INVESTIGATING PROKARYOTIC TRANSLATION ELONGATION AND INITIATION

LUC ARDEN HAMILTON ROBERTS

Bachelor of Science, University of Lethbridge, 2012

A thesis submitted in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

BIOMOLECULAR SCIENCE

Alberta RNA Research and Training Institute

Department of Chemistry and Biochemistry

University of Lethbridge

LETHBRIDGE, ALBERTA, CANADA

INVESTIGATING PROKARYOTIC TRANSLATION ELONGATION AND INITIATION

LUC ARDEN HAMILTON ROBERTS

Date of Defence: April 19th, 2023

Dr. Hans-Joachim Wieden	Professor	Ph.D.
Dr. Trushar Patel Thesis Co-Supervisors	Associate Professor	Ph.D.
Dr. Ute Kothe Thesis Examination Committee Member University of Manitoba	Professor	Ph.D.
Dr. Steve Wiseman Thesis Examination Committee Member	Associate Professor	Ph.D.
Dr. Jean-Denys Hamel Internal External Examiner Department of Chemistry and Biochemistry Faculty of Arts and Science	Assistant Professor	Ph.D.
Dr. Eric Jan External Examiner University of British Columbia	Professor	Ph.D.
Dr. Michael Gerken Chair, Thesis Examination Committee	Professor	Ph.D.

This dissertation is dedicated to my family.

No, not that kind of doctor.

Abstract

Protein synthesis or translation is a process critical to life performed by all living organisms. The rapid and accurate production of proteins allows organisms to perform cellular functions, reproduce, and react to environmental stimuli. The core theme of this thesis is prokaryotic translation, beginning with building tools to study the structural dynamics of elongation factor thermo unstable (EF-Tu) and using them to determine the order of events during its functional cycle. Additionally, the phenomenon of prokaryotic structure-based translation initiation by the intergenic region internal ribosome entry site was investigated and its mode of action determined to be different from the previously proposed model. Finally, the preparation and biophysical characterization of the translation initiation region of the *rpsA* gene from *Escherichia coli* was performed to begin the work towards an atomic level understanding of this mRNA molecule bound to the prokaryotic ribosome. Together, this thesis demonstrates the robustness of the prokaryotic translation machinery and highlights that, although significant insight has been gained into how these molecular machines operate, there is still additional information needed to fully understand this fundamental process.

Contribution of Authors

In Chapter 2 I developed the scientific concepts, wrote, and revised the chapter, and performed all wet lab experiments. Davinder Kaur Dhalla aided in optimizing the stopped-flow experimental conditions (not presented in this Thesis). Dylan Girodat performed the Gaussian fitting of the EF-Tu populations, MD simulations of EF-Tu (not presented in this thesis), and the subsequent analysis using the CINC pipeline (not presented in this thesis). Hans-Joachim Wieden and Dylan Girodat aided in the structuring of the ideas and edited the manuscript.

In Chapter 5 I developed the scientific concepts, wrote, and revised the chapter, and performed all presented wet lab experiments. When additional work by Justin Vigar, Jalyce Heller, and Dylan Girodat are discussed they are appropriately attributed in the text.

Acknowledgments

Thank you to those who assisted me in collecting the data presented in this thesis, your names are attributed in the corresponding sections. I would also like to thank those members of ARRTI (the Alberta RNA Research and Training Institute), the Department of Chemistry and Biochemistry, and the Wieden Lab who provided me with scientific support and feedback. Specific thanks are owed to Fan Mo who ensured the lights were on, the reagents stocked, and the lab culture was friendly.

Thank you to those on my supervisory committee who stayed with me to the end and to those who joined at the 1-yard line for the final push. Thank you to Dr. Trushar Patel for your easy-going nature, your support, and your willingness to always lend a hand. Thank you to Dr. Ute Kothe for your kindness, your mentorship, and your passion for teaching. Your biochemistry 2000 class led me to pursue an independent study and ultimately research; if I had one regret from my time here, it was that we never got to work together on a research project. Finally, thank you to Dr. Hans-Joachim Wieden for always having your students' best interest at heart, for not backing down from a fight, and for working your hardest to ensure the trainees in your lab got the best scientific training possible. Thank you for believing in me, for knowing when to be kind, and for knowing when not to sugar coat something I needed to hear.

I would like to specifically thank Dr. Dylan Girodat, Justin Vigar (cofounder of I^3), Rhys Hakstol, Dr. Taylor Sheahan, Dr. David McWatters, Jackson Knott, and Dr. Connor MacNeil for being an excellent support system and lifelong friends. Also, thank you to Richard LaFleur, Julian, and Bubbles for the laughs, it feels great to finally get my grade 10.

Thanks to my family for your unending love and support, I wouldn't be here without your help. Special thanks are owed to my wife Jennifer for being the rock on which my life is built, my best friend, and for being the primary bread winner. This could not have happened without you, I love you. Thank you to my dog Blue (Booper) as you always said "Bark, Bark, Woof, Grrrrr". Finally, thank you to my son Hudson, you're only four but you've taught me so much already and I can't wait to learn what you'll teach me next.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

aa-tRNA ..... Aminoacyl-Transfer RNA

aSD ..... anti-Shine-Dalgarno

ATP ..... Adenosine Triphosphate

cDNA ..... Complementary DNA

CDS ..... Coding Sequence

CHARMM ..... Chemistry at Harvard Molecular Mechanics

CINC ..... Computational Identification of Non-disruptive Conjugation sites

cIRES ..... Cellular IRES

CPM ..... Counts per Minute

CrPV ..... Cricket Paralysis Virus

DDPM ..... *N*-(4-dimethylamino-3,5-dinitrophenyl)maleimide

DNA ..... Deoxyribonucleic Acid

Da ..... Dalton

eCFP ..... Enhanced Cyan Fluorescent Protein

EDTA ..... Ethylenediaminetetraacetic Acid

EF .....Elongation Factor

EF-Tu .....Elongation Factor Thermo Unstable

EF-Ts..... Elongation Factor Thermo Stable

eGFP .....Enhanced Green Fluorescent Protein

EMCV .....Encephalomyocarditis Virus

F5M..... Fluorescein-5-maleimide

FMDV ..... Foot-and-mouth Disease Virus

fMET .....*N*-Formylmethionine

FRET .....Förster Resonance Energy Transfer

GAF .....GTPase Activating Factor

GDI .....Guanosine Nucleotide Dissociation Inhibitor

GDP ..... Guanosine Diphosphate

GDPCP..... 5'-Guanosyl-methylene-triphosphate

GDPNP ..... Guanosine 5′-[β,γ-imido]triphosphate

GEF .....GTPase Exchange Factor

GFP ..... Green Fluorescent Protein

gTIR ..... Genomic TIR

GTP ..... Guanosine Triphosphate

GTPase ..... Guanosine 5'-triphosphatases

HCV ..... Hepatitis C Virus

HIV ..... Human Immunodeficiency Virus

HeLa ..... Henrietta Lacks

IAEDANS ..... 5-(2-Iodoacetylaminoethyl)aminonaphthalene-1-Sulfonic Acid

IAPV ..... Israeli Acute Paralysis Virus

IC ..... Initiation Complex

IF ..... Initiation Factor

IGR ..... Intergenic Region

IPTG ..... Isopropyl β-D-1-thiogalactopyranoside

IRES ..... Internal Ribosome Entry Site

ITAFs ..... IRES Trans-acting Factor

LSU ..... Large ribosomal Subunit

m7G 7-Methylguanylate

MANT ..... *N*-Methylanthraniloyl

MD ..... Molecular Dynamics

mRFP ..... Monomeric Red Fluorescent Protein

mRNA ..... Messenger RNA

NAMD ..... Nanoscale Molecular Dynamics

nt ..... Nucleotides

PAGE ..... Polyacrylamide Gel Electrophoresis

PCR ..... Polymerase Chain Reaction

PEP ..... Phosphoenol Pyruvate

P_i ..... Inorganic Phosphate

PK ..... Pseudoknot

PSIV ..... *Plautia stali* Intestine Virus

PIC ..... Pre-initiation Complex

PTC ..... Peptidyl Transfer Center

RBS ..... Ribosome Binding Site

RF ..... Release Factor

RNA ..... Ribonucleic Acid

RNP ..... RNA and Protein

RRF ..... Ribosome Release Factor

RRM ..... Rhodamine-red-maleimide

rRNA ..... Ribosomal RNA

RT-qPCR ..... Reverse Transcription – Quantitative PCR

SD ..... Shine-Dalgarno

SDS ..... Shine-Dalgarno Sequence

SDS-PAGE ..... Sodium Dodecyl Sulfate–polyacrylamide Gel Electrophoresis

sfGFP ..... Superfolder GFP

SSU ..... Small Ribosomal Subunit

TE ..... Translation Efficiency

TIR ..... Translation Initiation Region

tRNA ..... Transfer RNA

tmRNA ..... Transfer-messenger RNA

trGTPase ..... Translational GTPase

UTR ..... Untranslated Region

vIRES ..... Viral IRES

VMD ..... Visual Molecular Dynamics

WT ..... Wild Type

Chapter 1.

Introduction

1.1 Prokaryotic Translation

Proteins are ubiquitous molecular nano-machines present in all domains of life and are responsible for a plethora of essential cellular processes and functions. As proteins are essential to all living organisms, it is not surprising that the cellular machinery, as well as the fundamental processes by which these nanomachines operate, is highly conserved between different domains of life. Protein synthesis requires the genetic information containing the instructions for the correct synthesis of a particular protein stored in deoxyribonucleic acid (DNA), where a series of enzymes and cellular components transcribe the stored genetic information into messenger ribonucleic acid (mRNA) by a process known as transcription [1, 2] (Figure 1.1). Next, that information must be translated into the corresponding amino acid sequences, a process completed by a large megadalton ribonucleoprotein (RNP) complex called the ribosome. The ribosome itself consists of two subunits, a small ribosomal subunit (SSU) and a large ribosomal subunit (LSU), with distinct functional roles [3-6] (Figure 1.2). Broadly, the SSU is responsible for accurately decoding the mRNA by ensuring proper base-paring is occurring between the mRNA and the aminoacyl-transfer RNA (aa-tRNA), while the LSU catalyzes peptide bond formation between amino acids, resulting in the corresponding polypeptide [7]. This massive complex of RNAs and proteins contains three tRNA binding sites termed the acceptor (A site), the peptidyl (P site), and the exit (E site) sites [8]. Translation is a cyclical (or iterative) process and is canonically divided into four distinct phases: initiation, elongation, termination, and recycling, where the resulting ribosomal subunits can subsequently reenter the translation process through the initiation step (Figure 1.2 and 1.3) [7]. Although the process is conserved among different species, the molecular details and involved macromolecules differ, with the bacterial system representing a minimal canonical set of

molecules and processes. To this end, I will primarily focus on the translational machinery from *Escherichia coli* and generalize where possible or demonstrate where it is not.

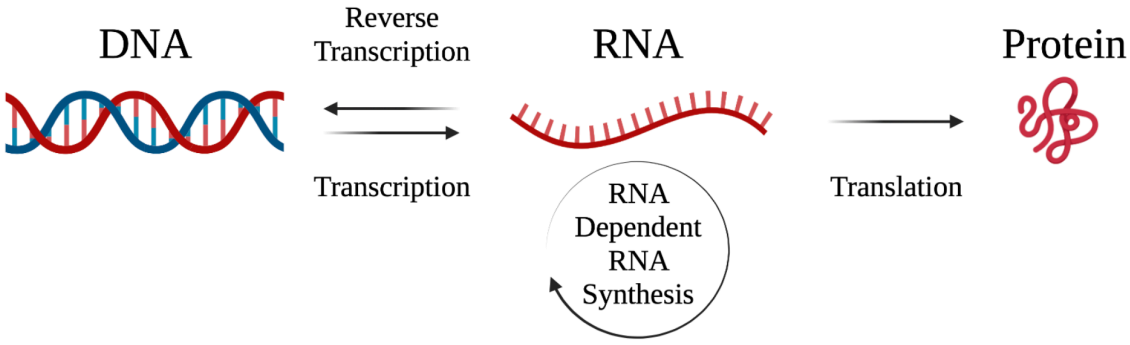

Figure 1.1. The flow of genetic information. In general, the genetic information stored as DNA is transcribed into RNA which is translated into a protein.

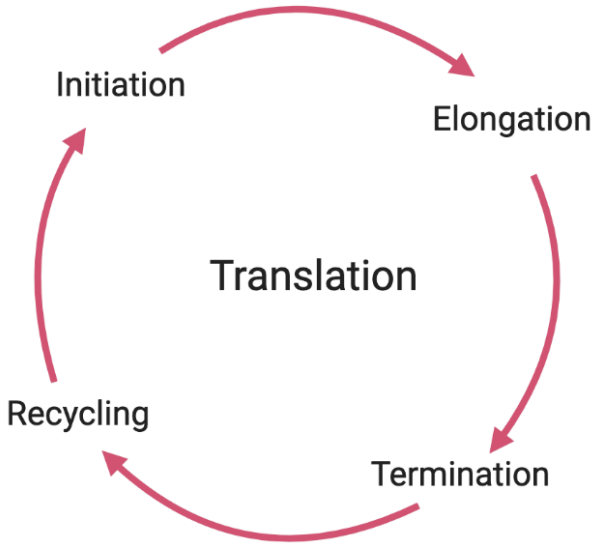

Figure 1.2. The steps of ribosome-dependent protein synthesis. The fundamental steps of the translation cycle are conserved among all domains of life.

1.1.1 Initiation

In prokaryotes, translation initiation is facilitated by three initiation factors (IFs): IF1, IF2, and IF3. IF3 binds to the 30S (small) subunit, commonly right after exiting the ribosome recycling step (*vide infra*) to promote 70S (ribosome) splitting and prevent premature rejoining with the 50S (large) ribosomal subunit [9]. Playing a similar role, IF1 binds to this complex to prevent aa-tRNA biding to the A site of the 30S subunit, while IF2 (complexed with guanosine triphosphate (GTP)) delivers the initiator tRNA (fMet-$tRNA_f^{Met}$) to the P site of the 30S [10-13]. Finally, the translation initiation region (TIR) of the mRNA binds, completing the 30S preinitiation complex [14, 15]. Following canonical codon-anticodon interactions between the start codon (AUG) of the mRNA and the anti-codon of fMet-$tRNA_f^{Met}$, the IFs are released and a 30S initiation complex can recruit the 50S ribosomal subunit, forming the 70S initiation complex [15-18].

1.1.2 Elongation

Translation elongation is also mediated by a series of protein factors appropriately termed elongation factors (EFs). Following 70S initiation complex formation, a ternary complex composed of EF-Tu, GTP, and aa-tRNA transiently binds to the A site of the ribosome in a codon-independent manner [19, 20]. The anticodon of the aa-tRNA samples the mRNA codon present in the 30S A site. If cognate codon-anticodon base pairs are formed (i.e., the correct amino acid), several RNA bases from the 16S rRNA (a component of the 30S) stabilize the newly formed mini-helix. The inability to stabilize this interaction (i.e. the incorrect amino acid) results in the dissociation of the EF-Tu•GTP•aa-tRNA ternary complex, constituting the first step of the aa-tRNA selection process at the heart of the molecular decoding process which ensures translational fidelity [20]. However, upon formation of a cognate codon-anticodon interaction, binding of the ternary complex is greatly stabilized [19, 21] and the GTP hydrolysis activity of EF-Tu is stimulated 10^5-fold [19, 22]. Following GTP hydrolysis, and inorganic phosphate (P_i) release, EF-

Tu undergoes a conformational change resulting in a reduced affinity for the bound aa-tRNA and the ribosome, subsequently dissociating [19, 23]. The CCA-end (carrying the amino acid) of the aa-tRNA is now free to accommodate into the peptidyl transfer center (PTC), where rapid peptide bond (i.e., amide bond) formation occurs, transferring the nascent polypeptide chain from the P site tRNA to the A site aa-tRNA.

GTPase activation, the conformational change of EF-Tu, and the accommodation of the aa-tRNA provide additional filters to prevent the misincorporation of non- or near-cognate amino acids able to pass the first selection step [24-26]. Following this. the translocase EF-G•GTP binds and translocates the aa-tRNAs (along with the mRNA) from the A and P sites to the P and E sites, respectively [27]. This entire process repeats as a new EF-Tu ternary complex can sample the now vacant A site to begin decoding the next codon, while deacylated tRNAs leave via the E site. Successive rounds of aa-tRNA delivery, peptide bond formation, and translocation continue until the ribosome encounters one of three stop-codons (UAA, UAG, or UGA) at the end of the mRNA coding sequence [28-31].

1.1.3 Termination

Translation termination is the final step in peptide synthesis, resulting in the release of the polypeptide from the P site tRNA and the ribosome with the aid of release factors (RFs). Depending on the stop codon presented, either RF1 (UAA and UGA) or RF2 (UAA and UGA) bind to the ribosome, the ester bond between the now complete polypeptide and the last tRNA is cleaved, and the newly synthesized polypeptide is released [32, 33]. Subsequently, RF3•GTP binds to the 70S•RF1/2•tRNA•mRNA complex and promotes the dissociation of the other bound RF (either 1 or 2) before it also dissociates [34].

1.1.4 Recycling

Following termination, the 70S ribosome is left bound to a deacylated tRNA and the previously translated mRNA. For the 70S ribosome to begin the cycle again, these factors are separated, allowing them to be recycled back into the pool of available translational components. This is facilitated by Ribosomal Release/Recycling Factor (RRF) and EF-G•GTP, which bind to the 70S•tRNA•mRNA complex and catalyze subunit splitting as well as tRNA and mRNA dissociation [35-38].

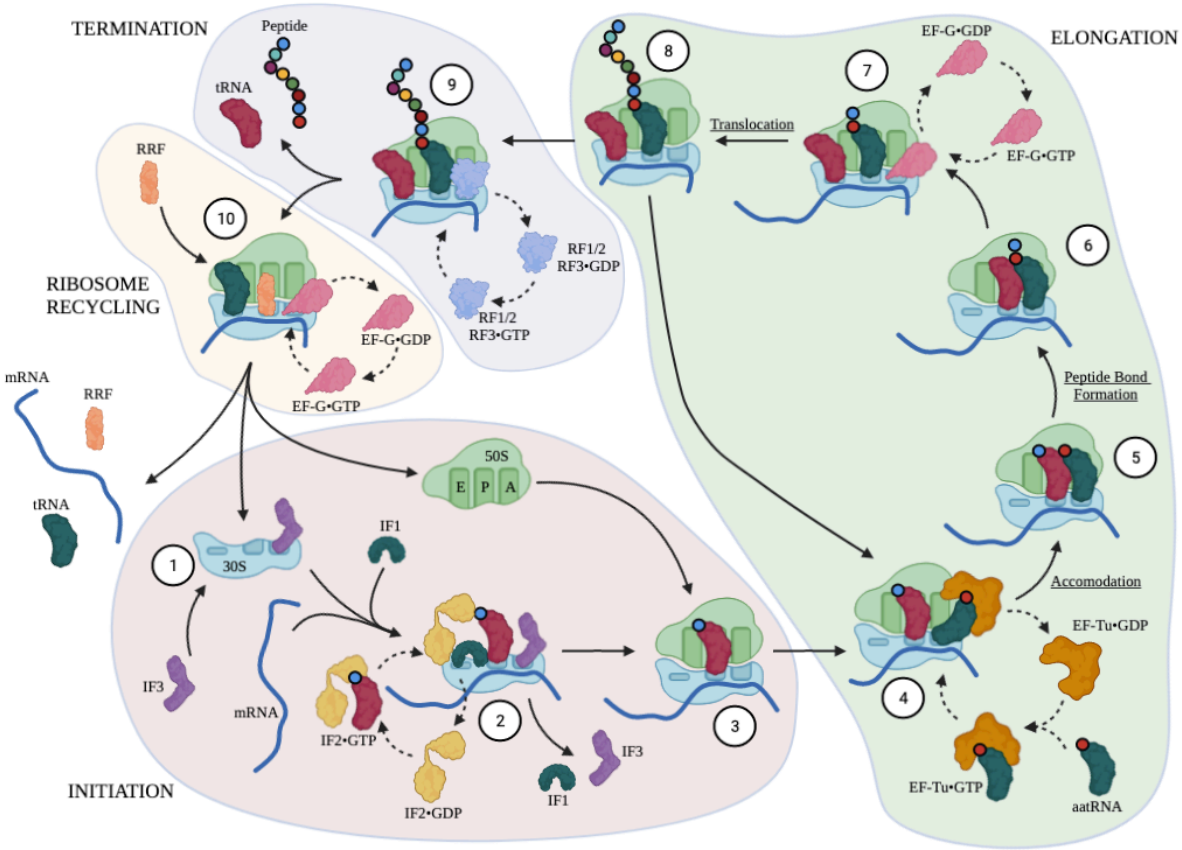

Figure 1.3. Overview of the prokaryotic translation cycle. Initiation (Steps 1-3), elongation (steps 4-8), termination (Step 9) and ribosome recycling (Step 10) are depicted. The major sub steps of elongation are designated in black underlined text.

1.2 Elongation Factor-Tu (EF-Tu)

Prokaryotic translation is a complex process involving a multitude of factors that ensure each step is performed correctly in order to generate the required protein product. The following section will focus on elongation, specifically the guanosine 5'-triphosphatase elongation factor Tu (EF-Tu) and its role in translation.

1.2.1 Guanosine 5'-triphosphatases (GTPases)

Guanosine 5'-triphosphatases (GTPases) are a class of proteins responsible for a wide variety of cellular functions. In the cell, GTPases function as molecular switches that regulate key biological processes (i.e., cell cycle, translation, etc.) [39]. A GTPase's functional state is determined by the associated nucleotide, for example, EF-Tu is in the "on" or active state when bound to GTP, whereas when bound to guanosine diphosphate (GDP), EF-Tu is in the "off" or inactive state [40-42]. The switch from on to off (and vice versa) is often tied to a conformational change relaying the state change to a downstream interaction partner [41, 43]. The ability to rapidly switch between on/off in response to stimuli is critical for the numerous cellular processes in which GTPases play a role (e.g., gene expression). In cases where the rate of GTP hydrolysis is rate limiting to the cellular process, the activation of GTPases can be facilitated by GTPase-activating factors (GAFs) increasing the rate of hydrolysis [20, 44, 45]. Equally critical to the activation of GTPases is the ability to recycle them by exchanging bound GDP for GTP. If the affinity for GDP is higher than that for GTP (e.g., for EF-Tu), or the rate of dissociation is prohibitively slow, a guanosine nucleotide exchange factor (GEF) can facilitate this exchange [46]. Additionally, nucleotide exchange can be regulated by guanine nucleotide dissociation inhibitors (GDI) which serve to inhibit the intrinsic and GEF-stimulated exchange of nucleotide [47].

1.2.2 EF-Tu is a translational GTPase (trGTPase)

EF-Tu is an essential, 43 kDa, three-domain protein and is one of the most abundant proteins found in prokaryotes (approximately 5% of total cellular protein in *Escherichia coli*) [48-50]. Domain 1 is a Rossman fold (i.e., alternating beta strand and alpha helical segments) containing a guanosine binding domain (G-domain) and is responsible for nucleotide binding and hydrolysis [41, 42, 45, 51, 52]. Domains 2 and 3 of EF-Tu are β-barrels which facilitate contacts with EF-Tu's many cellular partners and help define its nucleotide specificity [41-43, 53]. As a GTPase, EF-Tu hydrolyzes GTP to switch between the "on" and "off" state during its functional cycle, undergoing a large conformational change in the process (Figure 1.4, and described below) [41, 43]. The GTPase activity of EF-Tu is relatively slow (~10^{-4} s^{-1}) compared to the speed of translation (~15 amino acids s^{-1}), therefore this activity is enhanced by its associated GAF, the ribosome [22, 45, 54, 55]. Interestingly, the exact mechanism by which EF-Tu's GTPase activity is stimulated is still not clear [20, 56-58]. Finally, after EF-Tu has delivered its payload (i.e., aa-tRNA), it requires its GEF (EF-Ts) to exchange the bound GDP as EF-Tu has a 60-fold higher affinity for GDP than for GTP, a feature unique among trGTPases [45, 59-62]. Once free of GDP, the roughly 10-fold higher cellular concentration of GTP relative to GDP favors GTP binding to the EF-Tu•EF-Ts complex, ensuring active complex formation is favored over GDP rebinding [63, 64].

1.2.3 EF-Tu's Conformational Dynamics

Structural studies of EF-Tu have greatly improved our understanding of EF-Tu's structural changes during its functional cycle. Initial crystal structures of EF-Tu demonstrated that it adopts a "closed" conformation when bound to GTP and in the EF-Tu•GTP•aa-tRNA complex where the

three domains are in close proximity [43]. Conversely, EF-Tu displays a more "open" conformation when bound to GDP with domains 2 and 3 rotating away from the G domain (Figure 1.3) [41, 42]. More recent work has elucidated initial structural contacts and intermediates of EF-Tu•GTP•aa-tRNA as it binds and then subsequently delivers aa-tRNA to the ribosome [65, 66]. It is tempting to think of EF-Tu as existing in exactly two structural states determined by which nucleotide is bound, however it has been shown that EF-Tu bound to either nucleotide (GDP or GTP) can adopt multiple conformations in solution [67, 68]. Additionally, recent structural studies have demonstrated that EF-Tu can adopt both open and closed conformations when bound to a nucleotide triphosphate analog or EF-Ts [68, 69]. Finally, recent kinetic and molecular dynamic simulation data suggests that the unbound or *apo* state of EF-Tu adopts a unique conformation which binds GTP and GDP indiscriminately [62]. This implies that in the case of nucleotide dissociation, EF-Tu's conformational change should occur before nucleotide release but this has not been verified biochemically [62].

Together, this suggest that the classical view of EF-Tu as a binary switch is outdated and should be replaced with one that reflects the highly dynamic nature of this molecule. However, currently little is known about the detailed intramolecular structural transitions from one EF-Tu conformation to another and how nucleotide binding and conformational changes are coupled.

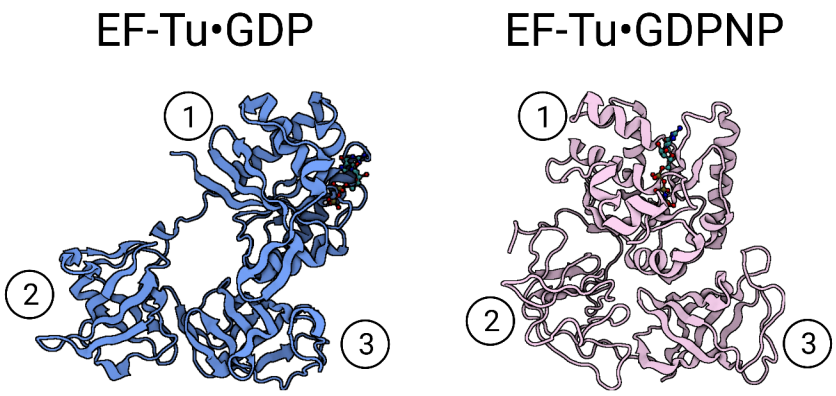

Figure 1.4. The conformation of EF-Tu depends on the bound nucleotide. EF-Tu•GDP from *E. coli* (Blue, 1EFC) and EF-Tu•GDPNP from *Thermus thermophilus* (Pink, 1EFT) are shown with the three domains numbered.

1.3 Objectives

This thesis aims to explore canonical and non-canonical facets of prokaryotic translation, shedding light onto this complex process. A particular focus focus will be on the evolution of the structural design principles exploited during two critical regulatory steps, translation initiation and EF-Tu mediated aa-tRNA delivery. To this end this thesis addresses two objectives:

1.) How does the structural dynamics of Elongation Factor Tu in solution relate to the available 3-dimensional information available and what are the mechanistic implications for aa-tRNA delivery?
2.) Does structure-based translation initial exist in bacteria, what are the evolutionary origins, and can it be used to design a minimal tuneable transition initiation signal?

To address these objectives, I will determine in Chapter 2 how the conformational changes of EF-Tu are related to nucleotide dissociation. In Chapter 3 I will describe eukaryotic translation in detail and how viruses are able to hijack translation machinery for their own proliferation. This chapter ends discussing the possibility of a universal structure-based translation initiation mechanism. In Chapter 4 I will answer questions posed in Chapter 3 by determining the molecular underpinnings of how a viral RNA is capable of initiating translation in prokaryotes. Finally, in Chapter 5 I will begin preliminary investigation of how a highly structured *E. coli* RNA, the *rpsA* TIR, can efficiently drive translation in *E. coli*.

Altogether, this thesis will provide information on the role of nucleotides in the conformational dynamics of EF-Tu, the robustness of the prokaryotic translational initiation machinery, and how the prokaryotic ribosome interacts with structured RNAs.

Chapter 2.

Nucleotide dissociation is limited by conformational change in EF-Tu

Roberts L., Kaur Dhalla D., Girodat D., and Wieden HJ.

2.1 Foreword

This chapter is written for submission to the Journal of Biological Chemistry and is therefore presented in manuscript form. As first author, I developed the scientific concepts, wrote and revised the manuscript, and performed all wet lab experiments. This includes but is not limited to cloning and site-directed mutagenesis, cell culture, protein purification, nucleotide purifications, fluorescent labelling, steady state fluorescent measurements, rapid kinetic measurements, and hydrolysis protections assays. Davinder Kaur Dhalla aided in optimizing the stopped-flow experimental conditions. Dylan Girodat performed Gaussian fitting of the EF-Tu populations, MD simulations of EF-Tu, and the subsequent analysis using the CINC pipeline. Hans-Joachim Wieden and Dylan Girodat aided in the structuring of the ideas and edited the manuscript.

