27846-8>.
- (16) Shimizu, Y.; Inoue, A.; Tomari, Y.; Suzuki, T.; Yokogawa, T.; Nishikawa, K.; Ueda, T. Cell-Free Translation Reconstituted with Purified Components. *Nat Biotechnol* **2001**, *19* (8), 751–755. <https://doi.org/10.1038/90802>.
- (17) Shimizu, Y.; Ueda, T. PURE Technology. In *Cell-Free Protein Production*; Endo, Y., Takai, K., Ueda, T., Eds.; Methods in Molecular Biology; Humana Press: Totowa, NJ, 2010; Vol. 607, pp 11–21. [https://doi.org/10.1007/978-1-60327-331-2\\_2](https://doi.org/10.1007/978-1-60327-331-2_2).
- (18) Zamecnik, P. C.; Keller, E. B. Relation between Phosphate Energy Donors and Incorporation of Labeled Amino Acids into Proteins. *Journal of Biological Chemistry* **1954**, *209* (1), 337–354. [https://doi.org/10.1016/S0021-9258\(18\)65561-9](https://doi.org/10.1016/S0021-9258(18)65561-9).
- (19) Zubay, G.; Lederman, M.; DeVries, J. K. DNA-Directed Peptide Synthesis. 3. Repression of Beta-Galactosidase Synthesis and Inhibition of Repressor by Inducer in a Cell-Free System. *Proceedings of the National Academy of Sciences* **1967**, *58* (4), 1669–1675. <https://doi.org/10.1073/pnas.58.4.1669>.
- (20) Spirin, A. S.; Baranov, V. I.; Ryabova, L. A.; Ovodov, S.; Alakhov, Y. B. A Continuous Cell-Free Translation System Capable of Producing Polypeptides in High Yield. *Science* **1988**, *242* (4882), 1162–1164. <https://doi.org/10.1126/science.3055301>.
- (21) Kazuta, Y.; Matsuura, T.; Ichihashi, N.; Yomo, T. Synthesis of Milligram Quantities of Proteins Using a Reconstituted in Vitro Protein Synthesis System. *Journal of Bioscience and Bioengineering* **2014**, *118* (5), 554–557. <https://doi.org/10.1016/j.jbiosc.2014.04.019>.
- (22) Lavickova, B.; Maerkl, S. J. A Simple, Robust, and Low-Cost Method To Produce the PURE Cell-Free System. *ACS Synth. Biol.* **2019**, *8* (2), 455–462. <https://doi.org/10.1021/acssynbio.8b00427>.
- (23) Sun, Z. Z.; Hayes, C. A.; Shin, J.; Caschera, F.; Murray, R. M.; Noireaux, V. Protocols for Implementing an Escherichia Coli Based TX-TL Cell-Free Expression System for Synthetic Biology. *JoVE* **2013**, No. 79, 50762. <https://doi.org/10.3791/50762>.
- (24) Caschera, F.; Noireaux, V. Synthesis of 2.3 Mg/ml of Protein with an All Escherichia Coli Cell-Free Transcription–Translation System. *Biochimie* **2014**, *99*, 162–168. <https://doi.org/10.1016/j.biochi.2013.11.025>.

- (25) Shin, J.; Noireaux, V. An *E. Coli* Cell-Free Expression Toolbox: Application to Synthetic Gene Circuits and Artificial Cells. *ACS Synth. Biol.* **2012**, *1* (1), 29–41. <https://doi.org/10.1021/sb200016s>.
- (26) Sun, Z. Z.; Yeung, E.; Hayes, C. A.; Noireaux, V.; Murray, R. M. Linear DNA for Rapid Prototyping of Synthetic Biological Circuits in an *Escherichia Coli* Based TX-TL Cell-Free System. *ACS Synth. Biol.* **2014**, *3* (6), 387–397. <https://doi.org/10.1021/sb400131a>.
- (27) Niederholtmeyer, H.; Stepanova, V.; Maerkl, S. J. Implementation of Cell-Free Biological Networks at Steady State. *Proceedings of the National Academy of Sciences* **2013**, *110* (40), 15985–15990. <https://doi.org/10.1073/pnas.1311166110>.
- (28) Jiang, L.; Zhao, J.; Lian, J.; Xu, Z. Cell-Free Protein Synthesis Enabled Rapid Prototyping for Metabolic Engineering and Synthetic Biology. *Synthetic and Systems Biotechnology* **2018**, *3* (2), 90–96. <https://doi.org/10.1016/j.synbio.2018.02.003>.
- (29) Bundy, B. C.; Hunt, J. P.; Jewett, M. C.; Swartz, J. R.; Wood, D. W.; Frey, D. D.; Rao, G. Cell-Free Biomanufacturing. *Current Opinion in Chemical Engineering* **2018**, *22*, 177–183. <https://doi.org/10.1016/j.coche.2018.10.003>.
- (30) Rasor, B. J.; Vögeli, B.; Landwehr, G. M.; Bogart, J. W.; Karim, A. S.; Jewett, M. C. Toward Sustainable, Cell-Free Biomanufacturing. *Current Opinion in Biotechnology* **2021**, *69*, 136–144. <https://doi.org/10.1016/j.copbio.2020.12.012>.
- (31) Pardee, K.; Green, A. A.; Takahashi, M. K.; Braff, D.; Lambert, G.; Lee, J. W.; Ferrante, T.; Ma, D.; Donghia, N.; Fan, M.; Daringer, N. M.; Bosch, I.; Dudley, D. M.; O'Connor, D. H.; Gehrke, L.; Collins, J. J. Rapid, Low-Cost Detection of Zika Virus Using Programmable Biomolecular Components. *Cell* **2016**, *165* (5), 1255–1266. <https://doi.org/10.1016/j.cell.2016.04.059>.
- (32) Beabout, K.; Bernhards, C. B.; Thakur, M.; Turner, K. B.; Cole, S. D.; Walper, S. A.; Chávez, J. L.; Lux, M. W. Optimization of Heavy Metal Sensors Based on Transcription Factors and Cell-Free Expression Systems. *ACS Synth. Biol.* **2021**. <https://doi.org/10.1021/acssynbio.1c00331>.
- (33) Noireaux, V.; Bar-Ziv, R.; Godefroy, J.; Salman, H.; Libchaber, A. Toward an Artificial Cell Based on Gene Expression in Vesicles. *Phys. Biol.* **2005**, *2* (3), P1–P8. <https://doi.org/10.1088/1478-3975/2/3/P01>.
- (34) Shi, X.; Wu, T.; M. Cole, C.; K. Devaraj, N.; Joseph, S. Optimization of ClpXP Activity and Protein Synthesis in an *E. Coli* Extract-Based Cell-Free Expression System. *Sci Rep* **2018**, *8* (1), 3488. <https://doi.org/10.1038/s41598-018-21739-6>.
- (35) Whitfield, C. J.; Banks, A. M.; Dura, G.; Love, J.; Fieldsend, J. E.; Goodchild, S. A.; Fulton, D. A.; Howard, T. P. Cell-Free Protein Synthesis in Hydrogel Materials. *Chem. Commun.* **2020**, *56* (52), 7108–7111. <https://doi.org/10.1039/D0CC02582H>.
- (36) Jiao, Y.; Liu, Y.; Luo, D.; Huck, W. T. S.; Yang, D. Microfluidic-Assisted Fabrication of Clay Microgels for Cell-Free Protein Synthesis. *ACS Appl. Mater. Interfaces* **2018**, *10* (35), 29308–29313. <https://doi.org/10.1021/acsami.8b09324>.

- (37) Sauer, R. T.; Baker, T. A. AAA+ Proteases: ATP-Fueled Machines of Protein Destruction. *Annu. Rev. Biochem.* **2011**, *80* (1), 587–612. <https://doi.org/10.1146/annurev-biochem-060408-172623>.
- (38) Baker, T. A.; Sauer, R. T. ClpXP, an ATP-Powered Unfolding and Protein-Degradation Machine. *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research* **2012**, *1823* (1), 15–28. <https://doi.org/10.1016/j.bbamcr.2011.06.007>.
- (39) Shin, J.; Noireaux, V. Study of Messenger RNA Inactivation and Protein Degradation in an Escherichia Coli Cell-Free Expression System. *J Biol Eng* **2010**, *4* (1), 9. <https://doi.org/10.1186/1754-1611-4-9>.