2.2 Abstract

The translational GTPase Elongation Factor Thermo unstable (EF-Tu) is an essential molecular switch that delivers aminoacyl-tRNA to actively translating ribosomes in a nucleotide-dependent manner. This role is dependent on EF-Tu's ability to undergo several conformational changes including domain separation and rotation while cycling through its functional states. Although the thermodynamically most stable conformers have been described structurally, the conformational dynamics (i.e., the existence of and transitioning between different conformers) of EF-Tu responsible for the transitions between these conformations remain unclear. A detailed molecular-level understanding of the conformational dynamics of EF-Tu (and by extension other molecular switches), is essential for describing their function as molecular switches, allowing for their rational design and forward engineering in, for example, synthetic biology applications. Furthermore, as EF-Tu is the target of several antibiotics, a detailed description of how EF-Tu's function is rooted in its conformational dynamics will be of great benefit for the development of

novel antibiotics targeting EF-Tu. Thus, I rationally designed two reporter systems utilizing intramolecular hetero-FRET and changes in the local dye environment to monitor EF-Tu's structural transitions. These reporters have been used in both equilibrium and in real-time measurements. Using these approaches, I demonstrate that EF-Tu exists in an equilibrium of multiple conformations (including the previously described open and closed conformations) in the presence of guanosine triphosphate or guanosine diphosphate, respectively. Interestingly, when bound to nucleotide triphosphate analogs, EF-Tu more closely resembles the diphosphate conformations. Pre-steady state investigations reveal that nucleotide dissociation is rate-limited by the inherent propensity of EF-Tu to undergo conformational changes between the diphosphate and triphosphate conformation, and not *vice versa*. Together, our conformational analyses demonstrate that EF-Tu samples a wide range of conformations and that upon nucleotide binding, the sampling of this conformational landscape is biased to reflect the respective nucleotide-bound state. Additionally, I show that EF-Ts binding does not induce a large conformational change in EF-Tu, and I propose a new model for EF-Ts-mediated nucleotide exchange in EF-Tu.

2.3 Introduction

Molecular switches are biomolecules that regulate critical cellular processes through the adoption of two or more interconvertible states. In the cellular context, molecules contribute, for example, to gene expression regulation through their essential roles in signal transduction cascades, as riboswitches, and during ribosome-dependent protein synthesis [39, 70]. Molecular switches provide powerful tools for the rational design and manipulation of genetic circuits, including tools like photoinducible switches, chemically activated switches, or ligand-based protein switches [71, 72]. Due to their foundational role in many essential cellular processes, they are important targets for therapeutical interventions with great societal impact such as antibiotics

and cancer therapies[73]. Understanding at the molecular level how molecular switches undergo conformational transitions and the underlying design principles that govern these is essential for their rational design and forward engineering in, for example, synthetic biology applications [62]. Additionally, detailed understanding of the structural transitions will provide critical constrains for the development of novel antibiotics, demonstrated by the fact that existing antibiotics target molecular switches by modulating their conformational landscape [74, 75].

The essential, and well-studied prokaryotic translational guanosine triphosphatase (GTPase) elongation factor EF-Tu, is a 43 kDa three domain protein responsible for the codon-dependent delivery of aminoacyl-tRNA (aa-tRNA) to actively translating ribosomes. During its functional cycle EF-Tu forms a ternary complex with guanosine triphosphate (GTP) and aa-tRNA, this complex transiently binds to the ribosomal A site in a codon independent manner [19, 20]. Upon cognate codon-anticodon base-pair interaction the affinity of the ternary complex is increased and hydrolysis of GTP by EF-Tu is stimulated several orders of magnitude [65, 76-79]. The hydrolysis of GTP to guanosine diphosphate (GDP) and subsequent dissociation of the formed inorganic phosphate induces a conformational change in EF-Tu that ultimately results in the release of the bound aa-tRNA into the ribosome [23, 41, 43]. During this conformational change, domain 1 of EF-Tu rotates approximately 90^{o} and is displaced relative to domains 2 and 3, resulting in an open conformation in comparison to the relatively compact closed conformation of EF-Tu•GTP (Figure 2.1) [41, 42]. Upon binding to elongation factor (EF)-Ts, its guanine nucleotide exchange factor (GEF), the bound GDP dissociates and allows a new GTP molecule to bind, facilitating another cycle of EF-Tu dependent aa-tRNA delivery [80, 81]. Traditionally it was thought that the bound nucleotide defined the conformational state of the GTPase through a nucleotide-dependent induced fit model where EF-Tu acts as a binary molecular switch

transitioning between the two structurally distinct “on/GTP” and “off/GDP” conformations. However, mounting evidence suggests that nucleotide-bound EF-Tu (both GTP and GDP) samples a variety of structural conformations in solution and that an effector (i.e., the ribosome or aa-tRNA) rather than the bound nucleotide determines the conformation of EF-Tu [67, 68]. These two models are at odds as the former conceptualizes the nucleotide to act as a “timer” needing to dissociate before EF-Tu can bind another nucleotide and change conformations. The latter would have no such requirement and possibly even result in conformational change as a limiting factor to nucleotide dissociation. This poses a “chicken or the egg” style of question as to whether conformational changes are controlled by the dissociation of the bound nucleotide or whether the change in conformation limits nucleotide dissociation? While it has been proposed that a conformational change of EF-Tu precedes nucleotide dissociation from the binary complex leading to the unbound or *apo* state, it has not yet been experimentally validated [62]. Answering the question if conformational changes control the dissociation of the bound nucleotide or vice versa is critical for understanding how GTPases and other molecular switches function within their cellular context.

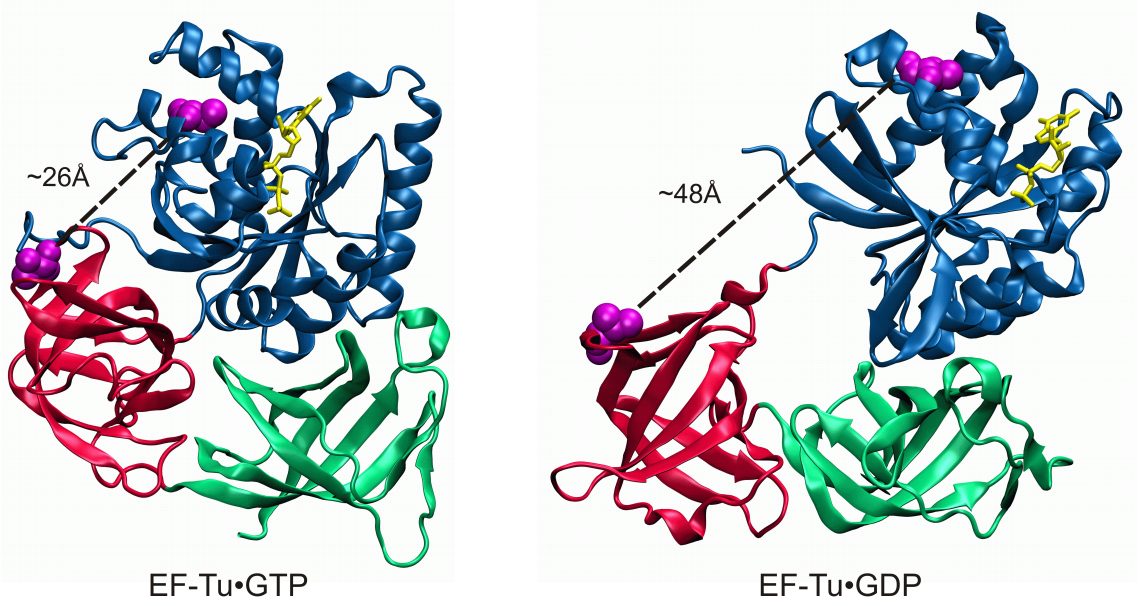

Figure 2.1. Structural comparison between the classical closed (GTP-like) and open (GDP-like) conformation of EF-Tu. The three domains of EF-Tu are shown: domain 1 or the G-domain (blue), domain 2 (red), and domain 3 (green). GDPNP and GDP are both shown as stick representations (yellow). Cysteines were introduced at positions T34 and L265 (purple spheres) and the spatial separation between the corresponding α-carbons is shown. Cartoon representations of EF-Tu bound to GDP and GDPNP were generated in VMD using PDB ID 1EFC and 1EFT respectively.

Here I investigate this question by using multiple fluorescently labeled EF-Tu variants with unique reporting properties (FRET and local fluorophore environment) in combination with rapid and equilibrium binding kinetics. I am able to confirm earlier findings that both GTP and GDP bound EF-Tu adopt multiple conformations in solution and show that EF-Tu's conformational changes (from nucleotide bound to *apo*) limit nucleotide dissociation, providing further evidence that the *apo* state adopts a unique state along EF-Tu's energetic landscape [62]. Additionally, I are able to support previous findings that the GTP analogs GDPNP and GDPCP both exhibit GDP-

like conformations in solution [67, 68]. Together, our data backs the model of "conformational selection", or population shift rather than an "induced fit" model for EF-Tu's conformational switching driven by bound nucleotides [68]. Additionally, I find that EF-Tu•GDP binding to EF-Ts does not induce a large conformational change, while EF-Tu•GTP binding to EF-Ts induces a very slow conformational change (~7400 fold slower than nucleotide release). This suggests EF-Ts does not significantly alter the conformation of EF-Tu upon nucleotide release, suggesting an alternate path for nucleotide exchange by sidestepping the rate limiting spontaneous changes in EF-Tu conformation before release of the nucleotide, and potentially following nucleotide binding (i.e., to and from the *apo* state).

2.4 Materials and Methods

2.4.1 Cloning and site-directed mutagenesis

All amino acid substitutions were introduced by mutagenesis of a pET21a plasmid with the coding sequence for a C-terminal His_6-tagged EF-Tu variant containing either a single cysteine (C81) or no cysteines using the Quikchange™ method (Agilent Technologies). All reactions were carried out using a $T_{Gradient}$ (Biometra) thermocycler and resulting mutant plasmids transformed into competent DH5α (New England Biolabs) cells as per manufacturer's instructions. Mutagenic primers were obtained from Integrated DNA Technologies and all mutations were confirmed by sequencing (Genewiz, DNA Sequencing Services). The forward (f) and reverse (r) primers used for mutagenesis are:

T34C-f 5`-GCAATCACT**TGC**GTGCTAGCTAAAACCTAC-3`

T34C-r 5`-GTAGGTTTTAGCTA**GCA**CGCAAGTGATTGC-3`

L265C-f 5`-ATGTTCCGCAAA**TGT**CTAGACGAAGGCCGTGCTGGT-3`

L265C-r 5`-ACCAGCACGGCCTTCGTCTAG**ACA**TTTGCGGAACAT-3`

E202C-f 5`-GGATTCTTACATTCCG**TGC**CCAGAGCGTGCG-3`

E202C-r 5`-CGCACGCTCTGG**GCA**CGGAATGTAAGAATCC-3`

2.4.2 Protein expression and purification

EF-Tu variants were expressed in *E. coli* strain BL21-DE3 (Novagen) and purified by Ni^{2+}-Sepharose (GE Healthcare) affinity chromatography followed by size exclusion chromatography (XK26/100 column; Superdex-75 prep grade (GE Healthcare)) as previously described [50]. EF-Tu purity typically >95% (Coomassie blue) and the final concentration was assessed by SDS-PAGE and densitometry.

2.4.3 Fluorescent labelling of EF-Tu variants

100,000 pmol of EF-Tu was thawed on ice and diluted 5-fold (~12 μM final concentration) in Buffer F (25 mM Tris-Cl pH 7.5 (4 ^{o}C), 7 mM $MgCl_2$, 30 mM KCl, and 20% (v/v) glycerol) before adding dropwise (on ice) a 10-fold molar excess of either a single dye (N-92-(Dansylamino)ethyl)maleimide (Dansyl)) or an equimolar mixture of two dyes (either Fluorescein-5-maleimide (F5M)) and rhodamine-red-maleimide (RRM) or 5-(2-iodoacetylaminoethyl)aminonaphthalene-1-sulfonic acid (IAEDANS) and *N*-(4-dimethylamino-3,5-dinitrophenyl)maleimide (DDPM)). The sample was then incubated at 4 ^{o}C for 4 hours or overnight with continuous inversion before being centrifuged at 21 000 xg for 5 min to pellet any precipitate that might have formed. The labelled protein was then separated from excess dye by size exclusion chromatography (XK16/20 column; Superdex-G25 (GE Healthcare)). Fractions

containing labeled EF-Tu were identified by SDS-PAGE, pooled and flash frozen with liquid nitrogen. Protein recovery from labeling was typically ≥80%.

2.4.4 Fluorescence spectroscopy

Labelled EF-Tu binary complexes were formed by incubating EF-Tu with a 100-fold molar excess of either GTP or GDP at 37 °C for 20 min. Complexes were then excited at either 493 nm (F5M/RRM), 560 nm (IAEDANS/DDPM), or 335 nm (Dansyl) and fluorescent emission was recorded using a fluorescence spectrophotometer (PTI) with either 2 nm/3 nm excitation and 2 nm/4 nm emission slit widths.

2.4.5 Gaussian fitting

Fluorescence spectra were plotted as histograms using a 1 nm bin size and Gaussian fitting was performed using the curve-fitting application in MATLAB ver. R2019b [82]. All graphs were fit with equation 1:

$$counts = Ae^{\frac{-(x-\bar{x}_a)}{2\sigma_a^2}} + Be^{\frac{-(x-\bar{x}_b)}{2\sigma_b^2}} \quad (1)$$

Where A and B are the peak heights of the histogram, $\bar{x}_a$and $\bar{x}_b$ are the mean value of the wavelength value, and σ_a and σ_b are the standard deviations for each histogram population as performed previously [62, 83]. During fitting, $\bar{x}_a$ and $\bar{x}_b$ were determined using a sliding window of 40 nm about the base of each distribution in the GTP distribution until they converged on a value, while σ_a, σ_b, A and B were boundless in the fitting. A 40 nm window around the mean value for GTP fluorescence was then used for fitting the fluorescence populations of different nucleotide complexes.

2.4.6 Theoretical FRET calculations

Theoretical FRET efficiencies (*E*) were calculated using equation 2:

$$E = \frac{1}{1+(\frac{R}{R_0})^6} \quad (2)$$

where R and R_0 are the inter dye distance and the Förster distance for the specific dye pair, respectively.

2.4.7 Purification of translational components

Ribosomes were purified from *E. coli* MRE600 as described previously [84]. [^{14}C]Phe-tRNA^{Phe} was prepared by aminoacylating *E. coli* tRNA^{Phe} (Sigma Aldrich) with [^{14}C]Phe (Roche Diagnostic) similarly to previously described [50, 84].

2.4.8 Rapid kinetic measurements

Fluorescence stopped-flow measurements were performed using a KinTek SF-2004 stopped-flow instrument. Nucleotide dissociation rate constants were determined as previously described [50]. In brief, EF-Tu was incubated with a 10-fold molar excess of MANT-GTP (2'-(or-3')-*O*-(*N*-Methylanthraniloyl) Guanosine 5'-Triphosphate, Thermo Fisher Scientific) or MANT-GDP (2'-(or-3')-*O*-(*N*-Methylanthraniloyl) Guanosine 5'-Diphosphate, Thermo Fisher Scientific) in TAKM$_7$ (50 mM Tris-HCl pH 7.6, 70 mM NH_4Cl, 30 mM KCl, 7 mM $MgCl_2$) at 37 °C for 20 min to form EF-Tu•MANT-GDP and EF-Tu•MANT-GTP binary complexes respectively. Twenty-five microliters of EF-Tu•MANT-GTP/MANT-GDP (0.15 μM after mixing) were rapidly mixed with 25 μL of a 10-fold molar excess of GTP/GDP at 20 °C in TAKM$_7$. Reactions containing MANT-GTP also contained 3 mM phosphoenolpyruvate (PEP) and 20 U/mL pyruvate

kinase (Roche Diagnostic) to ensure all nucleotides were in triphosphate form. MANT-nucleotides were excited via FRET from the tryptophan residue (W184) present in EF-Tu (λ_{ex} = 280 nm) and the resulting fluorescence was detected after passing through a LG-400-F long pass filter (NewPort) [45, 62]. The resulting fluorescent time courses were fit with a single exponential function (equation 3) using TableCurve software (Jandel Scientific) and, as the rate of dissociation is not dependent on concentration (first-order reaction), the apparent rate constant (k_{app}) is equal to the dissociation rate of the guanine nucleotides (k_{off}).

$$F = F_{\infty} + \mathrm{A}^{(-k_{app}\,\mathrm{t})} \tag{3}$$

To measure the conformational changes of EF-Tu, 25 µL of EF-Tu•GDP or EF-Tu•GTP (0.3 µM after mixing, prepared as described above) were mixed with 25 µL of EDTA (10 mM after mixing). Dansyl was excited either by FRET (as above) or directly at 335 nm and the resulting time dependence of fluorescence change was recorded after passing through a LG-350-F long pass filter (NewPort). The resulting traces were fit with a single exponential function (equation 3) or a single exponential function with an additional linear component (equation 4) using TableCurve (Jandel Scientific), yielding the apparent rate(s) of conformational change.

$$F = F_{\infty} + \mathrm{A}^{(-k_{app}\,\mathrm{t})} + Bx \tag{4}$$

2.4.9 Hydrolysis protection assay

Ternary complexes (EF-Tu•GTP•[^{14}C]Phe-tRNA^{Phe}) were formed as described previously[50]. Complexes were incubated at 37 °C and aliquots removed at various time points (0-100 min) spotted on trichloroacetic acid soaked and dried Whatman paper. The amount of [^{14}C]Phe retained on the filter was quantified using scintillation counting (Tri-Carb 2800TR Perkin Elmer Liquid Scintillation Analyzer). The obtained data was plotted as the natural logarithm of

the ratio of remaining [^{14}C]Phe-tRNA^{Phe} at any given time (c_t) relative to the initial concentration (c_0) as described previously [50].

2.4.10 Molecular dynamics (MD) simulations

Molecular dynamics simulations were performed as previously described [62, 83]. In brief, structural models of *E. coli* EF-Tu in the GDPNP-bound state were derived from homology modelling of the *T. aquaticus* EF-Tu obtained from protein data bank (PDB ID: 1EFT). In this model, GDPNP was manually converted to GTP. The GDP bound conformation of *E. coli* EF-Tu was obtained directly from the protein databank (PDB ID: 1EFC). Hydrogens were added to the system using the psfgen package within the NAMD software [85] and models were solvated with a TIP3P water box maintaining a 10 Å distance from any point on the protein using the Visual Molecular Dynamics (VMD) software [86]. The potential energy of each system was minimized using a two-fold reiterative process of minimizing the potential energy of water molecules followed by a minimization of the protein using CHARMM 27 parameters for 10,000 steps using NAMD [85]. Each EF-Tu conformation was subsequently neutralized with sodium and chloride ions followed by a final energy minimization of the entire system for 100,000 steps. EF-Tu systems were then equilibrated to 300 K and 350 K for 300,000 steps. Finally, MD simulations were performed at 300 K using velocities for the 350 K equilibration and atomic positions from the 300 K equilibration for 100 ns at a step size of 2 fs using Langevin dynamics to maintain temperature using the NAMD software.

2.4.11 Computational identification of non-disruptive conjugation sites (CINC)

Each simulation (100 ns) was put through the CINC pipeline (PCT Formal Patent Application No.: CA2021050110) to select positions for fluorescent labeling [87]. In brief, the

backbone flexibility, backbone dihedral angles, solvent accessibility, and distance to ligand were measured for each amino acid to develop a ranking score (Fscore). These properties were chosen as small dynamic features of the amino acid position that would have the largest impact on the environment of the conjugated fluorophore. Only amino acids with a conservation score of 1 from the ConSurf server [88, 89] and further than 5 Å from the ligand binding site were assigned an Fscore to ensure there would be no impact on protein function/stability and ligand binding properties. Flexibility was determined by measuring the root mean square fluctuation of the Cα atoms of each amino acid over the course of both the ligand bound and *apo* simulations. All parameters were measured using the VMD software [86].

2.5 Results

2.5.1 EF-Tu cysteine variant construction

To study the conformational dynamics of EF-Tu, I developed a fluorescent-based system to directly visualize changes in EF-Tu's conformation in real time. I designed double cysteine EF-Tu variants, based on a previously characterized Cys-less EF-Tu, which could be labeled with a pair of thiol specific and FRET compatible fluorophores [50]. In selecting positions for cysteine substitutions, a number of factors were considered. Firstly as EF-Tu is highly conserved, I selected residues which were less than 80% conserved between 125 bacterial species [50]. Secondly, residues were excluded if they were relatively large (i.e., tryptophan) or small (i.e., glycine), not surface exposed, or had been shown to have specific interactions with any of EF-Tu's interaction partners (aa-tRNA, EF-Ts, the ribosome, and guanine nucleotides). Finally, and perhaps most importantly, I selected amino acids pairs which had large (>15 Å) inter residue distance changes (measured from C_α to C_α using the crystal structures of EF-Tu•GDPNP•Phe-tRNA (1TTT) and EF-

Tu•GDP (1EFC)) between the two nucleotide bound states (Figure 2.1) [41, 43]. This left a limited number of locations for cysteine substitutions; therefore, the pair with the largest difference in inter residue distance between the two nucleotide bound states was selected (i.e., T34 and L265). In the GDP bound state, these positions are ~48 Å apart and in the GTP bound state, they are much closer at only ~22 Å (Figure 2.1). Consistent with our design strategy, these amino acid substitutions (i.e., EF-Tu $_{T34C\ L265C}$) and subsequent modification had no discernable effect on EF-Tu's interaction with guanine nucleotides or aa-tRNA (Table 2.1 and Figures S2.1 & S2.2).

Table 2.1. Nucleotide dissociation rates for various EF-Tu variants and fluorescently labelled EF-Tu. Dissociation rates measured via stopped-flow, the mean is reported and the error denotes one standard deviation.

EF-Tu	GTP dissociation k_{off} (s^{-1})	GDP dissociation k_{off} (s^{-1})
EF-Tu $_{WT}$	0.04 ± 0.01	0.002 ± 0.001
EF-Tu $_{T34C/L265C}$	0.02 ± 0.01	0.002 ± 0.001
EF-Tu ^{TDa}	0.02 ± 0.01	0.004 ± 0.001
EF-Tu ^{ID}	0.014 ± 0.002	0.003 ± 0.001
EF-Tu ^{Da}	0.015 ± 0.005	0.005 ± 0.001

2.5.2 Intramolecular hetero-FRET on double labelled EF-Tu

After confirming that EF-Tu $_{T34C\ L265C}$ did not affect EF-Tu's ability to bind nucleotides or aa-tRNA, I investigated whether FRET between the two dyes on a single EF-Tu molecule could be observed. Fluorescein-5-maleimide (F5M, λ_{ex} = ~493 nm, λ_{em} = ~518 nm) and rhodamine-red-maleimide (RRM, λ_{ex} = 560 nm, λ_{em} = 580 nm) were selected as a potential FRET pair with an R_0

(~60 Å) that should result in discernable FRET efficiencies for both of EF-Tu's nucleotide bound states. Using the distances derived from EF-Tu crystal structures bound to GTP/GDP and equation 2, I can estimate theoretical FRET efficiencies of 99% for the GTP-bound and 67% for GDP-bound conformations (Equation 2 and Table 2.2). Upon excitation of F5M/RRM labelled EF-Tu (EF-Tu^{FR}) at 493 nm, fluorescence emission peaks were observed at both ~525 nm and ~590 nm, representing fluorescent emission from both F5M and RRM, indicative of FRET (Figure S2.3A). The occurrence of FRET between the donor (F5M) and the acceptor (RRM) was validated by trypsin digestion, where a decrease in RRM fluorescence with a corresponding increase in F5M fluorescence (and blue shift back to 514 nm) is observed post digestion (Figure S2.3B). The donor dye alone (F5M) produced a signal peak with a maximum at 514 nm (Figure S2.3B) similar to the trypsin-digested EF-Tu sample.

While intramolecular hetero-FRET was observed, no significant differences between the GTP (EF-Tu^{FR}•GTP) and GDP (EF-Tu^{FR}•GDP) bound states were detectable (Figure S2.3A). This suggests that either our theoretical FRET efficiencies calculated based on the crystal structures of the proteins are not representative of the experimental conditions, that the distribution of conformations are similar under these conditions, or that our system is not sensitive enough to differentiate between the two conformations. The first option is likely partially true given I have incomplete information about the system (i.e., orientation of the dipole moments of the fluorophores), and given the heterogenous population of labelled EF-Tu's generated by the labelling procedure, it's likely I am observing an average of several distinct fluorescent signals. More accurate theoretical FRET calculations in this system would require MD simulations and quantum modeling as done previously in other systems [90, 91]. The latter, however, can be tested

by using a FRET pair with a shorter R_0, which would allow for the detection of conformation differences in EF-Tu based on the nucleotide bound state.

2.5.3 EF-Tu bound to tri-nucleotide analogs resemble EF-Tu●GDP in solution

To ensure that distinct conformations of EF-Tu can be detected using intramolecular FRET, the fluorescent dye 5-(2-iodoacetylaminoethyl)aminonaphthalene-1-sulfonic acid (IAEDANS , λ_{ex} = ~336 nm, λ_{em} = ~490 nm) and the fluorescent quencher *N*-(4-dimethylamino-3,5-dinitrophenyl)maleimide (DDPM) were selected as they have a reported R_0 that is roughly half the distance (~27 Å) of the F5M/RRM pair (Table 2.2). Upon excitation at 336 nm of GDP bound EF-Tu labelled with IAEDANS and DDPM (EF-Tu^{ID}•GDP), I observed a fluorescence peak at ~473 nm (Figure 2.2A). Interestingly, I observed ~4 times higher IAEDANS fluorescence (relative to GDP bound) when EF- Tu^{ID} is bound to GTP (Figure 2.2A) although the two positions should be ~22 Å closer (Figure 2.1). This was very surprising, as according to the theoretical FRET efficiency calculations (EF-Tu^{ID}•GDP, 4% and EF-Tu^{ID}•GTP, 76%), I would expect significantly less fluorescence in the "closed" GTP-like conformation when the quencher is closer to the fluorescent dye (Table 2.2).

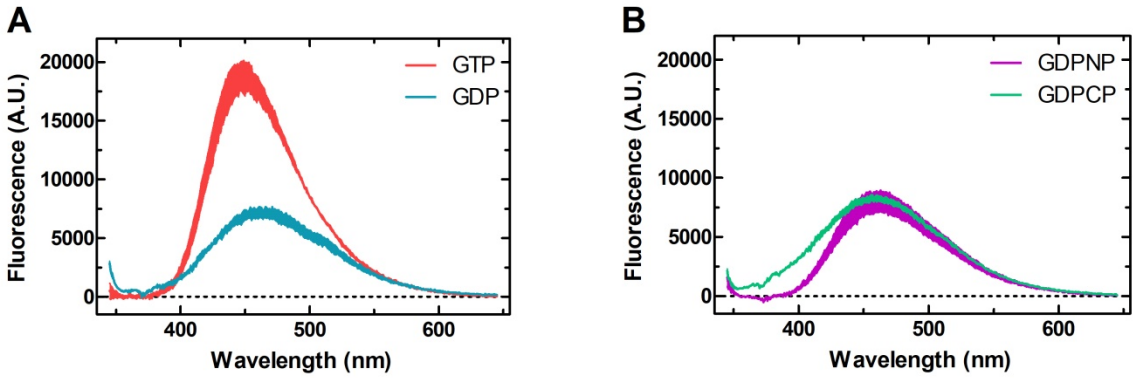

Figure 2.2. EF-Tu^{ID} produces distinct fluorescent levels in the GTP and GDP bound states. Fluorescent emission scans of double labeled EF-Tu^{ID} upon excitation at 336 nm. (A) Samples were incubated with either a 10-fold excess of either (A) GTP/GDP or (B) GDPNP/GDPCP prior to measurement, average of three replicates is plotted, the filled area denotes one standard deviation.

Table 2.2. Calculated theoretical FRET efficiencies for dye pairs on EF-Tu T34C L265C. Values calculated using Equation 2.

Fluorescent Dye Pair	Förster Distance (R_0)	Theoretical FRET Efficiency (EF-Tu•GTP)	Theoretical FRET Efficiency (EF-Tu•GDP)
F5M/RRM	60 Å	99 %	67 %
IAEDANS/DDPM	25-29 Å	76 %	4 %
Tryptophan/Dansyl	21 Å	31 %	1 %

Interestingly, when bound to GDPNP or GDPCP, EF-Tu^{ID} fluorescence levels appear GDP-like rather than GTP-like (Figure 2.2B). While this seems counterintuitive, it agrees with observations from Johansen *et al.* where a crystal structure of EF-Tu bound to GDPNP displayed an open conformation similar to the GDP-bound form [68]. This unexpected FRET was confirmed to not be the result of a difference in EF-Tu concentration but rather indeed a difference in FRET

using the trypsin digestion assay (Figure S2.4). Although a single Gaussian peak is observed for the trypsin digested EF-Tu^{ID} in the presence of either GTP or GDP, the broad emission peak in the intact protein with a maximum around 450 nm can be deconvoluted into two contribution populations using Gaussian fitting (Figures 2.3A and B). Comparing the GTP and GDP spectra reveals that they share statistically identical populations around 450 nm and 480 nm; the observed differences can be interpreted as a change in the population distributions (Figure 2.3B and Table 2.3). This suggests that EF-Tu exists in multiple conformations in solution, which is in agreement with previous single-molecule data [68] and supports the idea that the conformational dynamics of EF-Tu follows a conformational selection model in which the bound ligand biases the relative distribution across the respective conformational states [67, 68]. EF-Tu^{ID} bound to GDPCP and GDPNP also appeared to exist in multiple conformations in solutions (Figure 2.2B) and the Gaussian fit of EF-Tu^{ID}•GDPCP fluorescence spectrum is nearly identical to that of GDP (Figure 2.3C). Interestingly, EF-Tu^{ID}•GDPNP, while having a fluorescent signal similar to EF-Tu^{ID}•GDP, is unique its population distributions are "in between" those of EF-Tu^{ID}•GTP and EF-Tu^{ID}•GDP (Figure 2.3D and Table 2.3). Unfortunately, the detailed structural interpretation of this data is difficult as it could include four unique double labeled EF-Tu species resulting from the particular labelling approach and chemistry (i.e., EF-Tu with two donor dyes, EF-Tu with two acceptor dyes, and both variations of EF-Tu with an acceptor and donor dye).