- (40) Garamella, J.; Marshall, R.; Rustad, M.; Noireaux, V. The All E. Coli TX-TL Toolbox 2.0: A Platform for Cell-Free Synthetic Biology. *ACS Synth. Biol.* **2016**, *5* (4), 344–355. <https://doi.org/10.1021/acssynbio.5b00296>.
- (41) Glynn, S. E.; Martin, A.; Nager, A. R.; Baker, T. A.; Sauer, R. T. Structures of Asymmetric ClpX Hexamers Reveal Nucleotide-Dependent Motions in a AAA+ Protein-Unfolding Machine. *Cell* **2009**, *139* (4), 744–756. <https://doi.org/10.1016/j.cell.2009.09.034>.
- (42) Wang, J.; Hartling, J. A.; Flanagan, J. M. The Structure of ClpP at 2.3 Å Resolution Suggests a Model for ATP-Dependent Proteolysis. *Cell* **1997**, *91* (4), 447–456. [https://doi.org/10.1016/S0092-8674\(00\)80431-6](https://doi.org/10.1016/S0092-8674(00)80431-6).
- (43) Flynn, J. M.; Neher, S. B.; Kim, Y.-I.; Sauer, R. T.; Baker, T. A. Proteomic Discovery of Cellular Substrates of the ClpXP Protease Reveals Five Classes of ClpX-Recognition Signals. *Molecular Cell* **2003**, *11* (3), 671–683. [https://doi.org/10.1016/S1097-2765\(03\)00060-1](https://doi.org/10.1016/S1097-2765(03)00060-1).
- (44) Gottesman, S.; Roche, E.; Zhou, Y.; Sauer, R. T. The ClpXP and ClpAP Proteases Degrade Proteins with Carboxy-Terminal Peptide Tails Added by the SsrA-Tagging System. *Genes & Development* **1998**, *12* (9), 1338–1347. <https://doi.org/10.1101/gad.12.9.1338>.
- (45) Levchenko, I.; Seidel, M.; Sauer, R. T.; Baker, T. A. A Specificity-Enhancing Factor for the ClpXP Degradation Machine. *Science* **2000**, *289* (5488), 2354–2356. <https://doi.org/10.1126/science.289.5488.2354>.
- (46) Flynn, J. M. Modulating Substrate Choice: The SspB Adaptor Delivers a Regulator of the Extracytoplasmic-Stress Response to the AAA+ Protease ClpXP for Degradation. *Genes & Development* **2004**, *18* (18), 2292–2301. <https://doi.org/10.1101/gad.1240104>.
- (47) Bolon, D. N.; Grant, R. A.; Baker, T. A.; Sauer, R. T. Nucleotide-Dependent Substrate Handoff from the SspB Adaptor to the AAA+ ClpXP Protease. *Molecular Cell* **2004**, *16* (3), 343–350. <https://doi.org/10.1016/j.molcel.2004.10.001>.
- (48) Sun, Z.; Kim, J.; Singhal, V.; Murray, R. M. Protein Degradation in a TX-TL Cell-Free Expression System Using ClpXP Protease. *BioRxiv* **2015**. <https://doi.org/10.1101/019695>.
- (49) Rivas, G.; Minton, A. P. Macromolecular Crowding In Vitro , In Vivo , and In Between. *Trends in Biochemical Sciences* **2016**, *41* (11), 970–981. <https://doi.org/10.1016/j.tibs.2016.08.013>.
- (50) Sharp, K. A. Unpacking the Origins of In-Cell Crowding. *Proc Natl Acad Sci USA* **2016**, *113* (7), 1684–1685. <https://doi.org/10.1073/pnas.1600098113>.

- (51) Ge, X.; Luo, D.; Xu, J. Cell-Free Protein Expression under Macromolecular Crowding Conditions. *PLoS ONE* **2011**, *6* (12), e28707. <https://doi.org/10.1371/journal.pone.0028707>.
- (52) Jendrasiak, G. L.; Hasty, J. H. The Electrical Conductivity of Hydrated Phospholipids. *Biochimica et Biophysica Acta (BBA) - Lipids and Lipid Metabolism* **1974**, *348* (1), 45–54. [https://doi.org/10.1016/0005-2760\(74\)90091-5](https://doi.org/10.1016/0005-2760(74)90091-5).
- (53) Pautot, S.; Frisken, B. J.; Weitz, D. A. Engineering Asymmetric Vesicles. *Proceedings of the National Academy of Sciences* **2003**, *100* (19), 10718–10721. <https://doi.org/10.1073/pnas.1931005100>.
- (54) Pautot, S.; Frisken, B. J.; Weitz, D. A. Production of Unilamellar Vesicles Using an Inverted Emulsion. *Langmuir* **2003**, *19* (7), 2870–2879. <https://doi.org/10.1021/la026100v>.
- (55) Chowdhuri, S.; Cole, C. M.; Devaraj, N. K. Encapsulation of Living Cells within Giant Phospholipid Liposomes Formed by the Inverse-Emulsion Technique. *ChemBioChem* **2016**, *17* (10), 886–889. <https://doi.org/10.1002/cbic.201500643>.
- (56) Shin, J.; Noireaux, V. Efficient Cell-Free Expression with the Endogenous E. Coli RNA Polymerase and Sigma Factor 70. *J Biol Eng* **2010**, *4* (1), 8. <https://doi.org/10.1186/1754-1611-4-8>.
- (57) Pontes, M. H.; Sevostyanova, A.; Groisman, E. A. When Too Much ATP Is Bad for Protein Synthesis. *Journal of Molecular Biology* **2015**, *427* (16), 2586–2594. <https://doi.org/10.1016/j.jmb.2015.06.021>.
- (58) Elowitz, M. B.; Leibler, S. A Synthetic Oscillatory Network of Transcriptional Regulators. *Nature* **2000**, *403* (6767), 335–338. <https://doi.org/10.1038/35002125>.
- (59) Prindle, A.; Selimkhanov, J.; Li, H.; Razinkov, I.; Tsimring, L. S.; Hasty, J. Rapid and Tunable Post-Translational Coupling of Genetic Circuits. *Nature* **2014**, *508* (7496), 387–391. <https://doi.org/10.1038/nature13238>.
- (60) Kenniston, J. A.; Baker, T. A.; Fernandez, J. M.; Sauer, R. T. Linkage between ATP Consumption and Mechanical Unfolding during the Protein Processing Reactions of an AAA+ Degradation Machine. *Cell* **2003**, *114* (4), 511–520. [https://doi.org/10.1016/S0092-8674\(03\)00612-3](https://doi.org/10.1016/S0092-8674(03)00612-3).
- (61) Itoh, H.; Kawazoe, Y.; Shiba, T. Enhancement of Protein Synthesis by an Inorganic Polyphosphate in an E. Coli Cell-Free System. *Journal of Microbiological Methods* **2006**, *64* (2), 241–249. <https://doi.org/10.1016/j.mimet.2005.05.003>.
- (62) Resnick, S. M.; Zehnder, A. J. B. In Vitro ATP Regeneration from Polyphosphate and AMP by Polyphosphate:AMP Phosphotransferase and Adenylate Kinase from *Acinetobacter Johnsonii* 210A. *Appl Environ Microbiol* **2000**, *66* (5), 2045–2051. <https://doi.org/10.1128/AEM.66.5.2045-2051.2000>.
- (63) Davis, J. H.; Baker, T. A.; Sauer, R. T. Engineering Synthetic Adaptors and Substrates for Controlled ClpXP Degradation. *Journal of Biological Chemistry* **2009**, *284* (33), 21848–21855. <https://doi.org/10.1074/jbc.M109.017624>.

- (64) Kim, Y.-I.; Burton, R. E.; Burton, B. M.; Sauer, R. T.; Baker, T. A. Dynamics of Substrate Denaturation and Translocation by the ClpXP Degradation Machine. *Molecular Cell* **2000**, *5* (4), 639–648. [https://doi.org/10.1016/S1097-2765\(00\)80243-9](https://doi.org/10.1016/S1097-2765(00)80243-9).
- (65) Thompson, J. Testing the Conservation of the Translational Machinery over Evolution in Diverse Environments: Assaying *Thermus Thermophilus* Ribosomes and Initiation Factors in a Coupled Transcription-Translation System from *Escherichia Coli*. *Nucleic Acids Research* **2004**, *32* (19), 5954–5961. <https://doi.org/10.1093/nar/gkh925>.