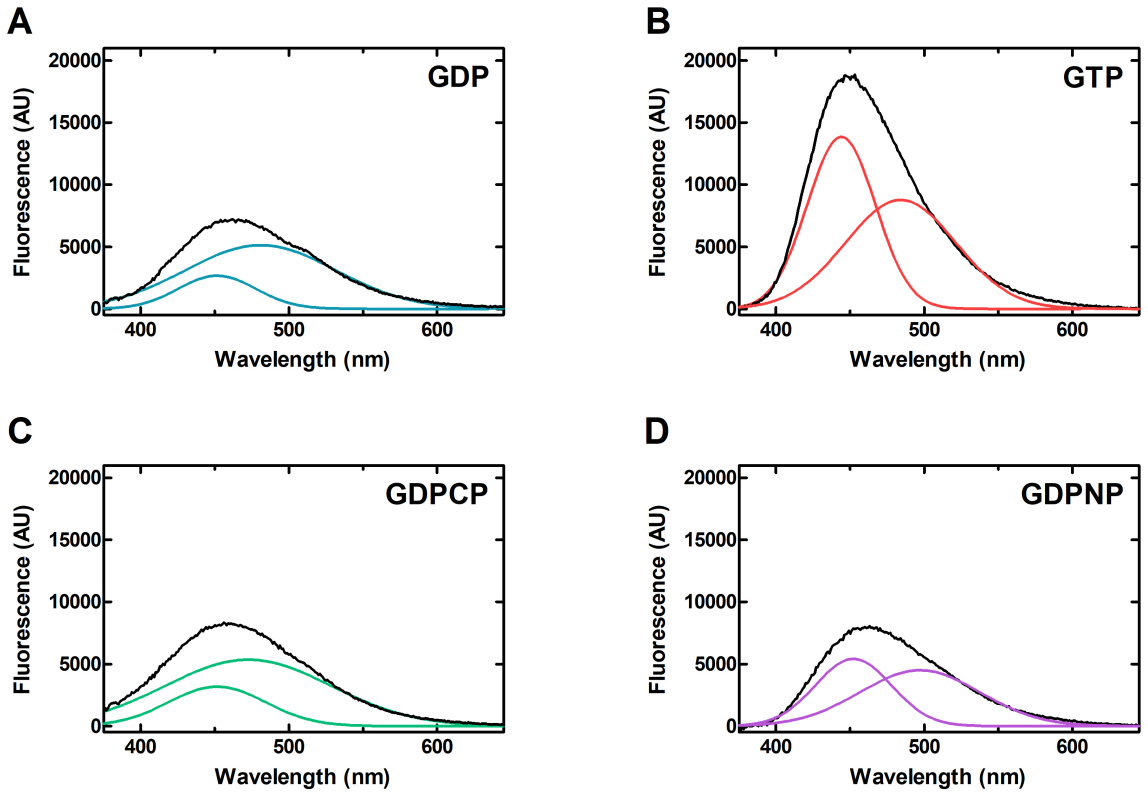

Figure 2.3. EF-Tu^{ID} exists in multiple conformations in all nucleotide-bound states. Fluorescent emission scans of double labeled EF-Tu^{ID} upon excitation at 336 nm (black traces). Gaussian fits showing possible populations for (A) GDP, (B) GTP, (C) GDPCP, and (D) GDPNP bound states (colored curves).

Table 2.3. Calculated fraction of EF-Tu^{ID} populations bound to guanine nucleotide and guanine trinucleotide analogs.

EF-Tu^{ID}	Fraction in Population 1 (λ_{max} = 450 ± 3 nm)	Fraction in Population 2 (λ_{max} = 484 ± 9 nm)
EF-Tu•GDP	0.21	0.79
EF-Tu•GTP	0.49	0.51
EF-Tu•GDPCP	0.26	0.74
EF-Tu•GDPNP	0.43	0.57

To help deconvolute these measurements, I decided to move to a dye system that exploits the presence of the naturally occurring tryptophan in *E. coli* EF-Tu, wherein the tryptophan is used as the fluorescence donor and *N*-[2-(dansylamino)ethyl]maleimide (Dansyl) as the acceptor. This allows us to measure intramolecular heteroFRET while covalently modifying only one position with a fluorescent dye. The selection criteria outlined above revealed the previously utilized position L265 as an excellent candidate. Upon excitation of EF-Tu^{TDa}•GDP at 280 nm, fluorescence emission for tryptophan (max ~325 nm) and little to no dansyl fluorescence were observed (Figure 2.4A), indicating no FRET is occurring in this state. Conversely, when EF-Tu^{TDa}•GTP is excited at 280 nm, a distinct dansyl peak (max ~440 nm) can be observed in addition to the tryptophan peak (Figure 2.4A). Interestingly (and consistent with EF-Tu^{ID} data), when EF-Tu^{TDa} is bound to the tri-phosphate analogs GDPNP and GDPCP, I observe minimal dansyl fluorescence, similar to the GDP-bound conformation (Figure 2.4B). This is again consistent with the nucleotide tri-phosphate analogs adopting a more GDP-like conformation in solution.

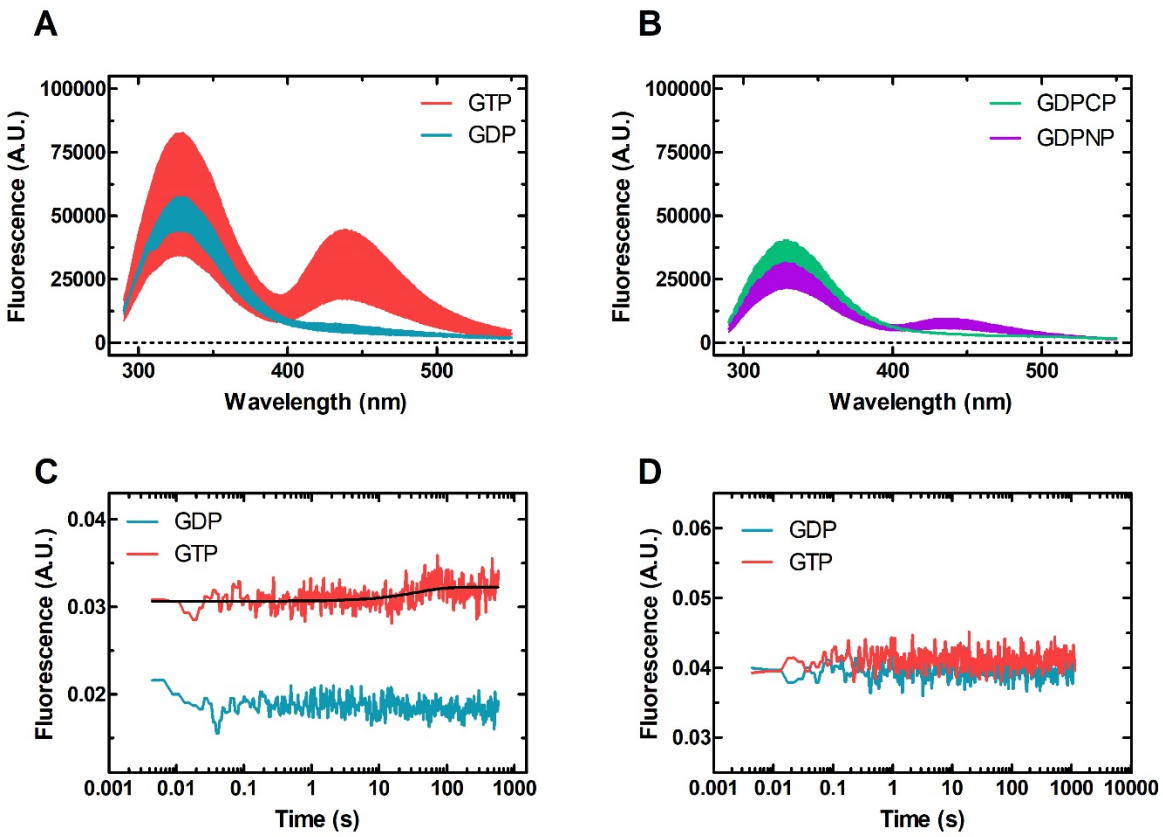

Figure 2.4. EF-Tu^{TDa} fluorescent signals in the GTP- and GDP-bound states. Fluorescent emission scans of single labeled EF-Tu^{TDa} upon excitation at 280 nm incubated with a 10-fold excess of either (A) GTP/GDP or (B) GDPNP/GDPCP prior to measurement; average of three replicates is plotted and filled area denotes one standard deviation. (C) Fluorescence change (FRET) of EF-Tu^{TDa} mixed with 10 mM EDTA upon excitation at 280 nm, data collected using a 400 nm long pass filter. For EF-Tu•GTP, a single exponential function fit to the data is shown (black trace). (D) Fluorescence change (direct dye excitation at 336 nm) of EF-Tu^{TDa} upon mixing with 10 mM EDTA, Fluorescence emission is recorded after passing through a 350 nm long pass filter.

2.5.4 EF-Tu conformational change in real time.

To investigate the kinetics of the conversion between the different structural states of EF-Tu (i.e., nucleotide bound to the *apo* state), I turned to pre-steady state kinetics using the stopped-flow approach. As nucleotide dissociation is naturally slow, I mixed EF-Tu bound to GDP or GTP against ethylenediaminetetraacetic acid (EDTA). EDTA will chelate the EF-Tu-associated magnesium, which coordinates the bound nucleotide phosphates, and thereby accelerating nucleotide dissociation [92]. Interestingly, when EF-Tu^{TDa}•GDP was mixed with EDTA, no change in dansyl fluorescence (i.e., a conformational change) was observed, although nucleotide dissociation was occurring (Figure 2.4C, Figure S2.5, & Table 2.4) on this time scale. This could suggest that the *apo* state and EF-Tu^{TDa}•GDP states may be similar in conformation, however this does not rule out conformational changes occurring outside our observable FRET-based "distance window" (e.g., starting at 0% FRET efficiency and moving farther away). EF-Tu^{TDa}•GTP however, in addition to starting at a higher initial fluorescence (consistent with equilibrium measurements, Figure 2.3A), undergoes a slow increase in fluorescence (Figure 2.4C) upon mixing with EDTA, with an apparent rate of 0.05 ± 0.02 s^{-1} (Table 2.4). This is interesting as this is an order of magnitude slower than nucleotide dissociation under the same conditions (Table 2.4) and is identical to the rate of spontaneous GTP dissociation (Table 2.1). These rates being indistinguishable suggests that the change in conformation is independent of nucleotide dissociation, inferring that EF-Tu conformational change may limit GTP dissociation in EF-Tu under normal conditions.

Table 2.4. Measured rates for EF-Tu^{TDa} rapidly mixed with 10 mM EDTA.

EF-Tu^{TDa}	Conformational Change k_{app} (s^{-1})	Nucleotide Dissociation k_{off} (s^{-1})
EF-Tu•GDP	-	0.45 ± 0.26
EF-Tu•GTP	0.05 ± 0.02	0.35 ± 0.19

While for EF-Tu^{TDa}•GDP no change in fluorescence was observed, EF-Tu^{TDa}•GTP revealed a small increase in FRET efficiency when moving from the GTP bound state to the *apo* state, suggesting a potential further compaction of EF-Tu following GTP release (Figure 2.4C). To assess, if the observed fluorescence change is caused by changes to the local environment of the acceptor dye, I directly excited the acceptor dye (dansyl), which resulted in no change in fluorescence for both GTP- and GDP-bound EF-Tu (Figure 2.4D), demonstrating that the observed signal is indeed due to changes in inter-dye distance.

2.5.5 Nucleotide dissociation is limited by conformational change in EF-Tu

While I successfully designed a FRET-based system capable of reporting conformational changes for EF-Tu^{TDa}•GTP, I failed to obtain such a system for EF-Tu^{TDa}•GDP using static structures (e.g., crystal structures, cryogenic electron microscopy) as a starting point to identify positions for engineering FRET pairs. I therefore turned to MD simulations to identify positions, that would give unique fluorescent signals based on local fluorophore environment, an approach previously developed in our lab [87, 93]. Using this MD simulation-based computational pipeline, I identified position amino acid E202 as a candidate (EF-Tu $_{E202C}$) based on a combination of its backbone flexibility, backbone dihedral angles, and solvent accessibility for further characterization (EF-Tu^{Da}).

Upon direct excitation of the dansyl group at position E202 (λ_{ex} = 335 nm), fluorescence with a λ_{max} = 500 nm is observed for EF-Tu^{Da}•GDP. However, when bound to GTP, a distinct increase in fluorescence (~1.5x) in addition to a large blue shift (λ_{max} = 475 nm) can be detected (Figure 2.5A). Consistent with all previous data, when EF-Tu^{Da} is bound to the tri-phosphate analogs GDPNP and GDPCP, I observe fluorescence spectra similar to the ones obtained for the GDP-bound conformation (Figure 2.5B). Although the fluorescence is slightly increased, no emission maxima blue shift can be observed.

Upon rapid mixing with EDTA, EF-Tu^{Da}•GTP undergoes a conformational change to a higher fluorescent state. The rate (0.018 ± 0.003 s^{-1}) of this conversion is nearly identical to the rate (0.05 ± 0.02 s^{-1}) measured using the FRET-based EF-Tu^{TDa} (Figure 2.5C). While the FRET-based assay (EF-Tu^{TDa}) did not exhibit a change in fluorescence for the EF-Tu•GDP complex, upon mixing with EDTA and using direct dye excitation (EF-Tu^{Da}), I was able to observe a larger change in fluorescence (~15%, compared to 5% for EF-Tu^{Da}•GTP) (Figure 2.5C). This conformational change is an order of magnitude slower (0.0016 ± 0.002 s^{-1}) than the previously measured rate for GTP-bound EF-Tu^{Da} and is identical to the native rate of GDP dissociation. Together, these data imply that the rates of spontaneous nucleotide dissociation and EF-Tu conformational change are linked.

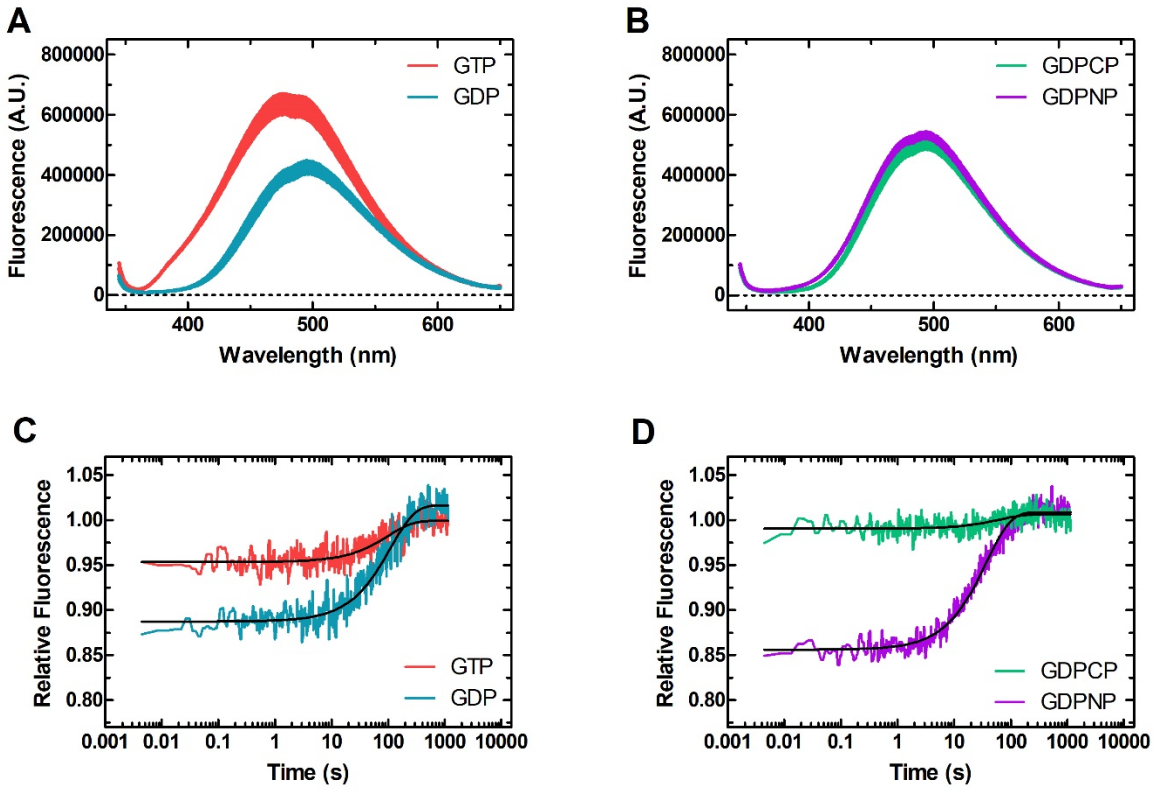

Figure 2.5. GTP- and GDP-bound states of EF-Tu^{Da} have different fluorescence properties. Fluorescent emission spectra of single labeled EF-Tu^{Da} upon excitation at 335 nm incubated with either a 10-fold excess of either (A) GTP/GDP or (B) GDPNP/GDPCP prior to measurement, average of three replicates is plotted and filled area denotes one standard deviation. Time resolved relative (to the end point) fluorescence change of EF-Tu^{Da} bound to (C) GDP/GTP or (D) GDPNP/GDPCP upon mixing with 10 mM EDTA.

Although EF-Tu^{Da}•GDPCP and EF-Tu^{Da}•GDPNP resemble the GDP-bound state in equilibrium measurements (Figure 2.5C), when they are rapidly mixed with EDTA, I observe a fluorescence increase suggesting conformational changes with apparent rates slightly faster (~2X) than for EF-Tu^{Da}•GTP, and an order of magnitude faster than for EF-Tu^{Da}•GDP (Table 2.5). While EF-Tu^{Da}•GDPCP undergoes a small conformational change (similar to GTP-bound), EF-

Tu^{Da}•GDPNP on the other hand undergoes a large change in fluorescence similar to GDP-bound when mixed with EDTA (Figure 2.5D).

As with EF-Tu^{ID}, the presence of multiple fluorescence peaks in the case of EF-Tu^{Da} indicates multiple conformations of EF-Tu contributing to both nucleotide-bound conformations. Similar to EF-Tu^{ID}, using Gaussian fitting, the recorded spectra can be best interpreted by a two-Gaussian fit representing two populations with maxima at ~484 nm and ~530 nm (Figure 2.6A). These two populations are present in EF-Tu^{Da}•GDP and EF-Tu^{Da}•GTP. Again, there are large shifts in the distribution, as the area under the 484 nm peak, representing the population on EF-Tu conformations contributing to this peak, is larger compared to the other population (i.e., that at 523 nm, Figure 2.6B). Gaussian fitting of the EF-Tu^{Da}•GDPCP and EF-Tu^{Da}•GDPNP reveals population distributions almost identical to those observed for EF-Tu^{Da}•GDP (Figure 2.6C, Table 2.6), again indicating that the trinucleotide analog bound EF-Tus appears GDP-like in solution.

Table 2.5. Rates of conformational change for EF-Tu^{Da} following nucleotide dissociation.

EF-Tu^{Da}	**Conformational Change k_{app} (s^{-1})**
EF-Tu●GDP	0.0016 ± 0.0002
EF-Tu●GTP	0.018 ± 0.003
EF-Tu●GDPNP	0.025 ± 0.005
EF-Tu●GDPCP	0.023 ± 0.005

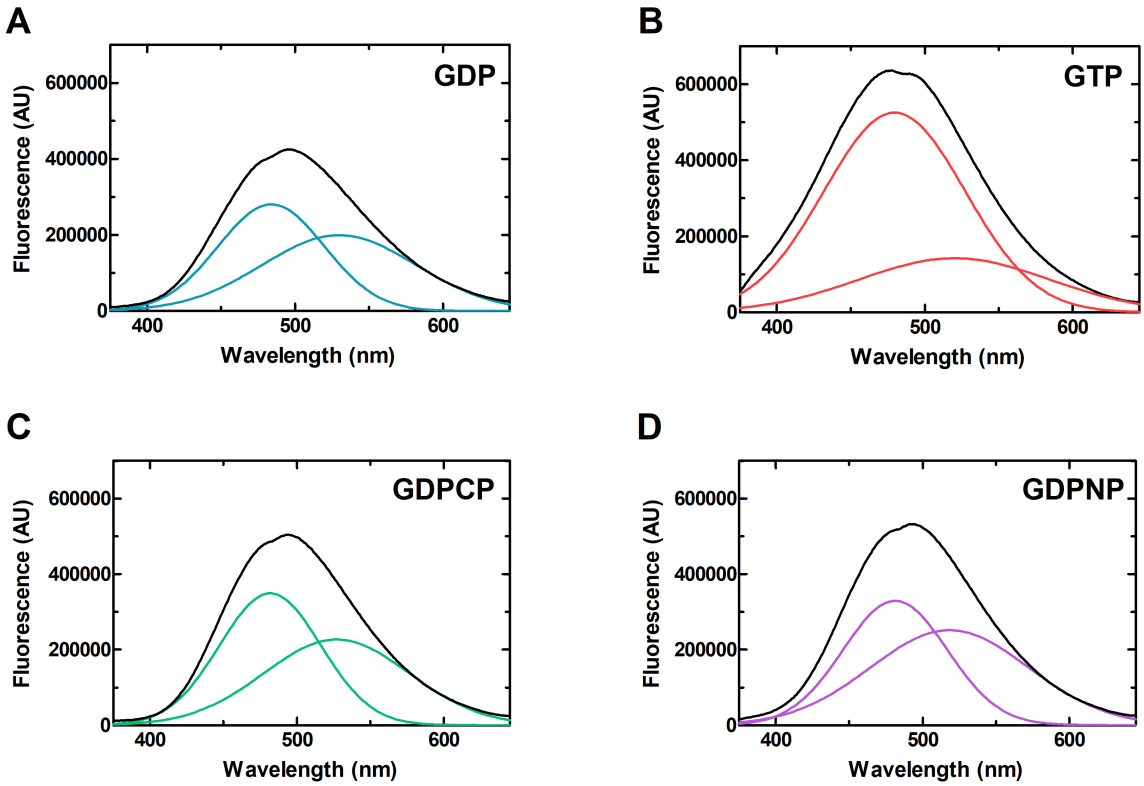

Figure 2.6. EF-Tu^{Da} exists in multiple conformations in all nucleotide-bound states. Fluorescent emission spectra of double-labeled EF-Tu^{Da} (λ_{ex} = 335 nm, black). Deconvolution of fluorescence emission using Gaussian fits (A) GDP, (B) GTP, (C) GDPCP, and (D) GDPNP bound states (colored curves).

Table 2.6. Calculated fraction of EF-Tu^{Da} populations bound to nucleotide and nucleotide analogs.

EF-Tu^{Da}	Fraction in Population 1 (λ_{max} = 480 ± 3 nm)	Fraction in Population 2 (λ_{max} = 524 ± 5 nm)
EF-Tu•GDP	0.49	0.51
EF-Tu•GTP	0.21	0.79
EF-Tu•GDPCP	0.46	0.54
EF-Tu•GDPNP	0.52	0.48

2.5.6 The rate of EF-Tu conformational change is independent of the EDTA concentration.

Under the experimental conditions used above (i.e., 10 mM EDTA), GTP and GDP are dissociating at the same rate ($k_{off, GTP}$ = 0.35 ± 0.19 s^{-1} and $k_{off, GDP}$ = 0.45 ± 0.26 s^{-1}, Table 2.4) which is ~10 and ~225 times faster than the measured conformational changes (Table 2.5). To ensure that the conformational changes I observed were not an artifact of our experimental set up (mixing with EDTA), I measured both nucleotide dissociation (MANT labelled nucleotides) and confirmation change at different concentrations of EDTA (direct dye excitation). Increasing EDTA concentration showed a corresponding change in nucleotide dissociation rate for both GTP and GDP, however the rates of conformational changes remained constant regardless of EDTA concentration (Figure 2.7, Table 2.7 and 2.8), demonstrating that the rates of conformational change are independent of EDTA concentration and the presence of the bound nucleotide. These conformational changes are also not observed when EF-Tu is mixed with buffer or with itself, demonstrating they are not due to dilution or the rapid mixing process (Figure S2.6).

Table 2.7. Concentration dependence of rates of conformational change and nucleotide dissociation for EF-Tu^{Da}•GDP.

[EDTA]	Conformational Change k_{app} (s^{-1})	GDP Dissociation k_{off} (s^{-1})
1 mM	0.0012 ± 0.0004	0.014 ± 0.002
5 mM	0.0009 ± 0.0001	0.36 ± 0.04
10 mM	0.0016 ± 0.0002	0.66 ± 0.05

Table 2.8. Concentration dependence of rates of conformational change and nucleotide dissociation for EF-Tu^{Da}•GTP.

[EDTA]	Conformational Change k_{app} (s^{-1})	GTP Dissociation k_{off} (s^{-1})
1 mM	0.013 ± 0.006	0.08 ± 0.03
5 mM	0.011 ± 0.006	0.30 ± 0.14
10 mM	0.018 ± 0.005	0.65 ± 0.20

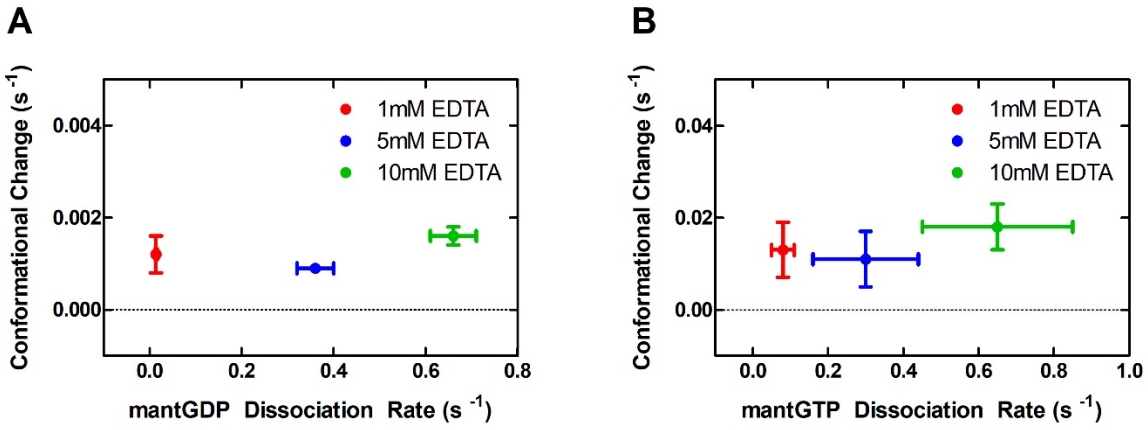

Figure 2.7. Nucleotide dissociation, but not conformational change, is sensitive to EDTA concentration. Rate of EDTA-induced nucleotide dissociation from EF-Tu^{Da} versus conformational change of EF-Tu^{Da} for bound to (A) GDP and (B) GTP upon mixing with increasing concentrations of EDTA.

2.5.7 EF-Ts-binding does not induce rapid conformational change in EF-Tu

While our system is able to accurately report the changes of EF-Tu from the nucleotide bound to the *apo* state, I was interested to see if our system could also report changes in EF-Tu's conformation when interacting with its cellular partners. Therefore, I investigated the conformational changes undergone by EF-Tu when interacting with its nucleotide exchange factor

EF-Ts. EF-Ts-stimulated nucleotide dissociation from EF-Tu^{Da} (GDP, k_{off} = 1.2 ± 0.3 s^{-1}) is as efficient as for wild type EF-Tu (GDP, k_{off} = ~1.6 s^{-1}, calculated from Gromadski *et al.* [45]), demonstrating that the introduced modification have not affected the ability of EF-Ts to interact with EF-Tu^{Da} (Figure S2.7) [45]. Equilibrium fluorescence measurements suggests that EF-Tu^{Da} in the EF-Tu^{Da}•EF-Ts complex behaves similar to the EF-Tu^{Da}•GDP-bound state under these conditions, which is in agreement with structural studies (Figure 2.8) [94]. Additionally, mixing EF-Tu^{Da}•GDP against EF-Ts resulted in no change in fluorescence, suggesting this complex does not undergo any structural changes (detectable by our system) upon EF-Ts binding (Figure 2.9A). Interestingly however, mixing EF-Tu^{Da}•GTP with EF-Ts results in a slow decrease in fluorescence, consistent with the equilibrium binding data (EF-Tu•EF-Ts complex is a relatively lower fluorescent state). However, this rate was too slow to be fit accurately at 22 °C, suggesting that although the nucleotide is dissociating significantly faster than in the presence of 10 mM EDTA, the change in conformation from GTP-like to EF-Tu•EF-Ts is slower than the change from EF-Tu-GTP to the *apo* state. Upon increasing the reaction temperature (to accelerate conformational changes) to 30 °C, I was able to fit the data with an exponential function (k_{app} 0.005 ± 0.001 s^{-1}) with an additional linear term which had not been required in previous lower-temperature experiments (Figure 2.9B). This linear term is likely either due to an increased rate of dansyl photo-bleaching, or there is an extremely slow change in conformation that appears linear over the time course of this measurement. This rate (0.005 ± 0.001 s^{-1}) is an order of magnitude slower than unstimulated GTP dissociation at 22 °C [92], and several orders slower than EF-Ts stimulated nucleotide dissociation at 30 °C (~37 s^{-1} estimated from Girodat *et al.*, Figure 3, panel B), again demonstrating that nucleotide dissociation does not limit conformational change in EF-Tu [62].

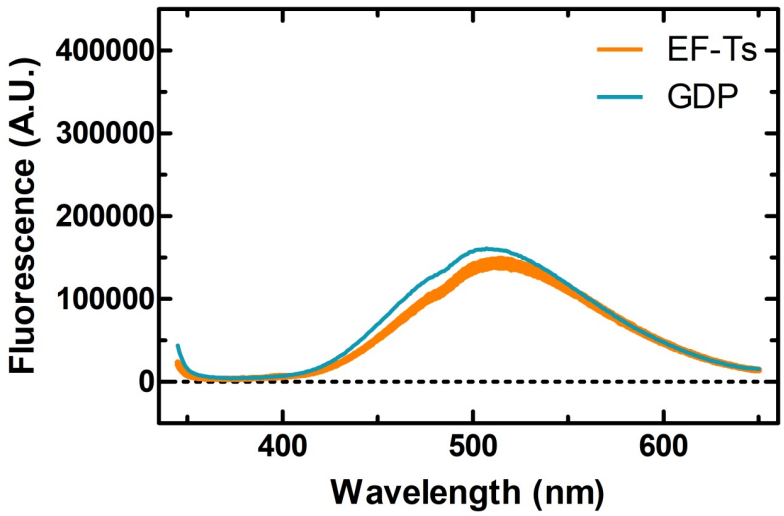

Figure 2.8. The EF-Tu^{Da} EF-Ts binary complex is GDP-like in steady state. Fluorescence emission scans of single labeled EF-Tu^{Da} upon excitation at 335 nm incubated with a 10-fold excess of either EF-Ts or GDP.

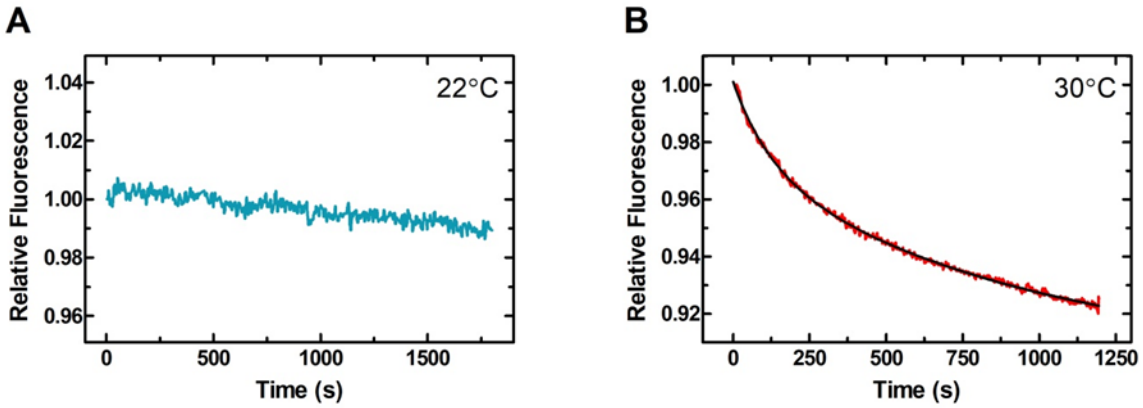

Figure 2.9. EF-Tu does not undergo rapid structural changes upon EF-Ts binding. Real-time fluorescence change of EF-Tu^{Da} bound to (A) GDP or (B) GTP upon mixing with EF-Ts.

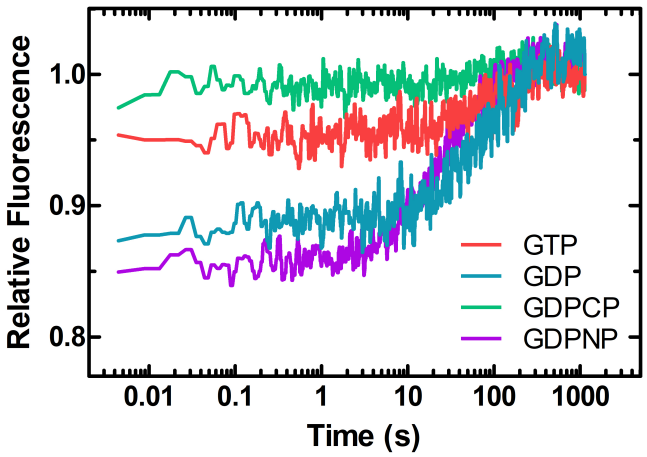

Figure 2.10. The *apo* state of EF-Tu is likely a unique conformation. Time courses of fluorescence change of EF-Tu^{Da} from its respective (GDP, GTP, GDPNP or GDPCP) nucleotide-bound form after induction of nucleotide dissociation by rapid mixing with 10 mM EDTA.

2.6 Discussion

EF-Tu exists in multiple conformations in all nucleotide-bound states

Considering the multitude of cellular partners EF-Tu interacts with (aa-tRNA, ribosomes, guanine nucleotides, etc.) and its moderate size (~43 kDa), it seems logical to infer that the conformational dynamics of EF-Tu play a key role in mediating these different interactions. Classically, EF-Tu (and other GTPases) have been thought of as a molecular switch that only exists in two states, where the bound nucleotide determines the structure and therefore the respective functional role of the GTPase (GTP=On/GDP=Off). However, work from us and others suggests that, rather than the single well-defined nucleotide-bound states observed in static structures, EF-Tu exists in a number of conformations in solution and likely only adopts its specific "active" conformation(s) upon interaction with a particular effector via conformational selection [42, 43, 68, 94]. Interestingly, the complexes containing the trinucleotide analogs (GDPNP and GDPCP)

also exist in multiple conformations in solution, but appear more GDP-like than GTP-like. This is again in agreement with previously reported single-molecule FRET data of doubly labelled EF-Tu, where multiple conformations were observed for EF-Tu in nucleotide-bound states predicted to be GTP-like [68].

Existing in multiple conformations would be beneficial from a design standpoint as EF-Tu could adopt a conformation to suit the current need (or cellular partner), regardless of the bound nucleotide. If EF-Tu was required to be in a distinct conformation to interact with each cellular partner, a significant number of encounters would be unproductive before a correct partner was identified. Conformational selection is also sensible when considering the speed at which EF-Tu must cycle through nucleotide exchange (~30 s^{-1} *in vivo*) to ensure protein synthesis is not bottlenecked [45]. A conformational selection (or induced fit) model of EF-Tu only adopting the "active" state once bound to the ribosome likely prevents incidental GTP hydrolysis and aa-tRNA dissociation.

Nucleotide dissociation is limited by conformational change to the apo state

Our data confirms a previous hypothesis that in the binary system (EF-Tu•nucleotide), the dissociation of the bound nucleotide would follow conformational change [62]. This suggests that the *apo* conformation of EF-Tu is a third unique conformation, separate from classical EF-Tu•GTP and EF-Tu•GDP. This is supported by our fluorescence data where all nucleotide-bound EF-Tu's undergo an increase in fluorescence rather than switching between two distinct states (Figure 2.10). However, it is possible that like the other observed nucleotide-bound states, the *apo* state may exist in two or more conformations, possibly selected by conformation induced by the previously bound effector (e.g., a nucleotide).

Interestingly, EF-Tu•GDPNP and EF-Tu•GDPCP exhibit rates of conformational change (to the *apo* state) slightly faster than EF-Tu•GTP although appearing GDP-like in equilibrium given the spectral resolution of our analysis (Table 2.6 and Figure 2.5). This increased rate of conformational change in addition to appearing more GDP-like in solution is likely due to the imperfect molecular mimicry of the non-hydrolysable analogs. However, this data suggests that the presence of a γ-phosphate plays a role in determining the rate of conformational change to the *apo* state. Switch I & II likely play a role in this as they have been shown to interact specifically with the γ phosphate of the trinucleotides [43] and their structural rearrangements are known to precede domain separation in EF-Tu [26, 95]. Switch I adopts two antiparallel β-strands in the GDP conformation that form contacts with domain 3 of EF-Tu. However, in the GTP conformation, it "switches" to two α-helices which coordinate the bound magnesium ion [42]. Switch II is an α-helix in both conformations and directly interacts with the γ-phosphate when bound to GTP [40]. Our equilibrium data is consistent with previous hypotheses that (for the native nucleotides) these motifs (Switch 1 & II) sense the nucleotide state of the molecule and help define the conformation of EF-Tu. However, I would expand on this hypothesis, suggesting that the initial structures of these motifs (and the corresponding energetic contributions [62]) are critical for determining the rate of conformational change and therefore nucleotide release. Additionally, the P-loop may be another factor that regulates nucleotide dissociation as it directly interacts with the γ-phosphate and its dynamics have been demonstrated critical for GTP dissociation [83]. Such considerations posit a mechanism where the P-loop may undergo behavioral changes when EF-Tu changes conformation that facilitates nucleotide release [83].

EF-Ts binding does not rapidly alter EF-Tu conformation

As EF-Tu is one of the most abundant components of translation, the ability to efficiently exchange nucleotide is extremely important. Interestingly, EF-Tu is unique among translational GTPases as it requires a GEF (EF-Ts), as it has unique thermodynamic contributions to nucleotide binding, which define its altered affinity for GDP compared to GTP (even compared to its eukaryotic homolog eEF1A) [62]. The co-evolution of an entirely novel factor (EF-Ts), rather than altering EF-Tu, suggests that evolutionary pressure exists to ensure EF-Tu's differential nucleotide affinities were maintained. While increasing the availability of nucleotides (GTP vs GDP) can drive the equilibrium of the EF-Tu•EF-Ts towards the binary complex (EF-Tu•GTP) that is able to bind to aa-tRNA (ensuring GTP binds EF-Tu after GDP release), it is unable to increase the turnover of the inactive binary complex (EF-Tu•GDP) or rescue inactive ternary complexes (EF-Tu•aa-tRNA•GDP). Therefore, in order to ensure translation is not bottlenecked while preserving the differential affinity EF-Tu has for GDP and GTP, a nucleotide exchange (i.e., EF-Ts) factor must be used [96]. Interestingly, the structure of the EF-Tu•EF-Ts complex has been solved in both the open GDP-like [81, 94] and closed GTP-like conformations [69], demonstrating its ability to exist in multiple conformations. Interestingly, our data demonstrates that EF-Ts does not induce a large change in conformation upon binding to EF-Tu•GDP or EF-Tu•GTP although nucleotide release is accelerated (Figure S2.7). This could be explained by small, relatively rapid structural changes (which would be invisible to our systems) in switches I, II, and the P-loop releasing the bound nucleotide. This would be consistent with many structures of EF-Tu•EF-Ts in which the switch regions are disordered, demonstrating that this complex destabilizes this region [69, 94].

Given the cellular concentration of GTP, the binding of GTP to the EF-Tu•EF-Ts complex would occur almost instantaneously, followed by EF-Ts dissociation [45], orders of magnitude faster than large-scale conformation changes measured here. This led us to propose a secondary

function for EF-Ts, in addition to its role in accelerating nucleotide dissociation, whereby it serves to "side step" the slow structural transitions of EF-Tu to and from the *apo* state (Figure 2.11). This would, at minimum, allow EF-Tu•GDP to avoid transitioning to the *apo* (or an *apo*-like) state for nucleotide dissociation, but would require conformational change upon GTP and subsequent aa-tRNA binding (Figure 2.11). This model also allows for rapid conversion to EF-Tu•GTP in the "open" conformation facilitating aa-tRNA binding, and active ternary complex formation. Given the strong affinity of aa-tRNA to EF-Tu•GTP (nanomolar), it is probable that EF-Tu•GTP in the open conformation is rapidly bound by aa-tRNA, ensuring EF-Ts does not accidentally rebind as the affinity of EF-Ts for both EF-Tu•nucleotide binary complexes is roughly the same [45, 97]. Such a model would prevent the association of deacylated tRNA to the EF-Tu•GTP complex as no stabilization of the "closed" conformation by the aa on CCA-end of the tRNA is possible.

This model is in line with the conformational selection model proposed for EF-Tu [68] and seems logical as requiring EF-Tu to fully adopt the open "GDP-like" conformation upon EF-Ts binding would significantly bottleneck translation. Additionally, this model would allow trapped EF-Tu•GDP•aa-tRNA ternary complexes to be easily rescued, potentially even while bound in the A site of the ribosome [96]. Finally, as EF-Ts stimulates nucleotide binding to EF-Tu roughly 10-fold (over EF-Tu alone), this "side step" pathway is even faster than avoiding the rate-limiting conformational changes alone [45].

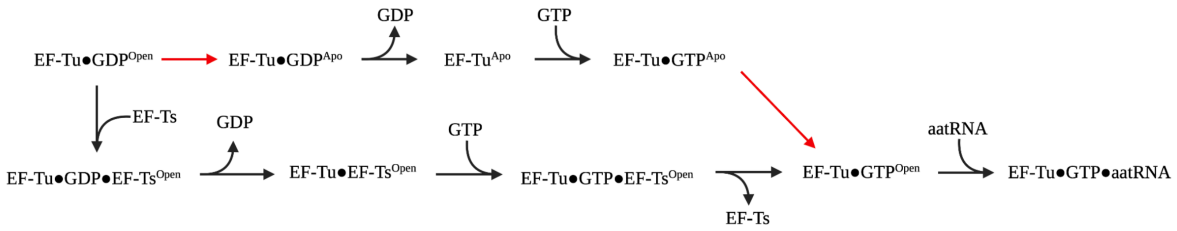

Figure 2.11. Proposed mechanism of EF-Ts mediated nucleotide exchange. Starting with EF-Tu•GDP in the open state (top left), the molecule can proceed through the *apo* pathway (top) with two slow conformational changes (red) or utilize the EF-Ts pathway (below) side stepping the *apo* state.

Understanding the role that protein conformational dynamics play for the function of biomolecular machines is currently at the forefront of biochemical research. Enzymes are dynamic molecules, and it is known that their ability to function relies heavily on this ability. The system reported here can be used to fill this knowledge gap for EF-Tu with respect to the nature of dynamic structures during its functional cycle. Using this method, the effects of sequence variation as well as physical parameters such as molecular crowding on EF-Tu conformational dynamics can be elucidated. Information about the molecular mechanisms governing EF-Tu's conformational dynamics can be extrapolated to similar GTPases and proteins in general. This system also has the potential to be used as a screening method for novel drug discovery, because the ability to monitor protein motions will allow us to identify small molecules capable of inhibiting or modulating these dynamic transitions. Finally, understanding the mechanisms which govern enzyme dynamicity would be a powerful tool for synthetic biologists as currently, rationally fine-tuning enzymatic activity and specificity is often achieved in an *ad hoc* manner.

2.7 Outlook

As it stands, a system for investigating the structural dynamics of EF-Tu has been generated and utilized to determine the order of events in nucleotide dissociation in the binary system. This data confirmed a previous theory from our lab in that the dissociation of the bound nucleotide would follow conformational change as the reverse is true for association to the *apo* state [62]. This suggests the *apo* state of EF-Tu is indeed one or more distinct conformations as opposed to being GDP or GTP-like. To follow up this work, it would be interesting to perform experiments with nucleotide-free EF-Tu^{Da} and monitor the conformational changes as nucleotides bind. Additionally, preliminary data suggest EF-Ts may "side step" the slow conformational changes to and from the *apo* state in order to further accelerate EF-Tu turnover. This system is now poised to investigate a number of questions about EF-Tu's structural dynamics and could even be used to screen compound libraries for potential inhibitors of EF-Tu's conformational dynamics (i.e., novel antibiotics). It would be interesting to determine nucleotide dissociation in the presence of the antibiotic kirromycin, which constrains EF-Tu structural dynamics. Additionally, of particular interest to myself would be the further investigation into EF-Tu's structural dynamics as a result of interactions with aa-tRNA, EF-Ts, and the ribosome (together and separately).

Chapter 3.

Viruses, IRESs, and a universal translation initiation mechanism

Reprinted with permission from “Viruses, IRESs, and a universal translation initiation mechanism”

Roberts L., Wieden HJ (2018). Viruses, IRESs, and a universal translation initiation mechanism.

Biotechnology and Genetic Engineering Reviews, Volume 34, Issue 1.

3.1 Foreword

This chapter is presented in the manuscript form that was published in the journal Biotechnology and Genetic Engineering Reviews (Biotechnol Genet Eng Rev) (Roberts & Wieden, 2018). As first author, I wrote and revised the manuscript. Hans-Joachim Wieden aided in the structuring of ideas and the writing of the manuscript. The style of citation has been updated from the original publication to maintain continuity in this thesis for the ease of the reader.

3.2 Abstract

Internal ribosome entry sites (IRESs) are cis-acting RNA elements capable of recruiting ribosomes and initiating translation on an internal portion of an mRNA. This is divergent from canonical eukaryotic translation initiation, where the 5' cap is recognized by initiation factors (IFs) that recruit the ribosome to initiate translation of the encoded peptide. All known IRESs are capable of initiating translation in a cap-independent manner, and are therefore not constrained by the absence or presence of a 5' m^7G cap. In addition to being cap-independent, IRES-mediated translation often uses only a subset of IFs allowing them to function independently of canonical initiation. The ability to function independently of the canonical translation initiation pathway allows IRESs to mediate gene expression when cap-dependent translation has been inhibited. Recent reports of viral IRESs capable of initiating translation across taxonomic domains (Eukarya and Bacteria) have sparked interest in designing gene expression systems compatible with multiple organisms. The ability to drive translation independent of cellular context using a common mechanism would have a wide range of applications ranging from agriculture biotechnology to the development of antiviral drugs. Here, we discuss IRES-mediated translation and critically compare the available mechanistic and structural information. A particular focus will be on IRES-mediated

translation across domains of life (viral and cellular IRESs), IRES bioengineering and the possibility of an evolutionary conserved translation initiation mechanism.

3.3 Background

3.3.1 Translation and translation initiation

The translation of information encoded in messenger RNA (mRNA) into proteins is an essential process performed by all living organisms. Translation is facilitated by the ribosome, a large ribonucleoprotein complex [3], and is divided into four major phases which are conserved throughout all domains of life: *initiation*, *elongation*, *termination*, and *ribosome recycling*. The ribosome (70S in prokaryotes, 80S in eukaryotes) is composed of a small (30S in prokaryotes, 40S in eukaryotes) and a large subunit (50S in prokaryotes, 60S in eukaryotes) [4-6] and contains three tRNA binding sites: the exit site (E site), the peptidyl site (P site), and the acceptor site (A site) [8]. During initiation, the ribosome is recruited and correctly positioned (i.e. the AUG start codon is situated in the ribosomal P site [98, 99]) on a target mRNA with assistance from several initiation factors (IFs) identifying the translational reading frame and the specific protein to be synthesized. Initiation is the most tightly regulated (and therefore a rate-limiting) step in protein biosynthesis [100], making it an effective target for bioengineering and medical applications.

Following initiation, the ribosome enters the elongation cycle, decoding the mRNA with the help of aminoacyl-transfer RNAs (aa-tRNAs) [101] and elongation factors (EFs) one codon at a time. A dedicated EF (EF-Tu in prokaryotes, eEF1A in eukaryotes) in complex with guanosine triphosphate (GTP) delivers aa-tRNAs to the translating ribosome [19, 102, 103] in a codon-dependent manner. If the anticodon of the aa-tRNA is complementary to the codon presented in

the ribosomal A site, GTP is hydrolyzed, the aa-tRNA is released into the ribosome [104] and the amino acid is added to the growing nascent polypeptide through the formation of a peptide bond [105]. Subsequently, the tRNAs, nascent polypeptide, and mRNA are translocated through the ribosome with the help of another GTP-dependent EF (EF-G in prokaryotes, eEF2 in eukaryotes) positioning the downstream codon in the A site, and the decoding process is repeated [106, 107]. This cycle continues until the end of the coding sequence (CDS) is reached (designated by a stop codon (UAA, UAG, or UGA)). Upon recognition of a stop codon by release factors, translation terminates and the newly synthesized protein is released from the ribosome. Finally, the translational components are recycled by ribosome recycling factors so that they can reassemble and begin a new round of translation [35, 37]. Although the general steps of protein synthesis are conserved between all organisms, the molecular mechanisms involved, as well as the associated cellular factors, vary drastically between taxonomic domains.

3.3.2 Prokaryotic translation initiation

As translation is cyclic, initiation begins with the final stage of translation, ribosome recycling. IF3 association to the 30S subunit of a 70S ribosome promotes the dissociation of the subunits [9], generating a complex (30S•IF3) to which IF1, an IF2 ternary complex (IF2•GTP•fMet-tRNA$_f^{Met}$) [10], and the translation initiation region (TIR) of an mRNA bind, forming the 30S preinitiation complex (30S PIC (30S•IF1•IF2•GTP•IF3•fMet-tRNA$_f^{Met}$•mRNA) [14]. The Shine–Dalgarno (SD) sequence on the mRNA interacts with the anti-SD sequence on the 16S rRNA, positioning the 30S PIC complex near the correct start codon, allowing the P site-bound fMet-tRNA$_f^{Met}$ to form a cognate codon–anticodon duplex [108]. Correct start codon recognition results in IF1 and IF3 dissociation and stable 30S initiation complex

(30S•IF2•GTP•fMet-tRNA$_f^{Met}$•mRNA) formation [16]. Finally, the 50S subunit associates, IF2 hydrolyzes GTP and dissociates [17], resulting in the 70S initiation complex (70S•fMet-tRNA$_f^{Met}$•mRNA) that proceeds to elongation [18].

3.3.3 Eukaryotic translation initiation

In eukaryotes, canonical translation initiation is complex and highly regulated compared to its prokaryotic counterpart. Initiation begins with the formation of a ternary complex composed of eukaryotic (e)IF2, GTP, and methionine-charged initiator tRNA (Met-tRNA$_i^{Met}$). With the help of eIFs 1 and 1A, the ternary complex associates with a complex containing the 40S ribosomal subunit and eIF3, forming the 43S pre-initiation complex (43S PIC (40S•eIF1•eIF1A•eIF3•eIF5•Met-tRNA$_i^{Met}$)) [109]. Formation of the 43S PIC results in the correct positioning of the Met-tRNA$_i^{Met}$ in the ribosomal P site. The 43S PIC is then recruited to the 5' end of an mRNA by the eIF4F complex (eIF4A, eIF4E, and eIF4G) and eIF4B in a 7-methylguanylate (m^7G) cap-dependent manner, forming the 48S PIC (40S•eIF1/A•eIF3•eIF4B•eIF4F•mRNA•Met-tRNA$_i^{Met}$). Once bound to the mRNA, the 48S PIC scans the mRNA in a 5'– 3' direction until a start codon (flanked by a Kozak sequence [110]) is located through a cognate codon–anticodon interaction with Met-tRNA$_i^{Met}$, resulting in the correct positioning of the ribosome at the beginning of the CDS to be translated. In addition to helping position the ribosome, the cognate codon–anticodon interaction also facilitates hydrolysis of GTP by eIF2 and the subsequent dissociation of eIF2•GDP. Finally, the remaining eIFs are displaced by the 60S ribosomal subunit with the help of eIF5B, forming the 80S initiation complex (80S•mRNA•Met-tRNA$_i^{Met}$) and translation proceeds to the elongation phase [109]. This method

of eukaryotic translation initiation is referred to as cap-dependent translation as it depends on the presence of a 5' m^7G cap on the mRNA.

3.4 Internal ribosome entry sites

While the cap-dependent mechanism is the primary mode of translation initiation in eukaryotes, a number of cellular mRNAs and viral RNAs are translated via specific RNA sequences referred to as internal ribosome entry sites (IRESs) [111]. IRESs are *cis*-acting RNA elements originally identified in the 5' untranslated region (UTR) of picorna virus RNAs [112, 113]; however, IRESs can be found far from the 5' end of an mRNA and even in the intergenic regions of polycistronic mRNAs [114]. IRESs are capable of initiating translation on an internal portion of an mRNA and are therefore not constrained by the presence/absence of a 5' m^7G cap [115]. In addition to being cap-independent, IRES-mediated translation often uses only a subset of the canonical IFs, allowing them to function independently of canonical initiation [115]. The ability to function independently of the canonical translation initiation pathway allows IRESs to mediate gene expression when cap-dependent translation has been inhibited (e.g. viral infection) [111].

3.4.1 Viral IRESs

Viral (v)IRESs are found in a wide variety of clinically relevant viruses (e.g. the Hepatitis C virus (HCV), the Human Immunodeficiency virus (HIV)), and have been classified into four mechanistic taxa based on structure, the IFs utilized, and the additional IRES trans-acting factors (ITAFs) required for functionality (Figure 3.1) [115, 116]. Despite the high-sequence diversity between mechanistic classes, there is a general trend that vIRESs with tightly folded RNA

structures (i.e. the most compact) require the fewest IFs [115]. While all vIRESs perform the same general task (cap-independent ribosome recruitment), IRESs from different mechanistic classes utilize different recruitment mechanisms owing to the differences in IF requirement [115].

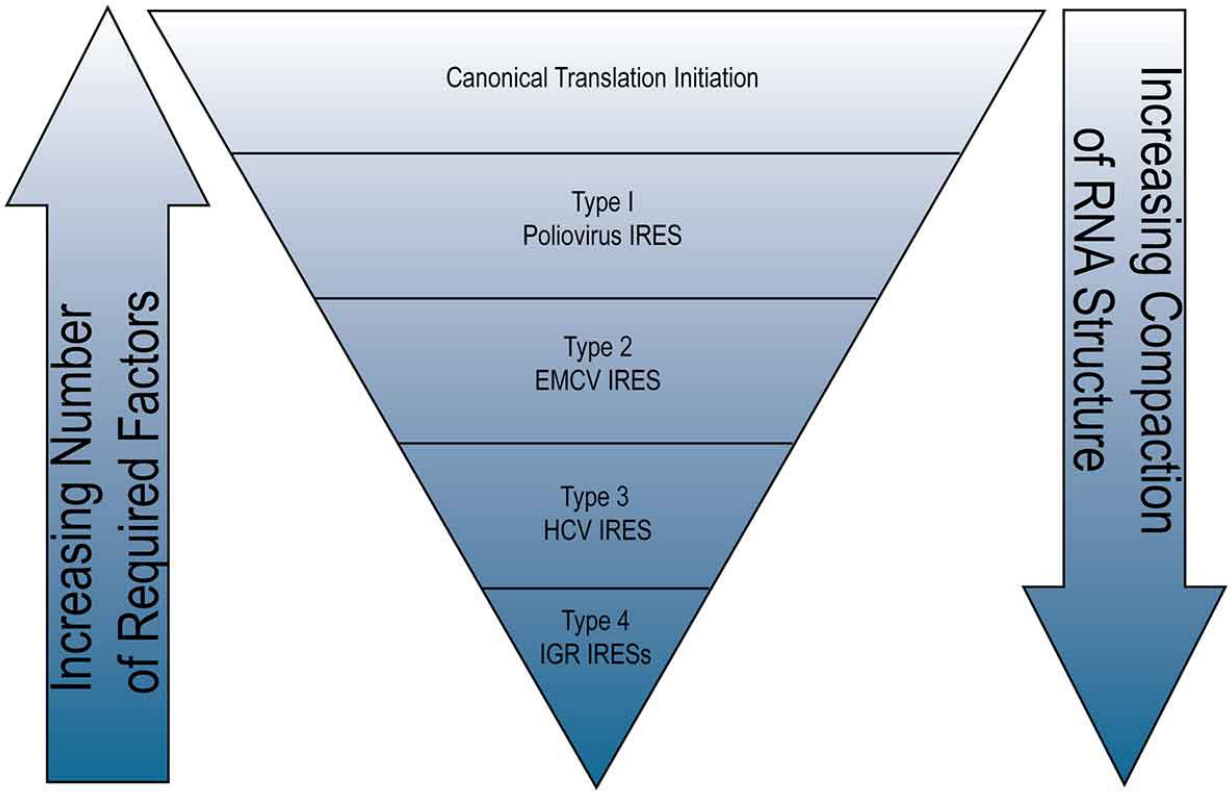

Figure 3.1. Comparison of the four classes of viral IRESs and canonical cap-dependent translation.

The first class of vIRESs is exemplified by the poliovirus IRES, which is composed of five core domains [117]. These domains facilitate interactions with eIF4A, eIF4B, eIF4G (either full length or truncated via viral protease 2C [118]), and the ITAF polyC binding protein-2 in order to recruit the 43S PIC [119]. Once recruited, this 48S-like complex then proceeds through the remainder of the canonical initiation pathway. The second vIRES class, represented by the encephalomyocarditis virus (EMCV) IRESs, contains four core domains and operates very

similarly to the first class in terms of mechanism. However, this class utilizes a different set of ITAFs. Instead, these IRESs utilize the polypyrimidine tract binding protein and ITAF45 in order to facilitate recruitment of the 43S PIC. Additionally, these vIRESs do not scan the mRNA. Instead, the 48S-like complex is recruited directly to the start codon [120, 121]. Viral IRESs belonging to the third class (HCV and HVC-Like IRESs) utilize a largely reduced set of protein factors (relative to the first two classes) and are composed of three domains which are responsible for direct (initiation factor-independent) small subunit recruitment. Following ribosome recruitment, an initiator tRNA is recruited through eIF1A, eIF2, eIF3, eIF5, and eIF5B to form a 48S-like complex directly on the start codon, similar to the second family of IRESs [122-124]. It is important to note that for the HCV and HVC-Like IRESs, multiple unique pathways have been described for competent ribosome complex formation [123, 125-128]. Under conditions with limited eIF2 the HCV IRES can utilize an alternate assembly pathway using eIF5B [123]. Additionally, *in vitro* (under high magnesium concentrations) the HCV IRES can initiate translation in the absence of initiation factors [129]. The fourth, and arguably best characterized class of vIRESs are the intergenic region (IGR) IRESs from the *Dicistroviridae* virus family, which do not require additional factors for translation initiation. The IGR IRESs are composed of three RNA pseudoknots (PK) and several stem loops [130, 131]. Mutagenic analysis [131, 132], as well as recent structural determinations of the IGR IRES in both its free [133]and ribosome bound [134-139] forms, have provided critical information regarding the function of each pseudoknot. PKI interacts directly with the decoding center (in the ribosomal A site), adopting a structure which mimics a cognate codon–anticodon interaction between a tRNA and an mRNA [135]. This precise molecular mimicry allows the IGR IRES to recruit a 60S subunit and initiate on a non-AUG start codon, essentially ‘jump-starting’ into elongation. For a detailed review of vIRES structures and

their intricate interactions with the ribosome, please see the recent review by Yamamoto and colleagues (2017) [140].