- (66) Rajagopal, V.; Lorsch, J. R. ATP and GTP Hydrolysis Assays (TLC). In *Methods in Enzymology*; Elsevier, 2013; Vol. 533, pp 325–334. <https://doi.org/10.1016/B978-0-12-420067-8.00025-8>.
- (67) Katzen, F.; Chang, G.; Kudlicki, W. The Past, Present and Future of Cell-Free Protein Synthesis. *Trends in Biotechnology* **2005**, *23* (3), 150–156. <https://doi.org/10.1016/j.tibtech.2005.01.003>.
- (68) Motomura, K.; Hirota, R.; Okada, M.; Ikeda, T.; Ishida, T.; Kuroda, A. A New Subfamily of Polyphosphate Kinase 2 (Class III PPK2) Catalyzes Both Nucleoside Monophosphate Phosphorylation and Nucleoside Diphosphate Phosphorylation. *Appl Environ Microbiol* **2014**, *80* (8), 2602–2608. <https://doi.org/10.1128/AEM.03971-13>.
- (69) Kameda, A.; Shiba, T.; Kawazoe, Y.; Satoh, Y.; Ihara, Y.; Munekata, M.; Ishige, K.; Noguchi, T. A Novel ATP Regeneration System Using Polyphosphate-AMP Phosphotransferase and Polyphosphate Kinase. *Journal of Bioscience and Bioengineering* **2001**, *91* (6), 557–563. [https://doi.org/10.1016/S1389-1723\(01\)80173-0](https://doi.org/10.1016/S1389-1723(01)80173-0).
- (70) Parnell, A. E.; Mordhorst, S.; Kemper, F.; Giurrandino, M.; Prince, J. P.; Schwarzer, N. J.; Hofer, A.; Wohlwend, D.; Jessen, H. J.; Gerhardt, S.; Einsle, O.; Oyston, P. C. F.; Andexer, J. N.; Roach, P. L. Substrate Recognition and Mechanism Revealed by Ligand-Bound Polyphosphate Kinase 2 Structures. *Proc Natl Acad Sci USA* **2018**, *115* (13), 3350–3355. <https://doi.org/10.1073/pnas.1710741115>.
- (71) Nocek, B. P.; Khusnutdinova, A. N.; Ruszkowski, M.; Flick, R.; Burda, M.; Batyrova, K.; Brown, G.; Mucha, A.; Joachimiak, A.; Berlicki, Ł.; Yakunin, A. F. Structural Insights into Substrate Selectivity and Activity of Bacterial Polyphosphate Kinases. *ACS Catal.* **2018**, *8* (11), 10746–10760. <https://doi.org/10.1021/acscatal.8b03151>.
- (72) Strohmeier, G. A.; Eiteljörg, I. C.; Schwarz, A.; Winkler, M. Enzymatic One-Step Reduction of Carboxylates to Aldehydes with Cell-Free Regeneration of ATP and NADPH. *Chem. Eur. J.* **2019**, *25* (24), 6119–6123. <https://doi.org/10.1002/chem.201901147>.
- (73) Cui, C.; Kong, M.; Wang, Y.; Zhou, C.; Ming, H. Characterization of Polyphosphate Kinases for the Synthesis of GSH with ATP Regeneration from AMP. *Enzyme and Microbial Technology* **2021**, *149*, 109853. <https://doi.org/10.1016/j.enzmictec.2021.109853>.
- (74) Tavanti, M.; Hosford, J.; Lloyd, R. C.; Brown, M. J. B. ATP Regeneration by a Single Polyphosphate Kinase Powers Multigram-Scale Aldehyde Synthesis *in Vitro*. *Green Chem.* **2021**, *23* (2), 828–837. <https://doi.org/10.1039/D0GC03830J>.

- (75) Wang, P.-H.; Fujishima, K.; Berhanu, S.; Kuruma, Y.; Jia, T. Z.; Khusnutdinova, A. N.; Yakunin, A. F.; McGlynn, S. E. A Bifunctional Polyphosphate Kinase Driving the Regeneration of Nucleoside Triphosphate and Reconstituted Cell-Free Protein Synthesis. *ACS Synth. Biol.* **2020**, *9* (1), 36–42. <https://doi.org/10.1021/acssynbio.9b00456>.