3.4.2 Cellular IRESs

IRESs have also been discovered in many cellular mRNAs; however, their mode of function and cellular role remain largely uninvestigated and highly debated (compared to vIRESs) [141]. Cellular (c)IRESs are often found in stress response genes as well as several oncogenes, and it has been proposed that this allows for their expression during global inhibition of translation (e.g. apoptosis) [142]. Interestingly, cIRESs share no detectable sequence similarity or predictable secondary structure, making their discovery and comparative analysis extremely challenging [141]. Despite the substantial differences, cIRESs do share ITAFs hinting toward at least partially common mechanisms [143]. Although this field is still very much in its infancy, one can speculate about cIRESs classification using insights from several works. Firstly, two classes of cIRESs, one with complex global RNA folds and a second composed of short sequence elements, were identified utilizing high-throughput screening assays [144]. Secondly, the X-linked inhibitor of apoptosis cIRES was shown to initiate translation in a structurally conserved viral-like mode [145]. Thirdly, Cellular IRESs have been suggested to recruit specialized ribosomes when cap-dependent translation is inhibited [146, 147]. Finally, mRNA modifications (specifically m^6A) have also been shown to enable mRNAs to bypass cap-dependent translation inhibition [148]. Taken together, this suggests that at least four classes of cIRESs exist, and that similar to vIRESs, the mechanisms of translation initiation (including the number and identity of required ITAFs) varies between the

classes. For a detailed overview of cIRES translation mechanisms and their implications, please see the review by Sharma and colleagues [149].

3.5 Universal translation initiation signals

Considering the existence of structure-based initiation in cIRESs and vIRESs, and the conserved structural features in the ribosomal core [150], an RNA structure-based evolutionary conserved cross-kingdom translation initiation signal, may exist. Experimental support for such a conserved signal has recently been described for two viral RNAs [134, 151].

Telpalo-Carpio *et al.* used a tricistronic reporter construct to assess whether the EMCV vIRES is capable of initiating translation in *E. coli* [151]. Although the authors observed production of fluorescent proteins (mCherry and eGFP) *in vivo*, it is not clear if this translation is driven by the structure of the vIRES or by the canonical bacterial translation mechanism (SD-mediated). While the authors did search for cryptic RBS sites using the Salis lab RBS calculator [152], they used only a subset of the 5' UTR for these predictions, excluding the majority of the 5' UTR. Therefore, we argue that a more detailed mechanistic analysis is required to determine if the vIRES has *bona fide* IRES activity in a prokaryotic context. Colussi *et al.* recently reported a crystal structure of an IGR IRES bound to a 70S ribosome [134]. The vIRES exhibited a binding mode consistent with similar IGR IRESs bound to 80S ribosomes (at the universally conserved tRNA binding sites) [135, 137, 139]. The authors also reported that the IGR IRES exhibits structure-dependent translational activity in *E. coli* similar to that described for the IGR IRESs in

eukaryotic contexts [132, 134]. The authors conclude that the conserved RNA structure present in the IRES constitutes a universally conserved structure-based translation initiation signal. However, contrary to the mechanism of IGR IRES initiation in eukaryotes, initiation in *E. coli* requires the presence of a downstream start codon. Interestingly, disruption of a pseudoknot required for eukaryotic IGR IRES activity (PK1) had little to no effect on the translation in a prokaryotic context [132, 134] suggesting unique mechanisms for translation initiation.

Currently, no naturally-occurring examples of structure-based initiation in bacteria or archaea have been reported. There are, however, a number of similarities between the domains of life with respect to the mechanisms by which translation is initiated. Several IRESs, including viral, mammalian, and synthetic IRESs, base pair to the 18S rRNA [153]. Interestingly, base-pairing serves different functional roles for different IRESs as the HCV IRES requires several other structured elements to initiate translation [126], while a 9 nt segment of the Gtx homeodomain mRNA with perfect complementarity to the 18S is seemingly able to initiate translation independent of other factors [153]. This indicates that viral and cellular IRESs may be able to (at least in part) utilize common mechanisms of translation initiation. These examples of Watson-Crick base pairing in eukaryotic IRESs are reminiscent of the prokaryotic SD:anti-SD mechanism [108, 154]. Further parallels between domains can be drawn when considering that specialized ribosomes can be recruited through non-canonical (non-SD) mechanisms to efficiently translate prokaryotic mRNAs [155]. For example, the stress response factor MazF cleaves the anti-SD sequences from *E. coli* ribosomes, allowing them to specifically translate likewise MazF-processed (leaderless) mRNAs [155]. Interestingly, most archaeal mRNAs are leaderless or lack canonical SD sequences. Whether their translation is due to specialized archaeal ribosomes or some other, structure-based, mechanism remains unclear [154].

3.5.1 Implications

The discovery that a structured viral RNA could universally initiate translation challenges the paradigm that prokaryotic and eukaryotic translation initiation are mutually exclusive. Considering that rRNA base-pairing and specialized ribosomes exist in both prokaryotes and eukaryotes, it is reasonable to assume that structure-based translation initiation may exist naturally in prokaryotes as well. The existence of a universal translation initiation signal is very appealing, in particular from a synthetic biology/bioengineering point of view. While a number of standardized post transcriptional gene expression control elements are available, their activities are often specific to organisms and sometimes even specific to cell lines. This limits the ability to port these designed biological devices from one organism to another. A multiplatform (multiorganism) gene expression system is highly desirable as fine-tuning gene expression is critical for biotechnology applications aimed at altering cellular behaviors, such as the production of value-added compounds (e.g. biofuels). Additionally, the ability to create, test, and debug a biological device in one organism (prototyping) and transfer the device into another organism without the need to change and inevitably re-test the actual device is highly desirable. A universal structure-based translation initiation signal would be a considerable step towards this goal.

3.6 IRESs in synthetic biology

Although a promising bioengineering target, surprisingly little IRES engineering has been reported to date. Current efforts can be classified into two major categories: identifying novel IRESs through randomized oligonucleotide pools (bottom-up) and re-engineering existing IRESs (top-down) (Table 3.1).

3.6.1 Bottom-up approaches: IRES discovery

Randomized oligonucleotide libraries have been used in combination with bicistronic reporter constructs to identify novel sequences that could act as IRES elements. The first to report this were Owens *et al.*, who utilized a bicistronic reporter construct with a randomized 18-nt library located in the intergenic region between genes coding for enhanced cyan fluorescent protein and enhanced green fluorescent protein [156]. Using fluorescence-activated cell sorting to isolate sequences with potential IRES activity, the authors identified two IRESs with complementarity to the 18S rRNA (IRES activity through Watson–Crick base pairing, *vide supra*). In the same year, Arun Venkatesan and Asim Dasgupta used an applied fluorescence-based screen to identify novel synthetic IRESs [157]. The authors again utilized a randomized oligo library, however, the sequence space was increased to 50 nt. Using this extended library, the authors were able to identify two novel IRESs. Interestingly, the identified sequences had no significant complementarity to the 18S rRNA, suggesting a unique mechanism of translation initiation [157]. Zhou *et al.* identified sequences that functioned as IRES elements in yeast using a unique bicistronic assay where randomized 18nt sequences were inserted between an upstream renilla luciferase gene and a downstream *HIS3* gene [158]. Transforming this plasmid into yeast with a *HIS3* knockout enabled IRES enrichment through viability selection instead of fluorescence output. The authors identified 56 sequences with IRES activity, many of which contained extensive complementarity to the 18S rRNA [158].

Table 3.1 Bioengineering of novel and natural IRESs

IRES	Approach	Methodology	Output	Regulation	Organism	Reference
Novel	Bottom-up	Randomized 18 nt	2 Novel IRESs	rRNA Base pairing	*M. musculus*	Owens *et al.,* 2001
Novel	Bottom-up	Randomized 50 nt	2 Novel IRESs	Unknown	*H. sapiens*	Venkatesan *et al.,* 2001
Novel	Bottom-up	Randomized 18 nt	56 Novel IRESs	rRNA Base pairing	*S. cerevisiae*	Zhou *et al.,* 2003
EMCV	Top-down	Random Mutagenesis	12 IRES Variants	N/A	*H. sapiens*	Li *et al.,* 2007
EMCV	Top-down	Directed Mutagenesis	24 IRES Variants	N/A	*H. sapiens*	Koh *et al.,* 2013
FMDV	Top-down	Computational Design	Thermo-IRESs	Temperature	*O. cuniculus*	Garcia-Martin *et al.,* 2016
IGR	Top-down	Rational Design	Ligand Responsive IRES	Ligand Binding	*T. aestivum*	Ogawa, 2011
IGR	Top-down	Rational Design	Ligand Responsive IRES	Ligand Binding	*T. aestivum*	Ogawa *et al.,* 2017

High-throughput methods are able to identify large numbers of possible IRES sequences in a short amount of time. However, randomized sequences must be relatively short in order to achieve sufficient library complexity, therefore limiting the discovery of longer IRESs. The shortest natural IRESs discovered to date are the IGR IRESs at ~190 nt [115]. Considering the evolutionary pressure to reduce genome size, it is possible that this is the lower limit for structure-based initiation. While increased gene synthesis capabilities may breathe new life into this method in the future, it is possible that below a certain RNA length, the IRESs selected for are biased towards those that base pair to ribosomal RNA.

3.6.2 Top-down approaches: reengineering natural IRESs

The engineering of naturally-occurring viral IRESs holds a unique potential in terms of providing tools for synthetic biologists, while also uncovering the mechanistic details of IRES-mediated translation. To date, the only reported IRESs that have been engineered are viral IRESs, likely due to their relative simplicity and the wealth of structural and mechanistic knowledge compared to their cellular counterparts [115]. In 2007, Li *et al.* utilized several previously-characterized and several novel EMCV IRESs variants generated through error-prone PCR to optimize the production of antibodies [159]. This technique allowed the authors to test a twenty-fold range of light chain to heavy chain ratios while simultaneously creating a number of IRES variants with distinct translational activities [159]. Additionally, in 2013, Koh *et al.* utilized another EMCV IRES library to optimize antibody production [160]. Single mutations and subsequent combinations of these allowed the authors to generate 24 unique IRES sequences with a 300-fold range in translational efficiencies [160]. The foot-and-mouth disease virus (FMDV) IRES was recently used in combination with RNA thermometers (cis regulatory elements that change structure based on the temperature) to rationally design temperature-sensitive IRESs (thermo-IRESs) [161]. The authors replaced a portion of the FMDV IRES with computationally-designed RNA thermometers, generating IRES that were more active at higher temperatures [161]. Finally, work from Ogawa and colleagues had demonstrated the IGR IRESs ability to be utilized in combination with RNA aptamers as a detection tool [162]. Originally, Ogawa showed that through careful placing of an RNA aptamer upstream of the IGR IRES, its translation could be regulated using a small molecule ligand [163]. Later, Ogawa *et al.* selected for a riboswitch aptamer inside the IGR IRES [162]. By placing a split theophylline aptamer in the middle of a key structural component (similar to the thermo-IRESs), the authors could use a ligand to control IGR

IRES-mediated translation [162]. Finally, Zhao *et al*. combined a toehold switch and the IGR IRES creating an RNA-responsive element for eukaryotic translational control [164].

3.7 Outlook

The ability to universally initiate translation is very intriguing from a synthetic biology and biotechnology point of view. Although only a handful of publications report successful IRES engineering projects, the non-peer reviewed literature reveals that several IRES-library projects were funded at the national and institutional level but remain either unpublished or are only available in the form of research theses. Thompson *et al.* gives a hint as to why so few IRES engineering projects have come to fruition considering the relatively large amount of bioengineering in yeast "We have expressed approximately two million dicistronic mRNA species in yeast, containing unique nucleotide sequences in the intercistronic spacer. However, none of those sequences functioned as an IRES to mediate translation of the second cistron (unpublished observation)." [165]. Interestingly, even IGR IRESs which require no additional factors function poorly in yeast, and only after hindering the canonical translation machinery can the IGR IRES function efficiently [165]. It may be the case that a predictable and multi-organism translation initiation signal may never be possible as the different expression systems have diverged too far over billions of years of evolution.

We believe that it is critical to end this review with a brief discussion around the 'umbrella' term IRES. Currently, in eukaryotes, the term IRES covers many non-canonical exceptions to the translation initiation mechanism and therefore has become somewhat synonymous with 'non-canonical'. This is due to the vague nature of the term, as technically all mechanisms which do not begin at the 5' end of the mRNA are IRES mechanisms. In the prokaryotic context, the term IRES

applies to practically all the described methods of translation initiation due to the absence of the 5' cap. Currently, IRESs are classified based on their mechanisms of translation and their origin (e.g. viral); however, with mounting evidence that cIRESs and vIRESs may use similar mechanisms, the boundaries between mechanistic classes is becoming increasingly blurred [145]. It may therefore be beneficial for bioengineers to adopt a generalizable TIR nomenclature based on translation initiation mechanism ((1) RNA structure, (2) rRNA base pairing, (3) specialized ribosomes, and (4) additional factor requirement) similar to what has been done for vIRESs. As this field continues to grow and evolve, we believe this will provide additional clarity (with respect to IRES translation initiation mechanism) as well as help pave the way for future rational engineering of IRES-based devices.

3.8 Acknowledgements

We thank Emily Wilton and Taylor Sheahan for critical reading of the manuscript and Justin Vigar for helpful discussions.

Chapter 4.

The prokaryotic activity of the IGR IRESs is mediated by ribosomal protein S1

Reprinted with permission from “The prokaryotic activity of the IGR IRESs is mediated by ribosomal protein S1”

Roberts L., Wieden H.-J. (2022). The prokaryotic activity of the IGR IRESs is mediated by ribosomal protein S1. Nucleic Acids Research, Volume 50, Issue 16, Pages 9355-9367.

4.1 Foreword

This chapter is presented in the manuscript form that was published in the journal Nucleic Acids Research (NAR) (Roberts & Wieden, 2022). As first author, I conceptualized and performed the experiments, I analyzed the results of the experiments, and wrote and revised the manuscript. Hans-Joachim Wieden aided in the structuring of ideas, the conceptualization of the experiments and the writing of the manuscript. The style of citation has been updated from the original publication to maintain continuity in this thesis for the ease of the reader.

4.2 Abstract

Internal ribosome entry sites (IRESs) are RNA elements capable of initiating translation on an internal portion of a messenger RNA. The intergenic region (IGR) IRES of the *Dicistroviridae* virus family folds into a triple pseudoknot tertiary structure, allowing it to recruit the ribosome and initiate translation in a structure-dependent manner. This IRES has also been reported to drive translation in *Escherichia coli* and, to date, is the only described translation initiation signal that functions across domains of life. Here, we show that unlike in the eukaryotic context, the tertiary structure of the IGR IRES is not required for prokaryotic ribosome recruitment. In *E. coli*, IGR IRES translation efficiency is dependent on ribosomal protein S1 in conjunction with an AG-rich Shine-Dalgarno-like element, supporting a model where the translational activity of the IGR IRESs is due to S1-mediated canonical prokaryotic translation.

4.3 Introduction

Translation is one of the most fundamental processes occurring in every living cell [7, 27, 109]. Consistent with its central role, translation is structurally and functionally conserved across all domains of life and can be divided into four main phases: initiation, elongation, termination,

and recycling [7]. Canonical eukaryotic translation initiation relies on the presence of the 7-methylguanylate cap at the 5′ end of messenger RNA (mRNA), which is recognized by initiation factors, that recruit the ribosome to initiate translation of the encoded peptide from the capped 5′ end of the mRNA [109]. While this is the predominant form of eukaryotic translation initiation, numerous non-canonical mechanisms exist [166, 167]. Among these, the internal ribosome entry sites (IRESs) are RNA elements present in a respective mRNA and that are capable of recruiting ribosomes. Through IRESs, translation initiation can begin on an internal portion of the mRNA [111]. While all IRESs function independently of the 5' cap, they often require additional protein factors for their function [7]. The intergenic region (IGR) IRES of the *Dicistroviridae* virus family is unique as it does not require any initiation factors and initiates translation on a non-AUG start codon [132, 168, 169]. This non-canonical activity relies on a triple pseudoknot (PK) tertiary structure, which recruits the ribosomal subunits and mimics a canonical tRNA-mRNA duplex [131, 133, 135, 137, 138, 170-172]. The IGR IRESs are phylogenetically classified into three subtypes based on the number of stem loops and the length of the loop regions [111, 173]. Subtypes I and II are well studied and similar in structure with subtype II having an extra hairpin in PKI, while subtype III is more recently discovered and significantly smaller than the other two subtypes [7, 111, 174-176]. Interestingly, and in tone with the overall evolutionary conservation of the translation machinery, it has been reported that the IGR IRES from the *Plautia stali* intestine virus is able to also initiate translation in *Escherichia coli* [134]. However, the reported prokaryotic IRES activity differs from that in the eukaryotic context. In *E. coli*, the AUG start codon is essential for IRES-mediated translation, additionally disrupting certain PK structures (demonstrated to be essential for translation in eukaryotes) has only a minimal effect on translation efficiency in *E. coli* [134]. As a consequence, the proposed mechanism for prokaryotic IRES activity is a hybrid of the

previously described eukaryotic IRES activity and the canonical bacterial Shine-Dalgarno (SD) initiation mechanisms, whereby bacterial ribosomes transiently interact (in a structure-dependent manner) with the IRES before repositioning to a downstream SD-like sequence [134, 177]. This suggests some structured mRNAs can indeed be specifically translated in bacteria by exploiting the evolutionary conserved structural features of the ribosomal core. Besides being interesting from an evolutionary perspective, this is also of great interest and utility for bioengineers and synthetic biologists alike [134, 150]. Interested in characterizing the molecular mechanism underpinning this phenomenon, we performed an in-depth characterization of prokaryotic IGR IRES ribosome recruitment (*in vitro*) and translation (*in vivo* and *in vitro*). Surprisingly, our results demonstrate that the tertiary structure of the IGR IRES is dispensable for prokaryotic ribosome recruitment and translation initiation in *E. coli*. Instead, we find that IGR IRES translation efficiency correlates positively with single-stranded AG-rich regions, is dependent on ribosomal protein S1, and requires a SD-like sequence upstream of the start codon. This challenges the previously proposed hybrid model in support of the canonical model of prokaryotic translation initiation for these IRESs where the bacterial translation machinery is not inherently biased to translate specific structured mRNAs, but rather seems to allow for translation of any mRNA for which the structure can be resolved by ribosomal protein S1 and where a SD-like sequence upstream of a start codon can be identified.

4.4. Materials and Methods

4.4.1 Fluorescent reporter construct design

IRES reporter constructs were designed to adhere to BioBrick engineering standards [178]. BioBrick Prefix and Suffix sequences (RFC 10) flank each construct for ease of cloning into

BioBrick vectors with standardized copy numbers. A T7 promoter (BBa_I719005) drives the transcription of optimized superfolder green fluorescent protein (sfGFP) coding sequence translationally controlled by an RBS or IRES sequences. Transcription is stopped by a transcriptional terminator (BBa_0015) downstream of the sfGFP coding sequence. Sequences for the strong (BBa_B0034), medium (BBa_B0032), and weak (BBa_B0033) RBSs were taken from the BioBrick part registry, and the 'dead' RBS is the reverse complement of BBa_B0034. Sequences for CrPV (AF218039), IAPV (NC_009025.1), and PSIV (AB006531) IGR IRESs with 18 nt of corresponding downstream coding sequences were obtained from GenBank (Summarized in Supplementary Table S4.1). We used the Salis Lab RBS calculator to ensure a strong RBS or an upstream start codon was not accidentally created during IRES mutagenesis and IRES scrambling [152].

4.4.2 Cloning and site-directed mutagenesis

Fluorescent reporter constructs were synthesized (Integrated DNA Technologies and Twist Biosciences) and subcloned into pSB3C5, a medium to low copy number plasmid. Pseudoknot (PK) mutations and deletions were introduced using the Quickchange™ method. All reactions were carried out using a TGradient (Biometra) thermocycler and resulting mutant plasmids transformed into electro competent BL21-Gold (DE3) cells (Agilent). The integrity of all constructs and PK mutations were confirmed by sequencing (Genewiz).

4.4.3 Cell growth

50 mL of *E. coli* BL21-Gold (DE3) cells containing fluorescent constructs were grown in LB media to mid log phase (0.5 OD_{600nm}) at 37 °C with shaking (200 rpm) in 125 mL Erlenmeyer flasks and expression induced with isopropylthio-β-galactoside (IPTG, 1 mM final concentration).

Cells were then harvested at distinct time intervals (fluorescent time courses) or grown for three hours before being analyzed by flow cytometry.

4.4.4 Flow cytometry

Cells were pelleted, washed twice with, and subsequently resuspended in, FACSFlow™ (BD Biosciences), and kept on ice until cytometric analysis. Flow cytometry was performed on a BD FACSAria Fusion cell sorter (488 nm excitation, observing sfGFP fluorescence in the FITC channel) and data analysis performed on Flowjo software (Flowjo, LLC). All flow cytometry was performed in biological triplicate, collecting 100,000 events per replicate.

4.4.5 sfGFP immunoblotting

Whole-cell lysate or 5 μL of PURExpress® (New England BioLabs) reaction was loaded onto a nitrocellulose membrane (Pall Corporation) using a Biodot SF microfiltration apparatus (BioRad) and the presence of sfGFP was detected using an anti-GFP antibody (Abcam, ab6556) and a peroxidase conjugated secondary antibody (Sigma, A0545). Chemiluminescence from three biological replicates was quantified using an Amersham Imager 600 (GE healthcare).

As a secondary check regarding the effect the maturation time of the sfGFP has on signal generation, we used the GFP specific antibody to probe protein levels during the expression time course, which confirmed that our live cell fluorescence assay was accurately reporting protein levels and not variations in the sfGFP maturation times (Supplementary Figure S4.1).

4.4.6 RT-qPCR

Total RNA from three biological replicates was extracted from *E. coli* using an EZ-10 total RNA purification kit (Bio Basic) and the integrity/purity confirmed using formaldehyde agarose

gel electrophoresis and A_{260}/A_{280} ratio (Biodrop). Using 100 ng of total RNA and the respective reverse primers (IDT) (sfGFP 5′-GATAACGAGCAAAGCACTGAAC-3′ and cysG 5′-ATGCGGTGAACTGTGGAATAAACG-3′), cDNA was generated using qScript cDNA Supermix (Quanta Biosciences) according to manufacturer's specifications. Quantitative PCR was performed according to manufacturer's specifications on a StepOnePlus Real-Time PCR System (Thermo Fisher) using PerfeCTa® SYBR® Green SuperMix (Quantabio) with the corresponding forward primers (IDT) (sfGFP 5′-GGTGACGCAACTAATGGTAAAC-3′ and cysG 5′-TTGTCGGCGGTGGTGATGTC-3′) and the above reverse primers. All sfGFP mRNA threshold values were scaled relative to the accompanying cysG reference mRNA threshold values to account for differences in cDNA input.

4.4.7 sfGFP degradation assay

50 mL of *E. coli* BL21-Gold (DE3) cell containing fluorescent constructs were grown in LB media to mid log phase (0.6 OD_{600nm}) at 37 °C with shaking (200 rpm) in 125 mL Erlenmeyer flasks in the presence of IPTG (1 mM final concentration). 50 mL of cells were pelleted and washed twice in AB minimal media (without IPTG) before being resuspended in 50 mL of AB minimal media and incubated at 37 °C with shaking (200 rpm) in 125 mL Erlenmeyer flasks [179]. Cells were then harvested at specific time intervals and sfGFP measured by flow cytometry (as above). The OD_{600nm} was constant at ~0.6 over the course of the experiment, ensuring cells were not actively dividing. No variations in sfGFP levels or degradation rates (Supplementary Figure S4.2) could be detected for the different constructs; in particular, the decay rate was so slow (sfGFP was stable over numerous days) that sfGFP degradation is negligible over the time of our experiments.

4.4.8 RNA in vitro transcription, [^{32}P] labelling, and purification

DNA templates for *in vitro* transcription were generated by PCR using plasmids containing wild type IRESs, mutant IRESs, or control RNAs (Supplementary Tables S4.1 and S4.2). The obtained DNA was used in subsequent *in vitro* transcription reactions, and the resultant RNA purified by nucleic acid spin column (Bio Basic). The purity and homogeneity of the RNA was assessed by urea PAGE and A_{260}/A_{280} ratio (BioDrop µlite, BioDrop).

Five hundred nanograms of IRES RNA in water was unfolded by heating to 95 °C for 2 min before being snap cooled on ice. RNA was then dephosphorylated by incubating at 37 °C with Shrimp Alkaline Phosphatase (0.001 U/µL final concentration, Fermentas) for 60 min. Two hundred and fifty nanograms of dephosphorylated RNA were incubated with T4 polynucleotide kinase (0.5 U/µL final concentration, Fermentas) and 1.5 µL of [^{32}P]-γ-ATP (30 µL total reaction volume) for 60 min at 37 °C. To quench the reaction, 1.5 µL of 0.5 M EDTA pH 8.0 was added and the reaction subsequently heated to 75 °C for 10 min before the RNA was purified via EZ-10 Spin Column RNA Cleanup and Concentration Kit (Bio Basic).

4.4.9 Purification of prokaryotic and eukaryotic ribosomes

Prokaryotic 70S ribosomes and 30S ribosomal subunits were purified from *E. coli* MRE600 as per Becker *et al.* [84, 180]. Eukaryotic 40S ribosomal subunits were purified from HeLa cells (National Cell Culture Laboratory) as previously described [145].

4.4.10 Removal of ribosomal protein S1 from 30S subunits (30S^{-S1} subunits)

30S ribosomal subunits were diluted tenfold in a high-salt dissociation buffer (20 mM Tris-HCl pH 7.5, 10 mM $MgCl_2$, 60 mM KCl, 1 M NH_4Cl, and 1 mM DTT). The mixture was incubated

at 37 °C for 10 min before being added to poly(U) (Sigma Aldrich, P8563) and incubated at 4 °C for 1 hr with inversion. The mixture was centrifuged at 500 × g for 5 min, and the supernatant collected. The $30S^{-S1}$ subunits were pelleted via ultracentrifugation with a Sorvall S55-S swinging-bucket rotor ultracentrifuge (Thermo Scientific) at 55 000 rpm, at 4 °C for 24 h and resuspended in TAKM5 (50 mM Tris-HCl pH 7.6, 70 mM NH_4Cl, 30 mM KCl, 5 mM $MgCl_2$) to a concentration of ~15 μM. S1 removal was confirmed via SDS-PAGE and mass spectrometry (U of L Mass Spectrometry Facility).

4.4.11 Nitrocellulose filtration assays

Radio-labeled RNA (50 nM final concentration) in TAKM5 buffer was heated to 95 °C for 10 min and slow cooled to room temperature. RNA was then incubated with increasing amounts of ribosomal subunits/ribosomes for 15 min at 37 °C before being rapidly filtered through a cellulose nitrate membrane filter (0.2 μm, GE Healthcare). The cellulose nitrate membranes were washed with 1 mL of cold TAKM5 buffer and placed into 10 mL of EcoLite (+) scintillation cocktail (MP Bio), vortexed for 30 s, and subsequently incubated at room temperature for 30 min, followed by vigorous mixing for 30 s. The retained radioactivity was quantified by scintillation counting (Tri-carb 2810 TR LSA, Perkin Elmer). To ensure our system was able to replicate previously reported data, we also performed binding assays with HeLa 40S subunits. Consistent with previous work, the WT CrPV IRES bound the 40S with a K_D of ~14 nM, while disruption of PK1 had no effect on 40S binding, and disruption PK1 and PK3 in combination abolished 40S binding (Supplementary Table S4.3) [132, 181].

4.4.12 RNA fluorescence titrations

Purified WT CrPV IGR IRES RNA was labelled at the 3′ end with pyrene as per Keffer-Wilkes *et al.* (65). Labeled RNA (50 nM final) in TAKM5 was heated to 95 °C for 10 min and slow cooled to room temperature. RNA was then incubated with increasing amounts of ribosomal subunits/ribosomal protein S1 before being excited at 341 nm. Ribosomal protein S1 was a generous gift from J.L.E. Heller and J.R.J. Vigar and was purified as previously described [182, 183]. The peak fluorescence at 391 nm was recorded and plotted as a function of increasing 30S or S1 concentration.

4.4.13 Circular dichroism (CD) spectroscopy

Circular dichroism was performed on a Jasco J-815 CD Spectrometer using a 1 s integration time over 200-320 nm [184]. WT CrPV IRES RNA (~150 nM) was folded as described previously (see nitrocellulose filtration assay) before being subjected to CD spectroscopy. Next, increasing amounts of ribosomal protein S1 was added to the IRES RNA, incubated at 37 °C for 15 min, and again subjected to CD spectroscopy. Ribosomal protein S1 alone was measured at each concentration and that data was subtracted from the RNA/protein mixture to ensure only RNA signal was being observed. All data were recorded 5 times; average traces are plotted.

4.4.14 Statistical information

For all statistical analyses, n = 3 unless otherwise stated.

A) Flow cytometry

Flow cytometry was performed as described in methods description. Mean fluorescence was calculated using FlowJo software (FlowJo, LLC). Standard deviation and relative significance (T-Test, two tailed) of each data set was calculated using Microsoft Excel.

B) Immunoblot

Immunoblot intensity was determined using the ImageJ gel analysis package [185]. Mean intensity and standard deviation were calculated using Microsoft Excel.

C) RT-qPCR

Threshold levels were generated by the StepOnePlus Real-Time PCR System and mRNA levels were calculated using the values generated from the StepOnePlus software (Applied Biosystems) and Microsoft Excel.