- (76) Jewett, M. C.; Swartz, J. R. Mimicking the Escherichia Coli Cytoplasmic Environment Activates Long-Lived and Efficient Cell-Free Protein Synthesis. *Biotechnol. Bioeng.* **2004**, *86* (1), 19–26. <https://doi.org/10.1002/bit.20026>.
- (77) Arnold, F. H.; Volkov, A. A. Directed Evolution of Biocatalysts. *Current Opinion in Chemical Biology* **1999**, *3* (1), 54–59. [https://doi.org/10.1016/S1367-5931\(99\)80010-6](https://doi.org/10.1016/S1367-5931(99)80010-6).
- (78) Packer, M. S.; Liu, D. R. Methods for the Directed Evolution of Proteins. *Nat Rev Genet* **2015**, *16* (7), 379–394. <https://doi.org/10.1038/nrg3927>.
- (79) Selvin, P. R. Fluorescence Resonance Energy Transfer. In *Methods in Enzymology*; Elsevier, 1995; Vol. 246, pp 300–334. [https://doi.org/10.1016/0076-6879\(95\)46015-2](https://doi.org/10.1016/0076-6879(95)46015-2).
- (80) Selvin, P. R. The Renaissance of Fluorescence Resonance Energy Transfer. *Nat. Struct Biol.* **2000**, *7* (9), 730–734. <https://doi.org/10.1038/78948>.
- (81) Trappl, K.; Joseph, S. Ribosome Induces a Closed to Open Conformational Change in Release Factor 1. *Journal of Molecular Biology* **2016**, *428* (6), 1333–1344. <https://doi.org/10.1016/j.jmb.2016.01.021>.
- (82) Zhang, G.; Zheng, S.; Liu, H.; Chen, P. R. Illuminating Biological Processes through Site-Specific Protein Labeling. *Chem. Soc. Rev.* **2015**, *44* (11), 3405–3417. <https://doi.org/10.1039/C4CS00393D>.
- (83) Hoesl, M. G.; Budisa, N. Recent Advances in Genetic Code Engineering in Escherichia Coli. *Current Opinion in Biotechnology* **2012**, *23* (5), 751–757. <https://doi.org/10.1016/j.copbio.2011.12.027>.
- (84) Johnson, J. A.; Lu, Y. Y.; Van Deventer, J. A.; Tirrell, D. A. Residue-Specific Incorporation of Non-Canonical Amino Acids into Proteins: Recent Developments and Applications. *Current Opinion in Chemical Biology* **2010**, *14* (6), 774–780. <https://doi.org/10.1016/j.cbpa.2010.09.013>.
- (85) Liu, C. C.; Schultz, P. G. Adding New Chemistries to the Genetic Code. *Annu. Rev. Biochem.* **2010**, *79* (1), 413–444. <https://doi.org/10.1146/annurev.biochem.052308.105824>.
- (86) McKay, C. S.; Finn, M. G. Click Chemistry in Complex Mixtures: Bioorthogonal Bioconjugation. *Chemistry & Biology* **2014**, *21* (9), 1075–1101. <https://doi.org/10.1016/j.chembiol.2014.09.002>.
- (87) Xiao, H.; Schultz, P. G. At the Interface of Chemical and Biological Synthesis: An Expanded Genetic Code. *Cold Spring Harb Perspect Biol* **2016**, *8* (9), a023945. <https://doi.org/10.1101/cshperspect.a023945>.
- (88) de la Torre, D.; Chin, J. W. Reprogramming the Genetic Code. *Nat Rev Genet* **2021**, *22* (3), 169–184. <https://doi.org/10.1038/s41576-020-00307-7>.

- (89) Sarathi Addy, P.; Italia, J. S.; Chatterjee, A. An Oxidative Bioconjugation Strategy Targeted to a Genetically Encoded 5-Hydroxytryptophan. *ChemBioChem* **2018**, *19* (13), 1375–1378. <https://doi.org/10.1002/cbic.201800111>.