4.5 Results

To quantify translation efficiency (TE) in live *E. coli*, we utilized a fluorescence-based reporter (superfolder green fluorescent protein (sfGFP)) assay for monitoring live cell fluorescence via flow cytometry. We opted to measure individual live cell fluorescence by flow cytometry to avoid potential averaging of distinct *E. coli* populations. To benchmark our reporter system, we selected three ribosome binding sites (RBS) (Strong B0034, Medium B0032, and Weak B0033) from the registry of standard biological parts (http://parts.igem.org) [186], as well as a 'dead' RBS (the reverse complement of B0034), to drive the expression of sfGFP (Supplementary Figure S4.3). Flow cytometry measurements correlate nearly perfectly ($R^2 = 0.99$) with the expression strength predicted using the ribosome binding site calculator, demonstrating the sensitivity and wide range over which our single-cell assay can accurately report translation efficiency *in vivo* without interference by endogenous mRNA expression (Figure 4.1A) [152, 187]. Using this assay and by benchmarking different IRESs against well-characterized RBSs, we are able to accurately measure their translation efficiency, allowing for the first time a direct comparison of IRES

translation to the canonical system and to assess their ability to compete with endogenous mRNAs *in vivo* (Figure 4.1B).

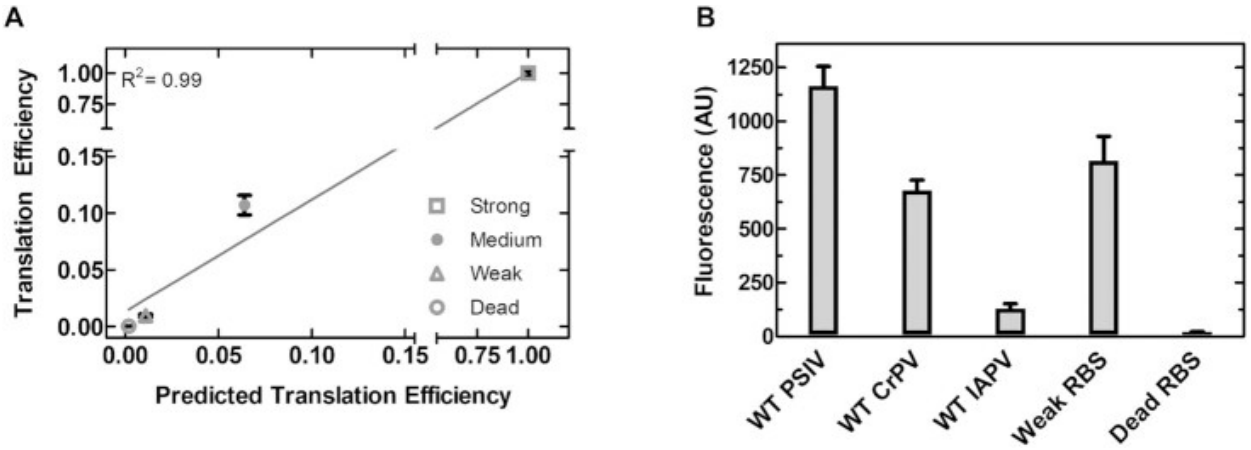

Figure 4.1. IGR IRES translation is weak compared to canonical ribosome binding sites. (A) Correlation of predicted and measured translation efficiency of standardized RBSs. Translation efficiency predicted using the Salis lab RBS calculator [152], translation efficiency measured by flow cytometry. Mean values of three biological replicates are plotted. (B) Translation efficiencies of the WT IGR IRESs compared to standardized RBSs. Translation efficiency measured by flow cytometry, mean values of three biological replicates are plotted and error bars indicate one standard deviation.

4.5.1 The IGR IRESs have translational efficiencies comparable to a weak RBS in vivo

To ensure accurate representation of the IGR IRESs, we selected two type I (*Plautia stali* intestine virus, PSIV and Cricket Paralysis Virus, CrPV) and one type II (Israeli Acute Paralysis Virus, IAPV) IGR IRESs. To accurately mimic the IRES expression, we included 18 nucleotides of viral coding sequence downstream of the IRES to keep the initiation element in its native context and our reporter system (Figure 4.2A and B) consistent with previous studies [132, 134]. Finally, we opted for a monocistronic IRES construct to avoid potential translational coupling as

downstream translation and the intergenic RNA structure can be influenced by upstream translation [188, 189]. The obtained live-cell fluorescence data revealed that the translation efficiency of the type I IRES constructs (CrPV and PSIV) are comparable to the weak RBS while the type II (IAPV) is roughly an order of magnitude lower (Figure 4.1B). Interestingly, the same trend is observed for these IRESs in yeast and rabbit reticulocyte lysates as the type I IRESs are more translationally efficient than type II IRESs [190, 191]. Deletion of the sfGFP start codon abolished translation in all IRES constructs, making them indistinguishable from cells with no fluorescent reporter (Figure 4.2C), and is consistent with previous work [134].

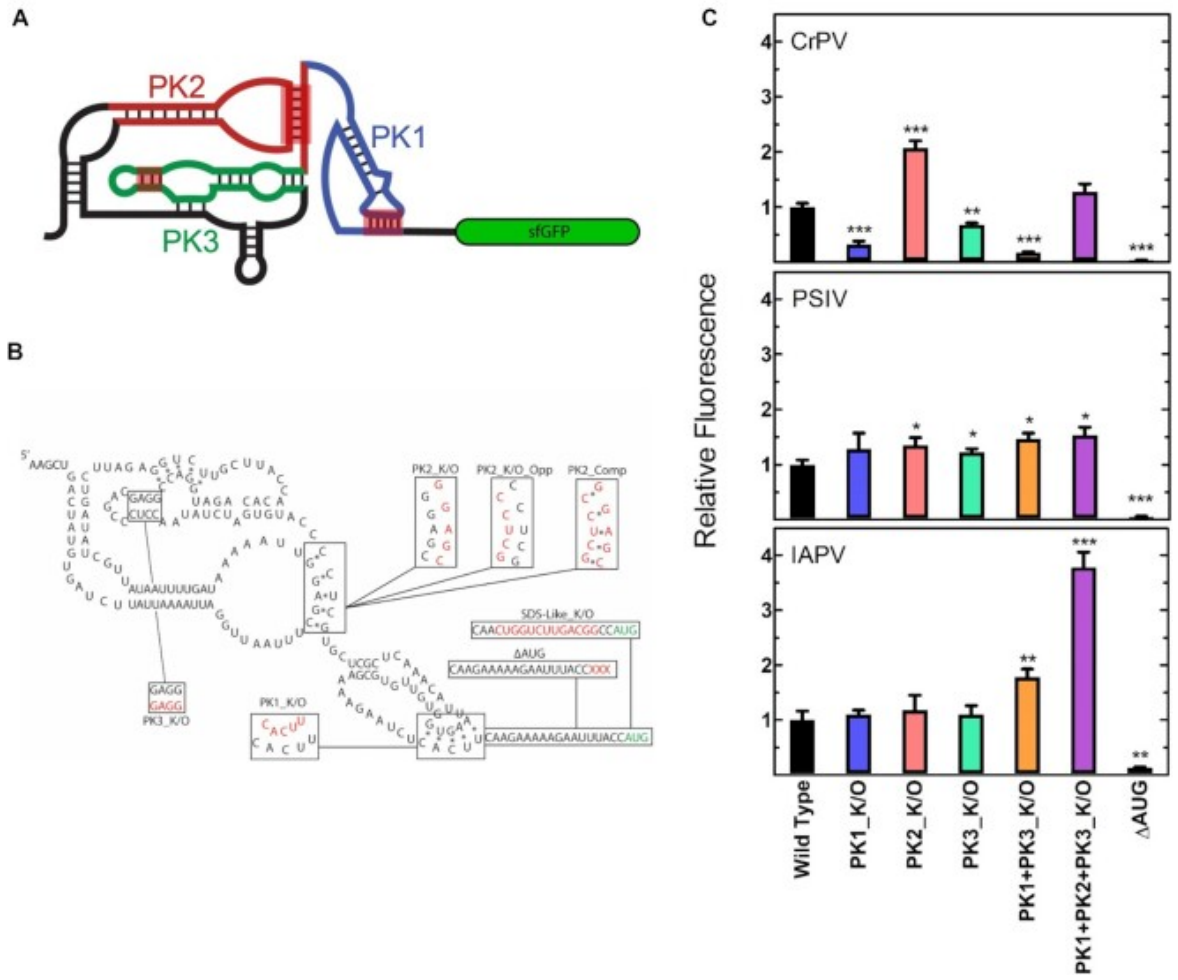

Figure 4.2. IGR IRES translation efficiency is independent of pseudoknot structure. (A) Cartoon representation of the monocistronic fluorescent reporter construct including the secondary structure of a generic IGR IRES, pseudoknots (PK) 1 (blue), 2 (red), and 3 (green) are indicated. (B) Secondary structure of the PSIV IRES with nucleotide substitutions for annotated mutations shown in red and the sfGFP start codon in green. (C) Translation efficiency of WT and IGR IRES variants measured by flow cytometry, mean values of three biological replicates are plotted relative to the respective WT IRES. For all panels, error bars indicate one standard deviation. Constructs with statistically significant differences from WT are indicated (*$P < 0.05$, **$P < 0.01$, ***$P < 0.001$, ****$P < 0.0001$).

4.5.2 Disruption of pseudoknots does not perturb type I or type II IGR IRESs activity in vivo

Although consistent with canonical bacterial translation, the observation that an AUG start-codon is required for translation of the IRES-containing mRNAs raises the question regarding the role of the IRES structure in recruiting the ribosome and ultimately translation initiation. To investigate the detailed mechanistic role that the structured elements of the IRES play in its translational efficiency, we systematically disrupted conserved structural features (e.g., PKs, Figure 4.2B) present in the IGR IRESs (Summarized in Table 4.1) and measured the corresponding expression levels of sfGFP using flow cytometry. The PKs are critical for ribosome recruitment and translation initiation in eukaryotes. It has been previously demonstrated that altering the sequence of the IRES to disrupt the Watson-Crick-Franklin base pairs affects only the respective PK structure, leaving the rest of the IRES structures intact [132]. Surprisingly, our measurements reveal (Figure 4.2C) that disruption of PK1, PK2, PK3, and combinations of these mutations have no negative effects on IAPV or PSIV IGR IRES mediated translation. In fact, several of the constructs exhibit increased translation when compared to their wild type (WT) counterparts (Figure 4.2C). However, this is in contrast to previous work where disruption of PK2 reduced PSIV mediated translation ~60% [134]. In this context, the CrPV IRES is unique as disrupting PK1 and PK3 decreases the translation efficiency by 70% and 30% respectively. When combined (PK1+PK3_K/O), the translation efficiency decreases by 80% (Figure 4.2C). Unexpectedly, disruption of PK2 increased translation efficiency (~100% relative to WT), and when combined with the other PK mutations (PK1+PK2+PK3_K/O), had a seemingly compensatory effect restoring activity to roughly WT level (Figure 4.2C).

Table 4.1. Relative *in vivo* translation efficiencies of IGR IRES constructs measured by flow cytometry. Mean values of three biological replicates are shown; error indicates one standard deviation.

	PSIV	IAPV	CrPV
Wild Type	1.00 ± 0.08	1.00 ± 0.17	1.00 ± 0.07
PK1_K/O	1.28 ± 0.29	1.10 ± 0.08	0.32 ± 0.06
PK2_K/O	1.35 ± 0.13	1.18 ± 0.03	2.07 ± 0.13
PK3_K/O	1.22 ± 0.06	1.10 ± 0.16	0.67 ± 0.04
PK1+3_K/O	1.47 ± 0.11	1.78 ± 0.15	0.18 ± 0.01
PK1+2+3_K/O	1.54 ± 0.14	3.78 ± 0.28	1.29 ± 0.14
SD-like_K/O	0.04 ± 0.01	-	-
sfGFP ΔAUG	0.06 ±0.01	0.14 ± 0.01	0.04 ± 0.01

The observation that disrupting pseudoknot structures often has no effect on IRES translation efficiency suggests that its tertiary structure is not required for the observed cross-kingdom expression activity of the IGR IRES. Complete deletion of PK elements from CrPV further supports this as deletion of the highly structured PK2 and PK3 (ΔPK2/3) results in a 300% increase in translation efficiency, while deletion of the relatively less structured PK1 (ΔPK1) abolishes translation (Supplementary Figure S4.4A). While these results do not align with a structure-based initiation mechanism, they can be interpreted through the lens of SD mediated translation of a highly structured 5′ untranslated region (UTR). In this context, removing a highly structured element (PK2/3) increases translation efficiency and placing this structured element near the start codon (by deleting PK1) hinders translation efficiency. To determine if the structure

of the IGR IRESs offers any advantage over similar mRNAs with no such tertiary structure, we randomized the CrPV and IAPV sequences (while maintaining nucleotide composition) and assessed their translational efficiency using our *in vivo* fluorescence assay. Interestingly, the randomized CrPV and IAPV translation efficiencies are ~200% and ~600% greater than their WT counterparts (Supplementary Figure S4.4 B and C), bringing CrPV translation efficiency to 200% of that of the weak RBS and elevating IAPV to the level of the weak RBS.

While the translational efficiencies of the IGR IRES PK variants do not agree with a structure-based interpretation, they correlate very well with the predicted translation efficiencies calculated using the Salis lab RBS calculator for bacterial mRNA expression strengths (PSIV R^2 = 0.98, CrPV R^2 = 0.76, and IAPV R^2 = 0.89) (Figure S4.5), suggesting canonical SD-based translation as the mechanism responsible for the observed expression of IRES containing mRNAs in bacteria. Together, these results indicate that the tertiary structure of the IRES is not responsible for, but rather is inhibitory to, efficient translation. Therefore, we hypothesize that the IGR IRESs are being treated as large structured 5′ UTRs and translated via the canonical processes of the translation machinery that deal with the expression of structured mRNA rather than through molecular mimicry exploiting the conserved structural core of the ribosome (as is the case in eukaryotes).

4.5.3 IRES translational efficiency is independent of bacterial growth phase

In our initial experiments, we measured fluorescence at a single time point three hours post induction, shortly after entry into the stationary growth phase. However, it is possible that during times of increased competition for ribosomes (rapid growth), the structure of the IRES provides a kinetic advantage to the mRNA by transiently interacting with ribosomes in a structure-dependent

manner and increasing the local ribosome concentration. In order to determine if IRES translation activity is indeed due to canonical translation and if this is affected by the growth phase, we performed a detailed characterization of global IRES driven gene expression from mRNA transcription to protein degradation (Supplementary Figure S4.6). For this analysis, we selected the WT PSIV and PK2_K/O PSIV IRES constructs as they have different translation efficiencies in our initial experiments and in previous work [134]. In agreement with our initial data, PK2_K/O reached a higher final fluorescence (Figure 4.3A) and the rate of sfGFP production during the exponential phase was 2-times greater for PK2_K/O (41.1 ± 1.4 min^{-1}) than for WT (19.9 ± 0.6 min^{-1}) (Figure 4.3B). This demonstrates that IRES translational efficiency is not affected by bacterial growth phase and again supports the notion that IRES tertiary structure is not responsible for translational activity.

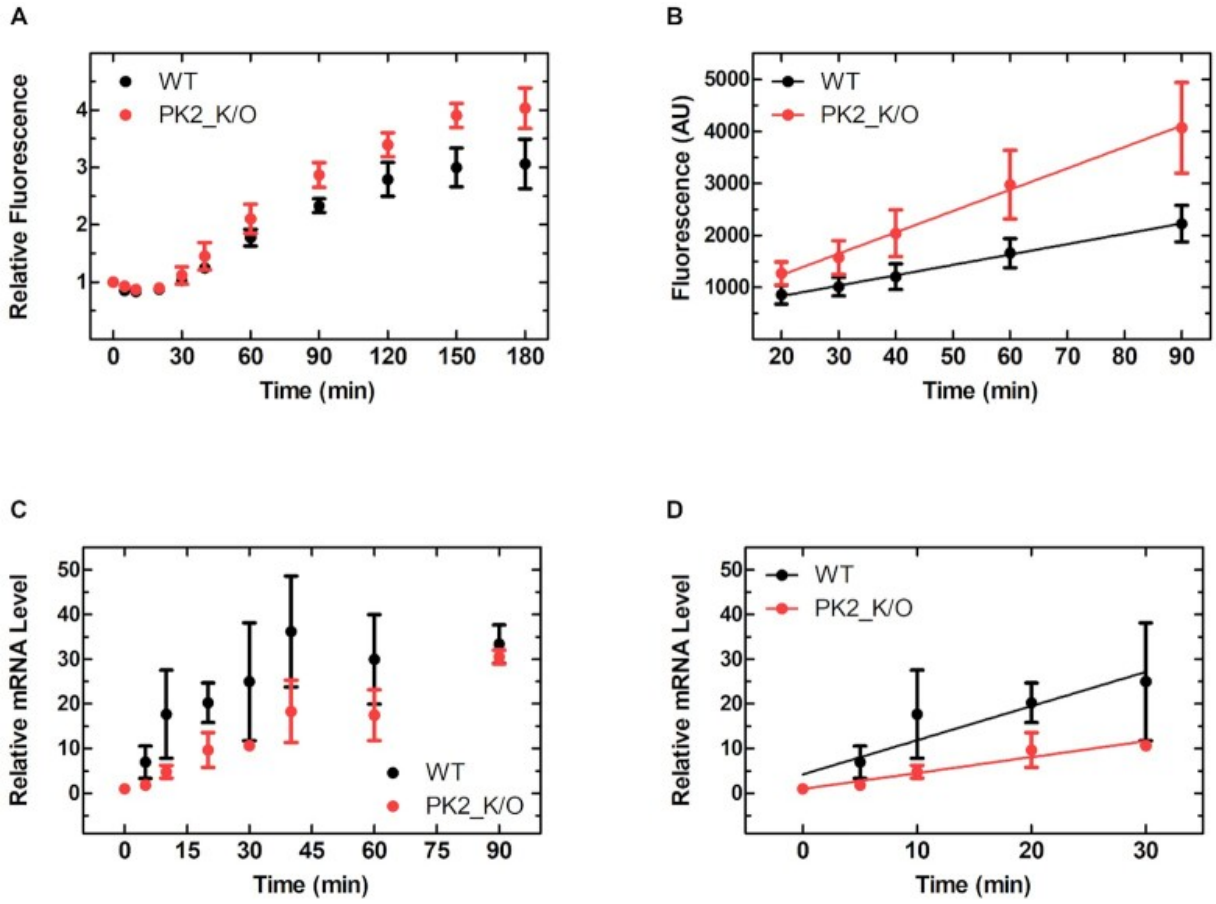

Figure 4.3. Differential PSIV IGR IRES translation efficiency is consistent over multiple growth phases and not an artifact of mRNA stability. (A) Relative fluorescence *in vivo* time course of *E. coli* containing PSIV IGR IRES constructs as measured by flow cytometry and (B) the linear portion of sfGFP expression in panel A. (C) Time course of relative mRNA level in live *E. coli* expression of PSIV IGR IRES constructs as measured by RT-qPCR. (D) Linear portion of mRNA expression from panel C. Mean values of three biological replicates are plotted; error bars indicate one standard deviation.

4.5.4 Disruption of IRES PK structure affects mRNA stability in vivo

While disruption of the PK structure has been shown previously to only affect local RNA structure [132], the PSIV PK2_KO mutation could be affecting the stability of the mRNA *in vivo*,

lowering the abundance of the respective mRNA and as a consequence alter the measured translational efficiency. To assess mRNA accumulation and stability, we measured the total mRNA levels via RT-qPCR on samples collected during the previous (*vide supra*) time course analysis. Interestingly, while mRNA levels were similar in the early stationary phase (Figure 4.3C), WT mRNA accumulated faster (~2×) than PK2_K/O mRNA, suggesting that disruption of PK2 (PK2_K/O) indeed decreases mRNA stability (Figure 4.3D). This result also suggests that PK2_K/O has an even higher relative translational efficiency than reported by our fluorescence data when factoring in the reduced abundance of the respective mRNA.

4.5.5 Location of the IRES element has no effect on translation initiation mechanism selection

In their natural context, the IGR IRESs initiate translation in the intergenic region of a bicistronic viral genome [7, 114]. While canonical Shine-Dalgarno based translation, according to this definition, could be considered 'IRES' translation, it is possible that the structure of the IRES will be more efficient at promoting initiation if located between two genes as opposed to the monocistronic reporter we utilized (*vide supra*). To test this, we designed a bicistronic reporter construct whereby an upstream monomeric red fluorescent protein (mRFP) is translated via the weak RBS (B0033, the most similar to the WT PSIV IRES expression level) and the PSIV IRES drives the translation of sfGFP downstream (Figure 4.4A). The general design of this dual reporter construct is often used to validate IRES activity in eukaryotes and allows calculation of the mRFP/sfGFP ratio to control for intrinsic cellular noise. Subsequent analysis using flow cytometry revealed that the relative mRFP/sfGFP expression levels of the WT PSIV and PK variants are identical to the monocistronic expression levels (Figure 4.4B). These results demonstrate that the location of the IRES (5′ or internal) does not bias against a structure-based initiation mechanism and reaffirms our findings that the structure of the IRES is not essential for translation.

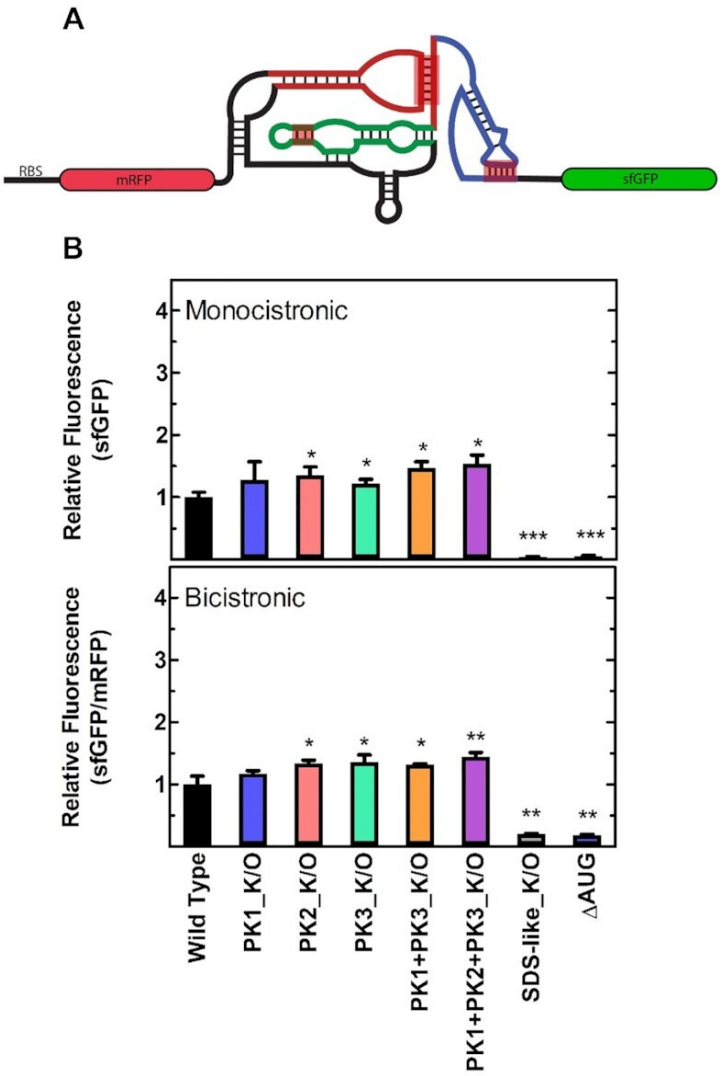

Figure 4.4. Relative PSIV IGR IRES translation efficiency is independent of its position in the mRNA. (A) Cartoon representation of the bicistronic fluorescent reporter construct including the secondary structure of a generic IGR IRES, pseudoknots (PK) 1 (blue), 2 (red), and 3 (green) are indicated. (B) Translation efficiency of the mono- and bicistronic PSIV IGR IRES constructs measured by flow cytometry, mean values of three biological replicates are plotted relative to the respective WT IRES. Error bars indicate one standard deviation. Constructs with statistically significant differences from WT are indicated (*$P < 0.05$, **$P < 0.01$, ***$P < 0.001$, ****$P < 0.0001$).

4.5.6 IRES translation efficiency correlates with single stranded AG-rich sequences

We noticed a trend in our data that PK variants with increased translational efficiency almost always contained substitutions that introduced new single stranded (or liberated previously base-paired) AG-rich sequences upstream of the start codon, which have been shown to promote initiation [192]. These single-stranded sequences could be acting as ribosome standby sites, S1 binding sites, or simply facilitate the breathing of neighbouring RNA structures [193, 194]. As a simple check, we introduced a compensatory mutation that re-establishes PK2 structure (PK2_Comp) in the PSIV and CrPV IRESs. This mutation returns PSIV translation efficiency to the wild type level and reduces the translation of the CrPV PK2_K/O IRES variant ~70% (Figure 4.2B and S4.5A and B). However, the compensatory mutation alone (PK2_Opp), while disrupting PK2 structure, does not introduce a single stranded AG-rich sequence (Figure 4.2B) and, interestingly, has little to no effect on translation efficiency (Supplementary Figure S4.7A and B). These further support (at least for these particular variants) the notion that not the disruption of the PK structure, but rather the presence of a single stranded AG-rich sequence upstream of the start codon is important for its increased translation efficiency.

4.5.7 Ribosomal protein S1 is required for efficient IRES translation

Our *in vivo* assays demonstrated that the structure of the IRES is non-essential and in fact even limits translational efficiency, suggesting that additional factors such as RNA helicases or chaperones (to resolve the structure of the IRES) might be required for their translation. To determine if cellular factors or the ribosome itself are responsible for this, we measured the rate of IRES-mediated sfGFP production *in vitro* using the highly purified and reconstituted PURExpress® system [195]. The respective sfGFP synthesis time courses mirror our *in vivo* data

as the rate of sfGFP production is 6-times greater for PK2_K/O (75.3 ± 25.3 min^{-1}) than for WT (11.5 ± 4.7 min^{-1}) (Figure 4.5A). This supports the idea that the component responsible for the observed effect is present in the recombinant, purified, and reconstituted transcription and translation system, and also backs our earlier observation that the *in vivo* translation efficiency of the PSIV IRES PK2_K/O might be limited by decreased *in vivo* stability of the respective mRNA (Figure 4.3C and D) as the contributing nucleases are not present in the PURExpress® system. Within the PURExpress® system, ribosomal protein S1 is the most likely candidate to resolve the IRES structure [196] as S1 binds to single stranded A-rich sequences and facilitates RNA unfolding essential for efficient canonical translation of mRNAs with structured 5′ UTRs in *E. coli* [194, 196, 197]. If the IRESs are indeed being treated as large structured 5′ UTRs by the bacterial translation machinery, their translational efficiency will be reliant on S1. However, if specific interactions of the triple PK structure of the IRES with the ribosome are responsible for its translational efficiency, the absence of S1 should have little to no effect. To probe this hypothesis, we monitored IRES translation efficiency using the PURExpress® Δribosome kit supplemented with either complete ribosomes (30S + 50S) or ribosomes lacking S1 ($30S^{-S1}$ + 50S). Both WT and PK2_K/O translation efficiencies were decreased ~90% when S1 was not present (Figure 4.5B), likely due to the ribosome no longer being able to efficiently bind and unwind the highly structured RNA. To ensure this is a S1 specific effect and not due to the treatment of the ribosomes during S1 removal, we supplemented stoichiometric amounts of recombinant S1 ($30S^{-S1}$ + 50S + S1), resulting in a 60% recovery in activity for both constructs (Figure 4.5B). Additionally, recombinantly-expressed ribosomal protein S1 binds the WT CrPV IGR IRES with an affinity roughly equivalent to the 30S ribosome (≤70 nM, Supplementary Figure S4.8), and is capable of unfolding the structure of the IRES in a concentration dependent manner (Supplementary Figure

S4.9). Together, this data demonstrates that S1 is indeed mediating the translation of the IGR IRESs by helping resolve their tertiary structure.

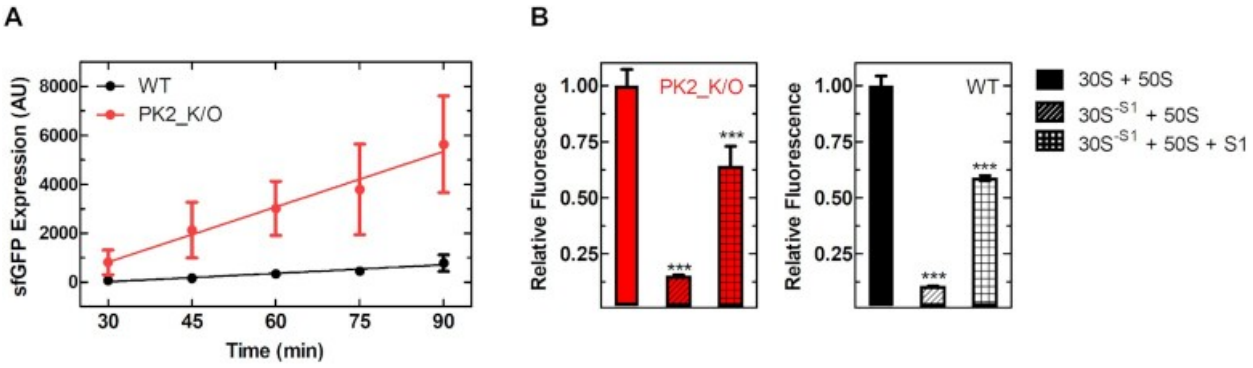

Figure 4.5. PSIV IGR IRES translation efficiency is strongly mediated by ribosomal protein S1. (A) Superfolder GFP production (*in vitro*) time course using the PSIV IGR IRES constructs and the PURExpress® system. (B) Relative fluorescence of PSIV IGR IRES constructs using the PURExpress® Δribosome system using complete ribosomes (30S + 50S), ribosomes lacking S1 ($30S^{-S1}$ + 50S), and ribosomes lacking S1 supplemented with recombinant S1 ($30S^{-S1}$ + 50S + S1). Mean values of three biological replicates are plotted, error bars indicate one standard deviation. Statistically significant differences between the WT and $30S^{-S1}$ samples and the $30S^{-S1}$ and $30S^{-S1}$ + S1 samples are indicated (*** $P < 0.001$).

4.5.8 IRES translation efficiency is dependent on a downstream SD-like sequence

If the IGR IRESs are being translated via the canonical translation initiation mechanism, then the 30S ribosome will require a SD-like sequence adjacent to the sfGFP start codon for efficient translation [192]. Interestingly, such a SD-like sequence is present in the viral coding region downstream of the PSIV IRES ([134] and Figure 4.2B). Altering this sequence to its reverse complement (SDS-like_K/O, Figure 4.2B) and leaving the IRES structure unchanged abolishes

translation in both the mono- and bicistronic reporters (Figure 4.4B) and further demonstrates that translation is proceeding via the canonical translation mechanism.

4.5.9 Pseudoknot mutations do not perturb IGR IRES binding to the prokaryotic ribosome

While IRES PK structure is not responsible for the observed translation activity in prokaryotes, it is still possible that the IGR IRES is able to transiently interact with prokaryotic ribosomes in a structure specific manner that may be overshadowed by the canonical translation activity. In such a model, disrupting the PKs structures will result in a reduced affinity of the IRES for the ribosome (as is the case for eukaryotic ribosome binding [132], Supplementary Table S4.3).

To investigate if the tertiary structure of the IGR IRES is responsible for prokaryotic ribosome binding, we determined the equilibrium binding constants (K_D) for the CrPV and IAPV IRES-ribosome complexes using nitrocellulose filter binding. WT CrPV and IAPV IRESs RNA binds to the 30S subunit and the 70S ribosome with comparable affinities of ~100 nM (Table 4.2). As expected, disrupting any of the PKs, individually or in combinations, has little to no effect on the affinity of either the CrPV or IAPV IGR IRESs for the 30S ribosomal subunit or the 70S ribosome (Table 4.2). This indicates that, in contrast to the eukaryotic system (Supplementary Table S4.3), the IGR IRESs bind to the prokaryotic ribosome independently of their tertiary structure [132]. This is consistent with the observation that prokaryotes do not have an equivalent to eS25, the eukaryotic ribosomal protein shown to be critical for ribosome binding and IRES activity [170, 198, 199]. To ensure that the use of nitrocellulose filtration to determine the affinity for the respective mRNAs is not biasing against a transient population of structurally bound IRESs, we also measured the affinity for the 30S ribosomal subunit using equilibrium fluorescence titrations with a pyrene-labeled WT CrPV IRES, which resulted in a nearly identical affinity of

~70 nM (Supplementary Figure S4.10). Finally, for comparison, we tested two native structured 5' UTRs (rpsO and sodB) and found that their affinities to the 30S ribosomal subunit (K_D = 10 ± 5 nM and K_D = 13 ± 2 nM, respectively) and the 70S ribosome (K_D = 94 ± 14 nM and K_D = 47 ± 13 nM, respectively) are on the same order of magnitude as the IRESs (Supplementary Tables S4.4 and S4.5).

Table 4.2. Dissociation constants (K_D) for CrPV and IAPV IGR IRES variants to 30S and 70S ribosomes as measured by nitrocellulose filter binding. Mean values of three biological replicates are shown; error indicates one standard deviation.

Construct	**Ribosome**	**CrPV**	**IAPV**
		K_D **(nM)**	K_D **(nM)**
WT	70S	69 ± 7	50 ± 9
PK1_K/O	70S	92 ± 9	67 ± 3
PK2_K/O	70S	100 ± 13	59 ± 11
PK3_K/O	70S	90 ± 15	59 ± 4
PK1+3_K/O	70S	78 ± 10	65 ± 4
PK1+2+3_K/O	70S	79 ± 10	72 ± 14
WT	30S	110 ± 23	93 ± 7
PK1_K/O	30S	55 ± 11	87 ± 14
PK2_K/O	30S	74 ± 13	81 ± 25
PK3_K/O	30S	48 ± 5	103 ± 42
PK1+3_K/O	30S	52 ± 12	80 ± 28
PK1+2+3_K/O	30S	48 ± 7	73 ± 14

4.6 Discussion

The ability of an RNA molecule to ubiquitously initiate translation in a structure dependent manner is an exciting prospect for a number of applications, including bioengineering and synthetic biology. While leaderless mRNAs have been described in all three domains of life [200], the PSIV IGR IRES would be the only RNA structure capable of initiating translation across domains and is currently the only example of structure-based initiation described in prokaryotes [134]. The proposed hybrid model of IRES mediated translation initiation in *E. coli* requires the IRES to first interact with the prokaryotic ribosome (in a structure-dependent manner) before it moves downstream to bind the RBS and initiate protein synthesis [134, 177]. This raises a number of interesting questions: do other examples of structure-based initiation exist in bacteria? How does the PSIV IGR IRES manipulate the bacterial ribosome? Is the ability to translate across domains of life an evolutionary-conserved feature of all the IGR IRESs or is it specific to PSIV? In order to better understand this phenomenon, we utilized a dual pronged approach to benchmark IGR IRES translation efficiency (*in vivo* and *in vitro*) and ribosome binding (*in vitro*). We opted to measure several IGR IRESs as it has been shown previously that the IGR IRESs from different viruses have unique translational efficiencies [165, 201]. Interestingly, our initial results demonstrated that disruption of the conserved PK elements does not consistently reduce IGR IRES translation efficiency and conversely often significantly increases translation efficiency. Therefore, results do not support the previously proposed model of IGR IRES mediated translation in *E. coli* and is at odds with previously published data where disruption of PK2 (and variants with a PK2 disruption) lead to a decrease in PSIV IGR IRES mediated translation [134]. We have ruled out the position of the IRES in the mRNA as the cause. However, it is possible that the reduced mRNA stability we observed for the PK2_K/O variant is more pronounced in a different mRNA

context (sfGFP vs Luciferase coding sequences), which could easily explain the lower translation efficiency given our observation that mRNA stability is sensitive to the presence of single stranded regions. Unfortunately, no mRNA stability information is provided in the previous report [134].

IRES variant translation efficiency data do correlate well with the mRNAs predicted free energy of folding ($R^2 = 0.61$, Supplementary Figure S4.11) using mFOLD [202] and the predicted translation efficiency using the Salis lab RBS calculator (Supplementary Figure S4.5), respectively [152]. This suggests that the IGR IRESs are being translated as structured mRNA through the canonical system rather than through a hybrid structure-based mechanism (Figure 4.6), highlighting the robustness of *E. coli*'s translation initiation machinery. This mechanism is supported by recent findings that ribosomes with altered anti-SD sequences (incapable of base-pairing to canonical SDs) are able to initiate at the correct codons, suggesting translation start sites are determined by inherent mRNA features such as upstream A-rich sequences and lower levels of surrounding mRNA structure [192, 203].

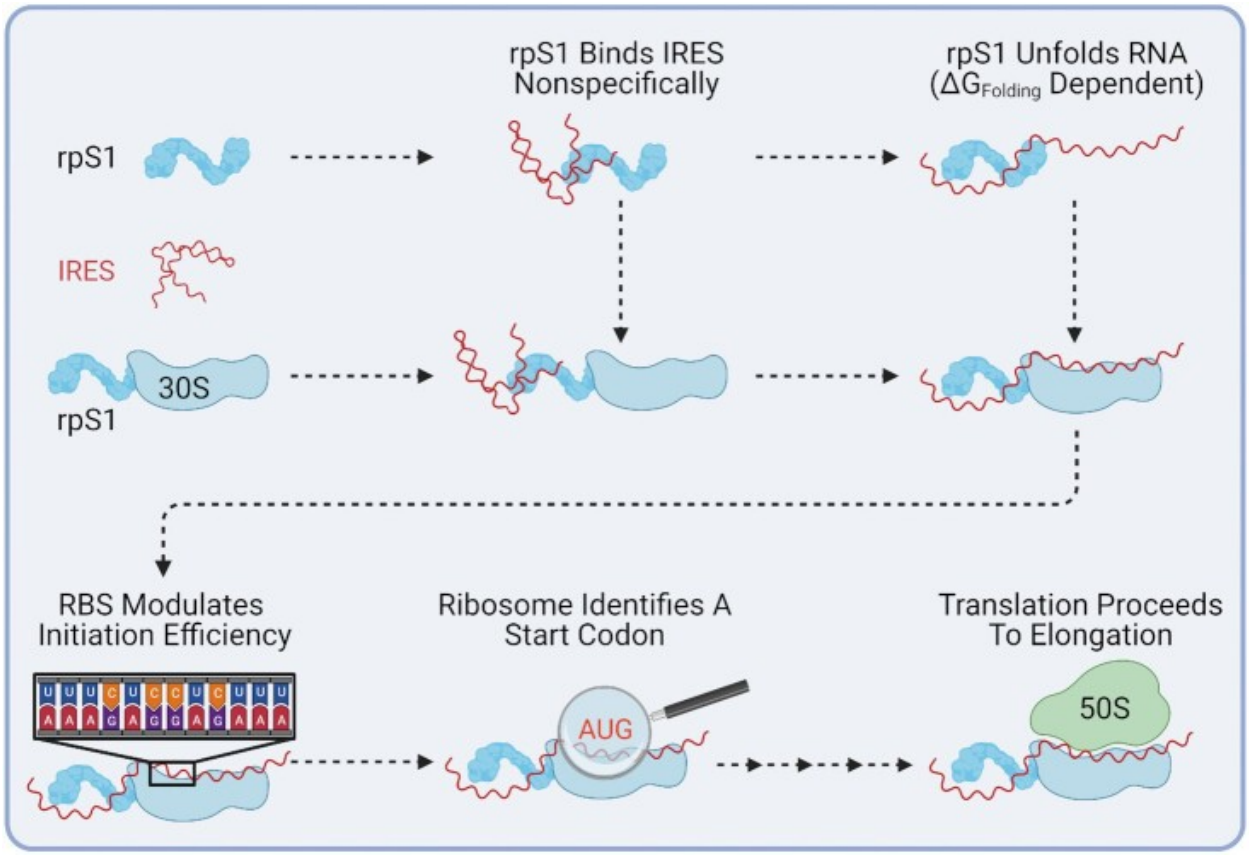

Figure 4.6. Ribosomal protein S1 mediates IGR IRES translation in *E. coli*. Ribosomal protein S1 non-specifically binds mRNA and assists in unfolding the mRNA structure. Next, the base pairing between the Shine-Dalgarno and anti-Shine-Dalgarno occurs, and the start codon is located. Finally, upon 50S subunit recruitment, the ribosome proceeds to translation elongation. For ease of view, initiation factors and initiator tRNA have been omitted.

While taken together, our data indicates that IGR IRES structure is not required and, in fact, even inhibitory to translation in *E. coli*, it is exceedingly difficult to prove that the IGR IRESs are not transiently interacting with prokaryotic ribosomes. In eukaryotes, the IGR IRES uses its triple pseudoknot structure to bind the conserved core of the ribosome and manipulate it into initiating translation in the P site without an initiator tRNA. While the same extent of manipulation does not occur in prokaryotes (an initiator tRNA and start codon are required), the IRES could be

binding to the ribosome in a structure-dependent manner. This claim is supported by the crystal structure of the IGR IRES from the PSIV bound to the *Thermus thermophilus* 70S ribosome [134], in which the IRES occupies the universally conserved tRNA binding sites, similar to what is observed for related IRESs bound to the eukaryotic ribosome [136, 137]. However, there are marked differences which are discussed in the original report [134]. The most significant is that only PK1 (tRNA-mRNA mimic) could be resolved using the density and the remaining PKs needed to be modeled. Additionally, PK1 was found bound to the P site, rather than the A site [134]. Given that deacylated tRNAs are known to be stably bound and copurify with ribosomes in crystallography experiments [204], it is not unreasonable to assume that a tRNA-mRNA mimic could bind to the conserved core of the ribosome in a highly purified and reconstituted system without being physiologically relevant [150]. Interestingly, IRES-ribosome complexes were only detectable in *E. coli* lysate when the PSIV IRES had an 88 nt tail and the elongation inhibitor Hygromycin B was present [134]. It is impossible to differentiate using this assay if the ribosome is bound to the PSIV IRES and not the SDS-like sequence present in the tail, which is why in our binding experiments (Table 4.2 and Supplementary Table S4.3, Supplementary Figures S4.8 and S4.10) we only included a small tail (8 nt) after the IRES. Additionally, Hygromycin B is a ubiquitous inhibitor of elongation and acts through stabilizing tRNA binding to the ribosome [205]. Considering the structural similarity of the IGR IRESs to a tRNA-mRNA duplex, it is possible that Hygromycin B could artificially increase the affinity of the IRES for the ribosome. This is supported by the fact that Hygromycin B does not interfere with the ability of the PSIV IGR IRES to bind ribosomes, but does suppress IGR IRES stimulated eEF2 GTPase activity and initial aa-tRNA binding to PSIV IGR IRES-ribosome complexes [140].

In agreement with previously published data [132, 206], disruption of PK structure significantly reduced the affinity of the IGR IRESs for eukaryotic ribosomes (Supplementary Table S4.3). However, disruption of PK structure had either no effect or a significantly milder effect on prokaryotic ribosome binding for both CrPV or IAPV IGR IRESs (Table 4.2). Disruption of PK1 or PK2 for CrPV and PK1 or PK1+PK3 for IAPV slightly reduced (10-15% and 4-8% respectively) the affinity of the IRESs for the 70S ribosome. Most disruptions had no effect on 70S binding, which is a rather unlikely pathway for translation initiation *in vivo*. More physiologically relevant are the effects of structure disruption on the interaction of the IRES with the 30S subunit. Interestingly, all PK disruptions (with the exception of PK2) for CrPV and disruption of PK1 + PK2 + PK3 for IAPV increased the affinity (25-40% and 10% respectively) for the 30S ribosomal subunit. This is in line with our *in vivo* data indicating that structure has an inhibitory effect on translational efficiency by reducing recruitment of the 30S. Together, this suggests that the IGR IRES are being bound non-specifically by the prokaryotic ribosome (Figure 4.6) rather than through specific structure-based contacts.

Finally, while attempting to determine factors responsible for the translation efficiency of the IGR IRESs, we identified that IGR IRES translation is strongly dependent on ribosomal protein S1 (Figure 4.5B) and the presence of the SD-like sequence upstream of the start codon (Figure 4.4B). This is consistent with S1's canonical role in host translation as it has already been shown to be essential for the translation of most endogenous *E. coli* mRNAs [207]. During translation initiation, S1 binds single-stranded RNA in a sequence-independent fashion (although preference for AU-rich regions has been reported [208, 209]), a critical role as many bacterial genes do not contain Shine-Dalgarno sequences [196, 207, 209-211]. Additionally, S1 binds RNA containing pseudoknots and has been shown to be essential for docking and unfolding of structured mRNAs

on the ribosome [196, 212]. Previous studies have also reported that S1 can allow foreign mRNAs devoid of guanines (and therefore no SD sequence) from a plant-infecting virus to form initiation complexes *in vitro* [210]. It is tempting to imagine ribosomal protein S1 extending into solution and non-specifically binding mRNAs and 'handing them over' to the ribosome, increasing the local concentration in much the same manner as the ribosomal L7/L12 stalk for Elongation Factors Tu and G [23, 213]. Interestingly, in systems where mRNA structure can greatly influence translation (e.g., the PURExpress® system), excess S1 (relative to ribosome concentration) has been shown to improve yield from structured mRNAs including the PSIV IGR IRES [214]. This suggests that the critical activity of ribosomal protein S1 (i.e. RNA unfolding or potentially stabilizing unstructured mRNAs, slowing degradation *in vivo*) may extend beyond ribosomal-bound S1 to free cytosolic S1. Collectively, our data demonstrate that the translational activity of the IGR IRESs is not due to their three-dimensional structure and is primarily the result of the activity of ribosomal protein S1 and the overall robustness of canonical SD-dependent translation (Figure 4.6), closing the door on exploiting these IRES elements for the design of cross-kingdom mRNA-based bioengineering and synthetic biology.

4.7 Acknowledgements

We thank C. B. DeSouza, Q. Daviduck, Y. Yao, and S. Alva for assistance with testing and data collection. We thank J.L.E. Heller and J.R.J. Vigar for the 30S subunits lacking ribosomal protein S1, as well as purified S1 protein. We thank the SynBridge synthetic biology maker space at the University of Lethbridge for access to scientific infrastructure.

4.8 Outlook

Given what we know about the IGR IRESs and the data presented here, it is evident that the prokaryotic activity of the IGR IRESs is simply due to the robustness of the translation

machinery and a conserved structure-based mechanism. Admittedly, this was disappointing as I had goals of characterizing the activity in prokaryotes using an approach where the sequence of the IRES would be randomly mutated and using florescence-activated cell sorting paired with deep sequencing. However, it is clear that this is not worth pursuing as the IGR IRESs are simply being translated by the canonical machinery and treated as large, structured UTRs. I do not foresee this project continuing, barring new evidence being reported in the literature.

Chapter 5.

Towards an atomic structure of the *rpsA* TIR

Roberts L., Vigar J., Heller J., Girodat D, and Wieden HJ

5.1 Foreword

This collaborative project was initiated based on data generated in the Wieden lab suggesting that the activity of the *rpsA* TIR might be dependent on its 3-dimensional structure. Given my interest in structure-based translation in *E. coli* and the determination that the prokaryotic IGR IRES activity was an artefact (see Chapter 4), we decided to collaborate on obtaining a high-resolution structure of the TIR. This chapter presents the work done to date and gives prospects on the continuation of that work. The work presented here was performed by myself, and when not, the work by Justin Vigar, Jalyce Heller, and Dylan Girodat are attributed in the text.

5.2 Abstract

The TIR of the *rpsA* gene in *E. coli* is one of the most efficient drivers of translation in prokaryotes, despite its highly structured nature and its lack of a canonical Shine-Dalgarno sequence [215]. Previous evidence suggests the TIR is utilizing its structure rather than utilizing a canonical SD sequence to drive translation, a phenomenon that has not previously been described in prokaryotes [182, 183]. To investigate this and to determine at atomic resolution how this RNA is interacting with the ribosome/ribosomal protein S1, we set out to obtain a high-resolution structure of a TIR•ribosome complex. Purified ribosomes and ribosomal subunits from *E. coli* as well as *in vitro* transcribed *rpsA* TIR RNA were used with a variety of biophysical techniques to characterize the interactions of these components. First, the binding affinity of the TIR for the 30S was determined using a fluorescently labeled TIR-construct and was estimated to be in the low nanomolar range. Nitrocellulose filter binding experiments further revealed that the affinities for the 30S and the $30S^{-S1}$ are not significantly different, however the fraction bound decreases upon

the removal of S1. Following this, affinity for *rpsA* TIR for the 70S ribosome was determined to be in the low nanomolar range and the affinity for the 50S ribosomal subunit to be closer to the micromolar range. Finally, we used sucrose density ultracentrifugation to demonstrate that the TIR was able to form stable complexes with the 30S, 50S, 70S, and the $30S^{-S1}$.

5.3 Introduction

5.3.1 Non-canonical bacterial translation initiation

While the majority of translation begins via the canonical method of translation initiation (see introduction), there are several examples of non-canonical translation initiation in bacteria. Briefly, canonical bacterial translation initiation is regulated through interactions between the SD located in mRNA TIR and the anti-Shine-Dalgarno (aSD) present in the 16S rRNA [108, 216]. Efficient translation is realized through modulating several factors such as the amount of secondary structure surrounding the start codon, the ability of the SD to bind the aSD, and the distance between the SD and the start codon, among many others [187, 217-221]. Careful biochemical assays and mathematical modeling have led to various tools capable of accurately predicting the translation efficiency of an mRNA based on its primary sequence [152, 222, 223]. These predictive tools, while incredibly powerful, are currently limited to accurately predicting the expression of a single gene or (two gene operon) and struggle to describe translation initiation across the transcriptome [224]. Importantly, they also cannot accurately predict the translation initiation efficiency of the seemingly non-canonical *rpsA* TIR [183]. This is due in part to the surprisingly inconsistent number of prokaryotic mRNAs that contain SDs (~12-91%, depending on the genome), suggesting that undescribed alternate translation initiation mechanism(s) are prevalent [187, 211].

Interestingly, alternate or non-canonical translation initiation mechanisms have been described (e.g., riboswitches, toehold switches, leaderless mRNAs, etc.). Several of these have been studied in detail and have been engineered into tools for synthetic biologists [225-228]. In general, non-canonical mechanisms can be classified into two distinct categories: dynamic regulators and static regulators [187]. The dynamic regulators involve interactions with an effector (e.g., small molecule, RNA, protein), occasionally with the assistance of a second factor (e.g., RNA chaperone), to change the RNA structure, resulting in down/up regulation of gene expression by occluding or exposing previously occluded SDs. The static regulators, however, do not utilize an effector and are less well understood [187]. This category includes the previously discussed IGR IRESs (see Chapters 3 and 4), which had been reported to initiate translation in a structure dependent manner in *E. coli*. This was the first example of structure-mediated translation initiation in bacteria, however as shown in Chapter 4, this was unfortunately only the result of the robustness of the prokaryotic translation machinery. Interestingly however, data from our lab [182, 183] and others [215] indicates that structure-mediated prokaryotic translation initiation may exist in the form of the *rpsA* TIR.

5.3.2. The rpsA TIR

The *E. coli rpsA* gene codes for ribosomal protein S1 (referred to hereafter as S1), a ribosomal protein in γ – proteobacteria [207, 229]. S1 is a component of the small ribosomal subunit (30S) and at 66 kDa is the largest protein in the ribosome [196, 230, 231]. S1 is an essential protein, critical for the translation of many genes in *E. coli* and interestingly utilizes a negative feedback loop at the translation level (binding to its own mRNA) to regulate its own expression [207, 232, 233]. In the cell, it functions as a RNA helicase, binding single-stranded regions of RNAs that are formed transiently during the natural rapid unfolding/refolding of secondary

structure [234-236]. This activity aids in the translation initiation of many structured mRNAs, however, S1 has several other reported roles in cellular processes including translation elongation, transcription, tmRNA translation, and phage replication (where, interestingly, it forms a complex with EF-Tu and EF-Ts) [207, 237-242].

The translation initiation region of the *rpsA* gene (i.e., the TIR) is roughly 90 nucleotides (around 3 times the length of the average *E. coli* UTR [243]) and is composed of three hairpins (Figure 5.1)[233, 244]. The hairpins are separated by single-stranded AU-rich regions, with hairpins I and II containing conserved GGA motifs in their respective loops [233, 244]. This highly structured TIR contains only a weak SD (in the 3' hairpin) but is surprisingly efficient at driving translation, making *rpsA* of one of the highest expressed genes in *E. coli* [244, 245]. Interestingly, the primary sequence of the TIR is not highly conserved, suggesting that the 3-dimensional structure rather than its sequence is the driving factor for the observed translation activity [244]. *In vivo* structural probing has shown that the third hairpin (which contains the AUG start codon and the weak SD) is readily unfolded (or unfolds), suggesting it could potentially be being translated via the canonical mechanism [246]. However, the weak SD has been shown to be dispensable for translational activity, again suggesting that a non-canonical mechanism is mediating the translation initiation activity [244].

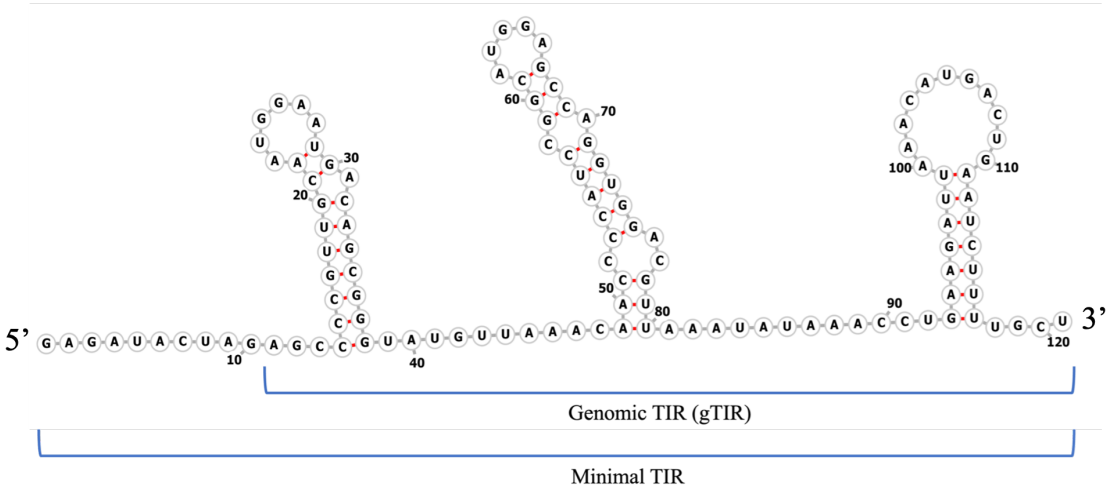

Figure 5.1. Secondary structure of the *rpsA* TIR generated using Forna™. The three hairpin stem loops can be seen: I (nucleotides 14-38), II (nucleotides 49-80), and III (nucleotides 93-117). Base pairs are denoted with red dashes. The minimal TIR is the nucleotides 1-120 and the genomic TIR sequence is nucleotides 11-120.

Previous work using high-throughput mutagenesis, phenotypic cell sorting, and next generation DNA sequencing resulted in a library of TIRs with differential translation efficiencies [183]. Interestingly, while a few expected results were observed (e.g., mutating the start codon abolishes translation) several mutations that could not be explained by the canonical method of translation initiation were observed. For example, a single point mutation ~70 nucleotides away from the start codon significantly reduced translation efficiency [183]. Furthermore, the effect of these mutations were demonstrated to not be due to mRNA stability (*in vivo*) or due to 30S binding (*in vitro*) [182]. However, later data suggests that 30S binding was indeed affected (J.L.E. Heller and J.R.J. Vigar, personal communication). Several of these variants were subjected to structural analysis by small angle x-ray scattering, however the flexible nature of the TIR prevented conclusive insights [182, 183].

Therefore, we set out to determine the atomic resolution model of the TIR. Given its flexible and small nature, it was decided to be complexed with the ribosome to (1) attempt to stabilize the structure and (2) gain insight into its mode of interaction with the ribosome.

5.4 Materials and Methods

5.4.1 Reagents and oligonucleotides

Reagents are as described in the materials and methods section of Chapter 4.

ORL015 – 5'-TAATACGACTCACTATAGGG-3'

ORL107 – 5'-ACTAGTCTTTTGCTTTCG-3'

5.4.2 Purification of the rpsA genomic TIR

Design

A 113 nt portion of the *rpsA* TIR (i.e. the genomic TIR or gTIR) was selected to be used as the RNA component of the TIR•Ribosome complex. Earlier biophysical work had utilized a wildtype "minimal" TIR (121 nt) with a 5' extension, or a wildtype TIR (eCFP) with a portion of a reporter (eCFP or mRFP) fused to the 3' end (154 nt). Because the *in vitro* transcription reaction is known to generate RNAs with varying 5' and 3' ends, it is important to make sure that the obtained RNA has defined ends to facilitate structure determination. The initial design of the genomic TIR utilized a hammerhead ribozyme at the 5' and 3' end which would self-cleave, resulting in RNA with defined ends. However, after months of effort, the 5' ribozyme could not be separated from the TIR and we opted to remove it from the construct.

Template DNA

The genomic *rpsA* TIR (113 nt) was PCR amplified from a plasmid (LR 164) using the primers ORL015 and ORL107. All reactions were carried out using a TGradient (Biometra) thermocycler using Phusion High-Fidelity Polymerase (Thermo-Fisher) according to manufacturer's specifications and the following cycling conditions: 1) 98 °C for 30 seconds, 2) 98 °C for 10 sec, 3) 51.6 °C for 30 sec, 4) 72 °C for 15 sec, 5) return to step 2 thirty (30) times, 6) 72 °C for 300 sec.

In vitro transcription and purification

RpsA TIR RNA was *in vitro* transcribed using the corresponding DNA and T7 polymerase (purified in house) and then purified using size exclusion chromatography as per D'Souza *et al.* [247]. Fractions containing TIR RNA were pooled, and ethanol-precipitated before being resuspended in H_2O.

5.4.3 Purification of ribosomal subunits

Ribosomal subunits, ribosomes, and 30S devoid of S1 were purified from *E. coli* MRE600 as described in Chapter 4.

5.4.4 Nitrocellulose filter binding

RNA was radioactively labelled, and nitrocellulose filter binding was performed as described in Chapter 4. In brief, radio-labeled RNA in $TAKM_5$ buffer was heated to 95 °C for 10 min and slow-cooled to room temperature. RNA was then incubated with increasing amounts of ribosomal subunits/ribosomes for 15 min at 37 °C before being rapidly filtrated through a cellulose nitrate membrane filter (0.2 μm, GE Healthcare). The cellulose nitrate membranes were washed

with 1 mL of cold $TAKM_5$ buffer and dissolved in 10 mL of EcoLite (+) scintillation cocktail (MP Bio), vortexed for 30 sec, and subsequently incubated at room temperature for 30 min, followed by vigorous mixing for 30 sec. The retained radioactivity was quantified by scintillation counting (Tri-carb 2810 TR LSA, Perkin Elmer). The data was fit with a hyperbolic function (equation 5) to determine the K_D for this interaction.

$$Y = \frac{B_{max}X}{K_D+X} \quad (5)$$

5.4.5 Sucrose density ultracentrifugation

RpsA TIR RNA was folded as per Chapter 4 before complexes were formed in $TAKM_5$ using 0.1 nM of radioisotope-labelled TIR RNA and 2,560 nM of ribosomal subunit or ribosome by incubating at 37 °C for 15 minutes. Sucrose gradients were formed using a Gradient Master 108 (BioComp), 17 mL of 50% sucrose was loaded into a 35 mL tube and 17 mL of 15% sucrose was layered on top. Then the pre-set 15-45% long sucrose gradient program was run to generate the gradient. Complexes were carefully pipetted on top of the sucrose gradient before being centrifuged at 26,000 rpm for 16 hrs in a SW28 rotor. Gradients were fractionated using a custom set up whereby a thin glass capillary was pushed to the bottom of the centrifuge tube being careful to avoid mixing. Next the pump and fraction collector of an AKTA Start was utilized to measure the OD_{260nm} and the conductivity while collecting 0.5 mL fractions. The OD_{260nm} and the contained radio-label activity each fraction was measured using a spectrophotometer and a scintillation counter (adding 100 µL of the fraction to 2 mL of scintillation cocktail) respectively.