- (90) Al-Kazwini, A. T.; O'Neill, P.; Adams, G. E.; Cundall, R. B.; Junino, A.; Maignan, J. Characterisation of the Intermediates Produced upon One-Electron Oxidation of 4-, 5-, 6- and 7-Hydroxyindoles by the Azide Radical. *J. Chem. Soc., Perkin Trans. 2* **1992**, No. 4, 657. <https://doi.org/10.1039/p29920000657>.
- (91) Wrona, M. Z.; Dryhurst, G. Oxidation of Serotonin by Superoxide Radical: Implications to Neurodegenerative Brain Disorders. *Chem. Res. Toxicol.* **1998**, *11* (6), 639–650. <https://doi.org/10.1021/tx970185w>.
- (92) Kodadek, T.; Durouxrichard, I.; Bonnafous, J. Techniques: Oxidative Cross-Linking as an Emergent Tool for the Analysis of Receptor-Mediated Signalling Events. *Trends in Pharmacological Sciences* **2005**, *26* (4), 210–217. <https://doi.org/10.1016/j.tips.2005.02.010>.
- (93) Capecchi, M. R. Polypeptide Chain Termination in Vitro: Isolation of a Release Factor. *Proceedings of the National Academy of Sciences* **1967**, *58* (3), 1144–1151. <https://doi.org/10.1073/pnas.58.3.1144>.
- (94) Rawat, U. B. S.; Zavialov, A. V.; Sengupta, J.; Valle, M.; Grassucci, R. A.; Linde, J.; Vestergaard, B.; Ehrenberg, M.; Frank, J. A Cryo-Electron Microscopic Study of Ribosome-Bound Termination Factor RF2. *Nature* **2003**, *421* (6918), 87–90. <https://doi.org/10.1038/nature01224>.
- (95) Laurberg, M.; Asahara, H.; Korostelev, A.; Zhu, J.; Trakhanov, S.; Noller, H. F. Structural Basis for Translation Termination on the 70S Ribosome. *Nature* **2008**, *454* (7206), 852–857. <https://doi.org/10.1038/nature07115>.
- (96) Weixlbaumer, A.; Jin, H.; Neubauer, C.; Voorhees, R. M.; Petry, S.; Kelley, A. C.; Ramakrishnan, V. Insights into Translational Termination from the Structure of RF2 Bound to the Ribosome. *Science* **2008**, *322* (5903), 953–956. <https://doi.org/10.1126/science.1164840>.
- (97) Korostelev, A.; Asahara, H.; Lancaster, L.; Laurberg, M.; Hirschi, A.; Zhu, J.; Trakhanov, S.; Scott, W. G.; Noller, H. F. Crystal Structure of a Translation Termination Complex Formed with Release Factor RF2. *Proceedings of the National Academy of Sciences* **2008**, *105* (50), 19684–19689. <https://doi.org/10.1073/pnas.0810953105>.
- (98) Brown, A.; Shao, S.; Murray, J.; Hegde, R. S.; Ramakrishnan, V. Structural Basis for Stop Codon Recognition in Eukaryotes. *Nature* **2015**, *524* (7566), 493–496. <https://doi.org/10.1038/nature14896>.
- (99) Fu, Z.; Indrisiunaite, G.; Kaledhonkar, S.; Shah, B.; Sun, M.; Chen, B.; Grassucci, R. A.; Ehrenberg, M.; Frank, J. The Structural Basis for Release-Factor Activation during Translation Termination Revealed by Time-Resolved Cryogenic Electron Microscopy. *Nat Commun* **2019**, *10* (1), 2579. <https://doi.org/10.1038/s41467-019-10608-z>.
- (100) Svidritskiy, E.; Demo, G.; Loveland, A. B.; Xu, C.; Korostelev, A. A. Extensive Ribosome and RF2 Rearrangements during Translation Termination. *eLife* **2019**, *8*, e46850. <https://doi.org/10.7554/eLife.46850>.

- (101) Mohammadi, F.; Prentice, G. A.; Merrill, A. R. Protein–Protein Interaction Using Tryptophan Analogues: Novel Spectroscopic Probes for Toxin–Elongation Factor-2 Interactions. *Biochemistry* **2001**, *40* (34), 10273–10283. <https://doi.org/10.1021/bi011035u>.
- (102) Graille, M.; Heurgué-Hamard, V.; Champ, S.; Mora, L.; Scrima, N.; Ulryck, N.; van Tilbeurgh, H.; Buckingham, R. H. Molecular Basis for Bacterial Class I Release Factor Methylation by PrmC. *Molecular Cell* **2005**, *20* (6), 917–927. <https://doi.org/10.1016/j.molcel.2005.10.025>.
- (103) Yu, X.; Wu, X.; Bermejo, G. A.; Brooks, B. R.; Taraska, J. W. Accurate High-Throughput Structure Mapping and Prediction with Transition Metal Ion FRET. *Structure* **2013**, *21* (1), 9–19. <https://doi.org/10.1016/j.str.2012.11.013>.
- (104) Hogue, C. W. V.; Rasquinha, I.; Szabo, A. G.; MacManus, J. P. A New Intrinsic Fluorescent Probe for Proteins Biosynthetic Incorporation of 5-Hydroxytryptophan into Oncomodulin. *FEBS Letters* **1992**, *310* (3), 269–272. [https://doi.org/10.1016/0014-5793\(92\)81346-N](https://doi.org/10.1016/0014-5793(92)81346-N).
- (105) Ross, J. B. A.; Senear, D. F.; Waxman, E.; Kombo, B. B.; Rusinova, E.; Huang, Y. T.; Laws, W. R.; Hasselbacher, C. A. Spectral Enhancement of Proteins: Biological Incorporation and Fluorescence Characterization of 5-Hydroxytryptophan in Bacteriophage A Ci Repressor. *Proc. Natl. Acad. Sci. USA* **1992**, *5*. <https://doi.org/10.1073/pnas.89.24.12023>.
- (106) Alexander Ross, J. B.; Szabo, A. G.; Hogue, C. W. V. Enhancement of Protein Spectra with Tryptophan Analogs: Fluorescence Spectroscopy of Protein-Protein and Protein-Nucleic Acid Interactions. In *Methods in Enzymology*; Elsevier, 1997; Vol. 278, pp 151–190. [https://doi.org/10.1016/S0076-6879\(97\)78010-8](https://doi.org/10.1016/S0076-6879(97)78010-8).
- (107) Addy, P. S.; Erickson, S. B.; Italia, J. S.; Chatterjee, A. A Chemoselective Rapid Azo-Coupling Reaction (CRACR) for Unclickable Bioconjugation. *J. Am. Chem. Soc.* **2017**, *139* (34), 11670–11673. <https://doi.org/10.1021/jacs.7b05125>.
- (108) Birdsall, T. C. 5-Hydroxytryptophan: A Clinically-Effective Serotonin Precursor. *Altern Med Rev* **1998**, *3* (4), 271–280.
- (109) Lee, T. C.; Moran, C. R.; Cistrone, P. A.; Dawson, P. E.; Deniz, A. A. Site-Specific Three-Color Labeling of  $\alpha$ -Synuclein via Conjugation to Uniquely Reactive Cysteines during Assembly by Native Chemical Ligation. *Cell Chemical Biology* **2018**, *25* (6), 797-801.e4. <https://doi.org/10.1016/j.chembiol.2018.03.009>.
- (110) Yang, J.-Y.; Yang, W. Y. Site-Specific Two-Color Protein Labeling for FRET Studies Using Split Inteins. *J. Am. Chem. Soc.* **2009**, *131* (33), 11644–11645. <https://doi.org/10.1021/ja9030215>.
- (111) Kozlowski, L. P. Proteome-PI: Proteome Isoelectric Point Database. *Nucleic Acids Research* **2017**, *45* (D1), D1112–D1116. <https://doi.org/10.1093/nar/gkw978>.
- (112) Green, M. R.; Sambrook, J. Transformation of Escherichia Coli by Electroporation. *Cold Spring Harb Protoc* **2020**, *2020* (6). <https://doi.org/10.1101/pdb.prot101220>.

(113) Simpson, R. J. Staining Proteins in Gels with Coomassie Blue. *Cold Spring Harbor Protocols* **2007**, 2007 (8). <https://doi.org/10.1101/pdb.prot4719>.