5.4.6 Fluorescent labelling of the rpsA TIR

RpsA TIR RNA was fluorescently labelled a the 3' end with pyrene or fluorescein as described in Keffer-Wilkes *et al.* [248] and Chapter 4 for the IGR IRES RNA.

5.4.7 Steady State

Labeled RNA in $TAKM_5$ was heated to 95 °C for 10 min and slow-cooled to room temperature. Labeled RNA was then incubated with increasing amounts of ribosomal subunits before being excited at 341 nm. The peak fluorescence at 391 nm was recorded and plotted as a function of increasing 30S concentration using equation 5.

5.4.8 Stopped-flow

Fluorescence stopped-flow measurements were performed using a KinTek SF-2004 stopped-flow instrument. Twenty-five microliters of *rpsA* TIR containing solution (50 nM after mixing) was rapidly mixed with 25 μL of 30S (600 nM after mixing) in $TAKM_5$. Pyrene was excited (λ_{ex} = 341 nm) and the resulting fluorescence was recorded after passing through a LG-400-F long pass filter (NewPort) (42,58). The resulting fluorescent time courses were best fit with a single exponential function (equation 3) using TableCurve software (Jandel Scientific), yielding an apparent rate of association (k_{app}) at this concentration.

5.4.9 Cryogenic electron microscopy

Sample Preparation

Ribosomes or the 30S ribosomal subunit were mixed with *rpsA* TIR in $TAKM_5$ and incubated at 37 °C for 15 min before being flash frozen and stored at -80 °C. Purified ribosomes and *rpsA* TIR were sent for complex formation and structure determination (D. Girodat).

Sample preparation and data collection/processing (Performed by D. Girodat)

Quantifoil R 1.2/1.3 copper grids (Jena Bioscience) were functionalized with a thin mica carbon film and glow discharged with a PELCO easiGlow device. Grids were loaded into a Vitrobot (ThermoFischer) at 100% humidity where 3 µL of TIR-ribosome complexes were added to the grid and subsequently blotted for 4 sec before rapid freezing in liquid ethane. Images were taken on a JEOL-3200 microscope at 300 kV using a Gatan K2 camera at 30,000 x magnification and a total dose of ~ 50 e/Å^2.

Micrographs were compiled in Relion using Motioncor2 and CTFfind 4. Micrographs with greater than 7 Å resolution were selected for further analysis. All particles were selected using Relion template particle picking algorithm using previously determined 2D averages of ribosomes. Particles were extracted with a box size of 512 pixels at a resolution of 1.2 Å/pixel. An initial 2D classification was used to remove cubic ice and debris from ribosomes. A 3D classification, 3D refinement, and post processing were performed using Relion.

5.5 Results

See the design section of the materials and methods for the differences between the gTIR and the minimal TIR (Figure 5.1).

5.5.1 The minimal TIR binds the 30S with a low nanomolar affinity

To ensure homogenous complex formation (complete occupancy of ribosomal subunits/ribosomes with TIR), it is critical to determine the binding affinity. Therefore, I determined the equilibrium binding constant (K_D) by titrating fluorescently-labelled (pyrene) minimal TIR with 30S ribosomal subunits and measuring the resulting change in fluorescence. Plotting the concentration dependence of the observed change in fluorescence allowed for the construction of a binding curve (Figure 5.3).

This initial test using 50 nM TIR suggested the affinity was in the low nanomolar range and therefore accurate determination would require using less TIR. Unfortunately, at 5 nM TIR the data is quite noisy however, but the K_D was determined to be 12 ± 10 nM (Figure 5.2). While the error is large compared to the measurement, it still suggests that the affinity is in the low nanomolar range. However, this means our starting TIR concentration of 5 nM (the same as the K_D within error) is too high to accurately determine affinity as we are starting with at least ~50% RNA bound upon the addition of any amount of 30S. This is consistent with the observation that the initial increase in fluorescence (upon 30S addition) is always the largest, followed by several small changes in fluorescence (upon more 30S addition). Therefore, the affinity of the minimal TIR for the 30S is reported here as ≤ 12 nM.

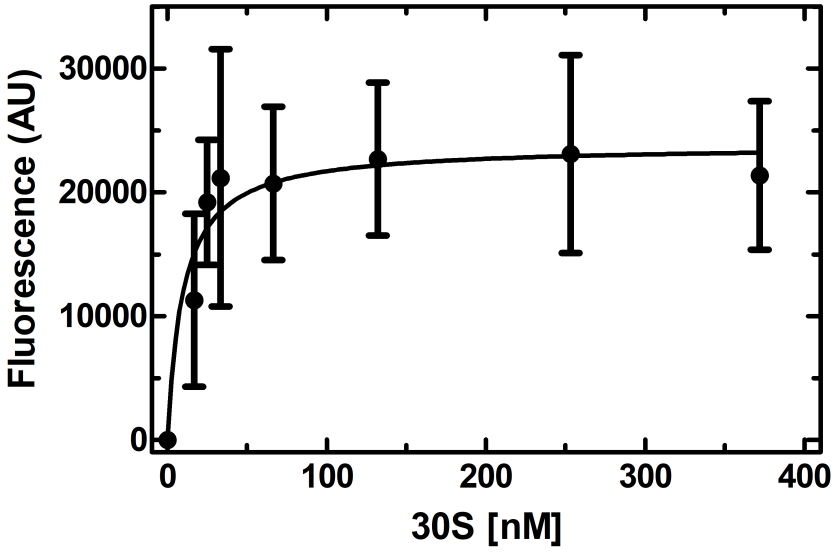

Figure 5.2. The minimal TIR binds the 30S ribosomal subunit with nanomolar affinity. Titration of fluorescently-labelled (pyrene) minimal TIR with 30S ribosomal subunit. Fluorescently labeled RNA is incubated with increasing amounts of 30S. Relative fluorescence emission at 391 nm is shown (λ_{ex} = 341 nm).

5.5.2 Removal of S1 affects fraction bound but not affinity of the rpsA gTIR for the 30S

Interested in the interplay between the TIR and S1, we opted to form several complexes for initial particle screening before high resolution structure determination. Unfortunately, limited structural information regarding S1 is available due to its elongated and flexible nature, however, it is known that S1 anchors onto the small subunit by binding to ribosomal protein S2 near the mRNA exit channel. In order to ensure complex formation, it was critical to ascertain the ability of the TIR to bind the 30S subunit with and without S1. Interestingly, using nitrocellulose filter binding, the affinity of the minimal TIR was determined to be 99 ± 10 nM [182], significantly higher than the estimate generated from the fluorescence equilibrium titration (≤ 12 nM). This

could be due to the non-equilibrium nature of nitrocellulose filtering binding as it does include several buffer wash steps.

The affinity of the genomic TIR (gTIR) for the 30S was determined to be 890 ± 120 nM, suggesting that this "trimmed down" version of the TIR may be less efficient at binding the 30S, perhaps due to fewer single stranded regions (for S1 binding) on each end of the three-hairpin structure. Upon the removal of S1, the affinity of the gTIR for $30S^{-S1}$ is 1120 ± 320 nM, the same within error as the 30S measurements, suggesting that S1 does not modulate the affinity of the gTIR for the 30S (Figure 5.3). However, the fraction of TIR bound decreases to ~23% compared to 75% for intact 30S. The K_D being identical indicates that S1 is not affecting the k_{on} or k_{off} and therefore is not influencing the stability of the complex after binding of the TIR. It is possible given S1's helicase activity that by resolving the structure of the TIR, it is shifting the equilibrium of the TIR able to bind the 30S and as such increasing the fraction bound.

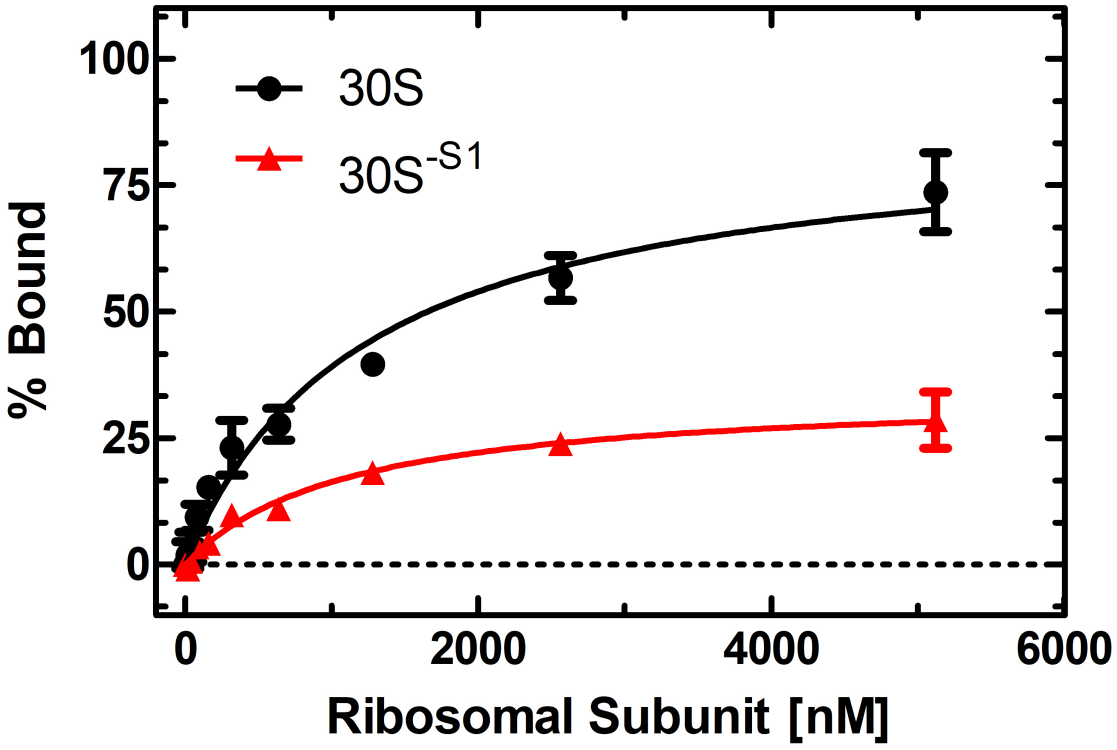

Figure 5.3. The gTIR has a similar affinity but lower occupancy for ribosomes devoid of S1. Nitrocellulose filter binding of gTIR RNA with increasing amounts of 30S or $30S^{-S1}$. Mean values of three replicates are plotted, error bars indicate one standard deviation.

The different affinities between the minimal and gTIR for the 30S prompted measurement of the affinity of the gTIR for the 70S ribosome. This affinity has been reported to be 18 ± 3 nM for the minimal TIR [182], however again the gTIR had a significantly higher affinity of 133 ± 58 nM (Figure 5.4). Together with the 30S data, this suggests that coding sequence downstream of the TIR is important for ribosome binding. Interestingly the fraction of TIR bound to the 70S is only ~38%, this could be due to the mRNA needing to be unfolded and threaded through the ribosome to remain stably bound as the subunits were already bound to each other prior to incubating with the TIR. As the affinity for the 30S and 70S is much higher for the gTIR, the

minimal TIR was selected for continued studies. As a final check of the system, the affinity of the minimal TIR for the 50S subunit was measured and, not surprisingly, the minimal TIR has a much lower affinity of (668 ± 49 nM) for the 50S subunit than it does for the 30S subunit or the 70S ribosome (Figure 5.5).

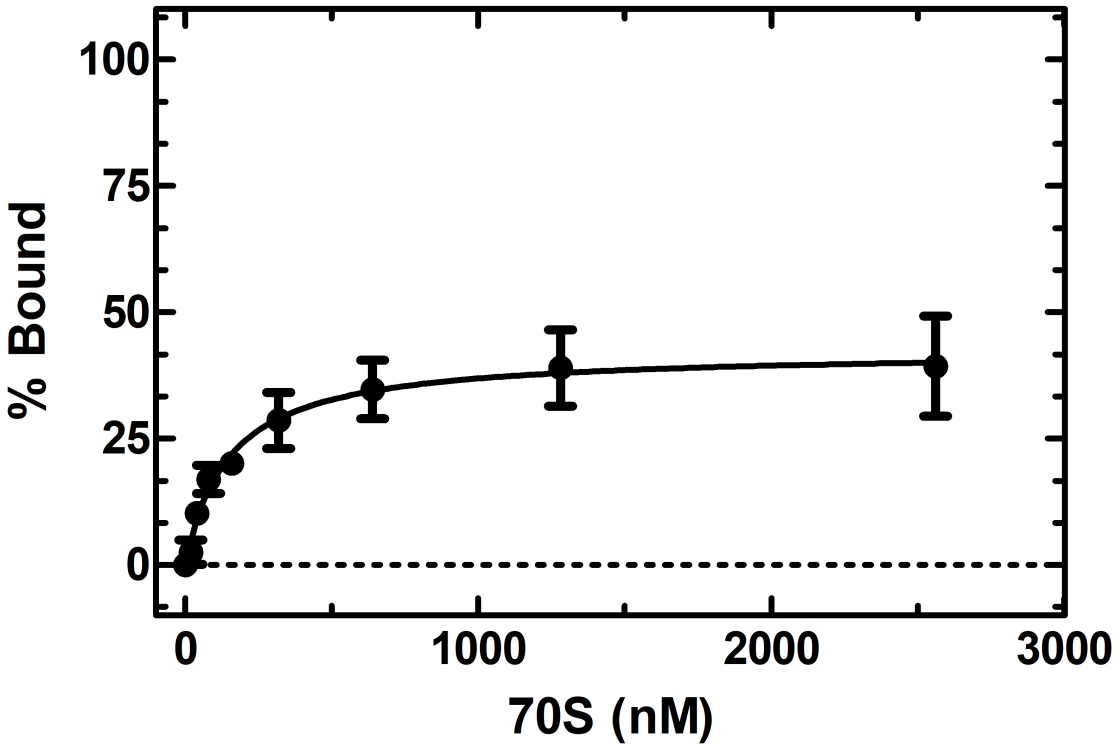

Figure 5.4. The gTIR binds the 70S on the same order of magnitude as the 30S. Nitrocellulose filter binding of gTIR RNA with increasing amounts of 70S. Mean values of three replicates are plotted, error bars indicate one standard deviation.

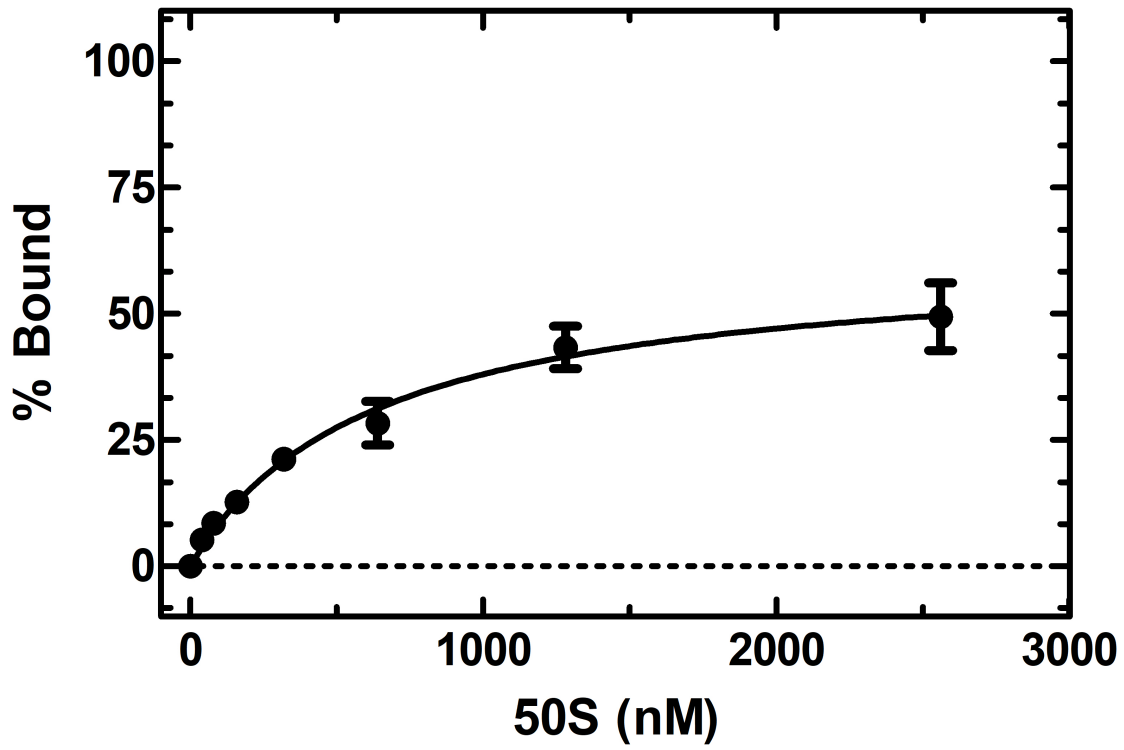

Figure 5.5. The minimal TIR binds the 50S ribosomal subunit. Nitrocellulose filter binding of the minimal TIR RNA with increasing amounts of 50S. Mean values of three replicates are plotted, error bars indicate one standard deviation.

5.5.3 The TIR and the 30S form a stable complex

In order to ascertain that the TIR and the 30S form stable complexes and that the majority of the 30S ribosomal subunits will contain the TIR RNA, we subjected radioactively-labelled TIR and TIR•ribosome complexes to sucrose density ultracentrifugation.

Following the absorbance at 260 nm, we observed that both ribosomal particles the 30S and the $30S^{-S1}$ yield single peaks at fraction 22 (Figure 5.6A and 5.6B). This suggests that S1 removal has left the 30S ribosomal particle largely intact as the migration appears unaffected. This

is not surprising as S1 is small (66 kDa) compared to the ribosome (roughly 2.5 MDa). Either in the presence or absence of S1, a fraction of the TIR is able to bind (indicated by counts in fractions containing the ribosomal particles) and co-migrates with the ribosomal subunit while the TIR alone migrates to fractions ~34-35 (figure 5.6C). The latter also demonstrates that the counts observed in the ribosomal subunit fractions are not simply free TIR RNA or larger TIR aggregates. Together, this confirms that the TIR and the 30S are capable of forming a stable complex during sucrose gradient centrifugation. Interestingly, $30S^{-S1}$ seems to bind a similar amount of the TIR compared to the 30S in contrast to the reduction in binding seen upon S1 removal in our filter binding experiments (Figure 5.3 and 5.6). This could be explained perhaps by the less harsh conditions (no wash steps, as in the filter binding) in the sucrose density ultracentrifugation allowing the TIRs to remain bound.

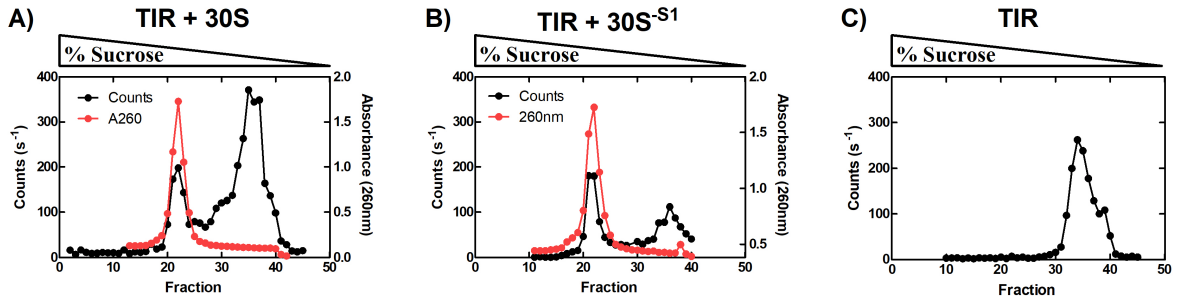

Figure 5.6. The minimal TIR and the 30S subunit form a stable complex. Sucrose density ultracentrifugation of TIR and ribosome TIR complexes. Complexes centrifuged through a 50%-15% sucrose gradient at 26,000 rpm for 16 hrs in a SW28 rotor.

Given the success of the 30S sucrose gradient centrifugation experiments, we decided to see if the TIR can also form stable complexes with the 70S ribosome and the 50S ribosomal subunit. Interestingly, the TIR does co-migrate with 50S and 70S particles, suggesting that even

the lower affinity 50S•TIR complex is relatively stable once formed (Figure 5.7). These experiments were run over the course of two weeks, therefore to account for radioactive decay and compare between the different runs, the CPM was converted to mol of TIR using the specific activity (CPM/mol) for the radioactively labelled TIR. Again, we observe a similar fraction of TIR bound to the $30S^{-S1}$ compared to the other ribosomal subunits and the ribosome (Figure 5.7). This is interesting as it has been reported that S1 is required for the translation of its own mRNA, therefore further investigation is required [207, 244].

Looking at absorbance at 260 nm alone, we observe a shift in the position of the "ribosome" peak upon TIR binding. This is consistent for all three ribosomal particles analysed, the 30S, $30S^{-S1}$, and the 70S, suggesting that TIR binding allows the ribosome to enter further into the sucrose gradient potentially through structural compaction (Figure 5.8). Considering that removal of S1 (66 kDa) did not significantly alter the migration of the 30S, it is unlikely that binding of the minimal TIR (39.2 kDa) to the ribosomal particles is causing this shift due to addition of mass. Unfortunately, the 260 nm data file for the 50S run was corrupted and could not be analyzed.

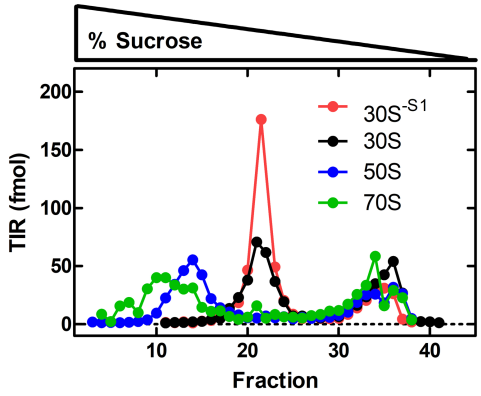

Figure 5.7. The minimal TIR and the ribosome/ribosomal subunits form stable complexes. Sucrose density ultracentrifugation of TIR and ribosome/ribosomal subunit TIR complexes. Complexes centrifuged through a 50%-15% sucrose gradient at 26,000 rpm for 16 hrs in a SW28 rotor. Fractions 21 and 22 from the 30S^{-S1} run (red) were accidentally combined and therefore it appears as a single point at 21.5 in this figure.

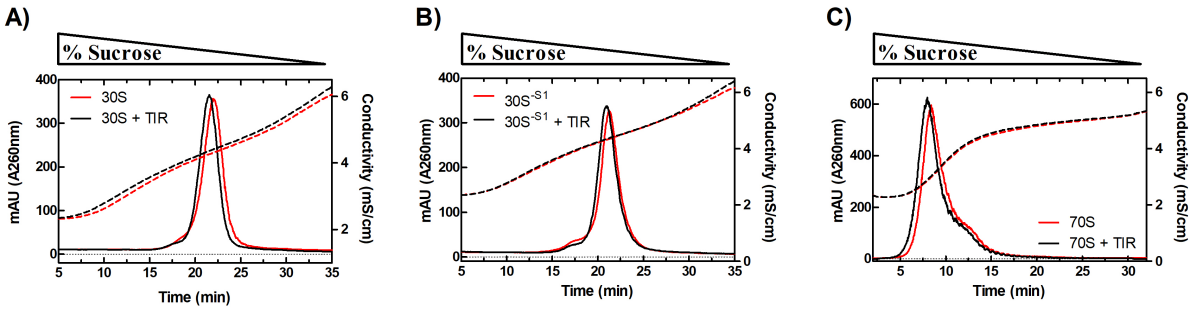

Figure 5.8. The minimal TIR and the ribosome/ribosomal subunits complexes shift in migration compared to ribosomes alone. Sucrose density ultracentrifugation of TIR and ribosome/ribosomal subunit TIR complexes. Complexes centrifuged through a 50%-15% sucrose gradient at 26,000 rpm for 16 hrs in a SW28 rotor. Absorbance at 260 nm (solid lines) and conductivity (dashed lines) are shown.

5.5.4 Initial characterization of the TIR binding to the 30S ribosomal subunit.

I was interested to know if the fluorescently-labelled TIR can be used for time-resolved fluorescence measurements with rapid mixing (stopped-flow), reporting the kinetics of the TIR interacting with the 30S ribosome.

When mixing the *rpsA* TIR with buffer or itself, no meaningful changes in fluorescence (Figure 5.9A) were observed, even over long timescales (1800-2400s). Mixing the TIR against itself resulted in a higher initial fluorescence (as expected because the fluoresce TIR is not diluted with buffer upon mixing) and a small apparent decrease in fluorescence with an extremely slow rate (~2.5/hour), which likely can be attributed to photobleaching (Figure 5.9C). This small change was difficult to fit with such a small amplitude and therefore the error in the measurement was larger the than determined rate ($0.00012 \pm 0.00015 s^{-1}$). To determine if this is a real signal (e.g., photobleaching), more data must be collected to bring the error below that of the measured rate. However, mixing the TIR with 30S resulted in a decrease in fluorescence with an apparent rate of $0.042 \pm 0.014 s^{-1}$ (Figure 5.9B). This indicates that either the fluorescent dye is in a new local environment upon 30S binding or that the TIR is changing conformation after 30S binding, again resulting in a local change of environment. This would need to be repeated at different ribosome concentrations to determine if this rate is concentration dependent (binding vs conformational change). Unfortunately, these experiments could not be continued. However, in case of a concentration-dependence for the signal, measuring the change in apparent rate for the observed fluorescence decrease will allow for the determination of the association rate constant (k_{on}) and with the K_D / Y-axis intercept will allow the calculation of the dissociation rate constant (k_{off}). These kinetic parameters, in addition to those for $30S^{-S1}$ and ribosomal protein S1, will be helpful

in dissecting the translation initiation and the mechanism of gene expression autoregulation of the *rpsA* TIR.

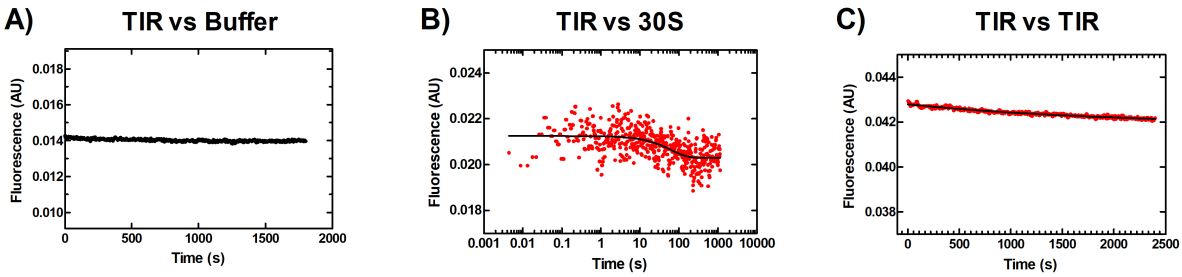

Figure 5.9. The TIR undergoes a change in fluorescence upon binding to the 30S ribosomal subunit. Time resolved change in fluorescence of *rpsA* TIR upon rapid mixing with (A) Buffer, (B) 30S ribosomal subunits, or (C) *rpsA* TIR.

5.6 Outlook

A biophysical characterization of the *rpsA* TIR interacting with the 30S subunit and ribosomal protein S1 is within reach. It would be very interesting to know the difference (if any exist) between the TIR encountering cytosolic S1 and ribosome associated S1 as perhaps this could shed light onto its negative feedback auto regulation capabilities. Additionally, investigating the *rpsA* TIR interaction with ribosomes devoid of S1 ($30S^{-S1}$) will be particularly interesting as it has been reported that ribosomal protein S1 is critical for the translation of its own mRNA [207]. If no cellular S1 is available, how does the cell produce more S1 if it's needed for its own translation?

Chapter 6.

Conclusions

6.1 Concluding Remarks

In this thesis I have described several research projects all focused around the theme of protein translation. To understand the role that conformational dynamics plays in the ability of EF-Tu to execute its critical cellular functions in translation elongation, I developed several fluorescent assays to monitor intramolecular movements in real time. Finding positions to label EF-Tu and have it report changes in conformation was extremely difficult, there are many variants (not described in this thesis) which were either insoluble, inactive, or completely precipitated upon fluorescent labelling. During these experiments I also had to completely disassemble and reassemble the stopped-flow apparatus by myself to service it, which was extremely informative, and I would recommend this to any future graduate students (assuming you are allowed). However, using these assays I was able to determine the order of events for nucleotide release in EF-Tu in that first the protein must undergo a structural change before the bound nucleotide can dissociate. This was postulated to be the case by Dr. Dylan Girodat in our lab based on his work determining the energetic landscape of EF-Tu's nucleotide binding. That my independent work using completely different reporter system primarily looking at the conformational dynamics of EF-Tu reaches identical conclusions is very reassuring and strong evidence that our data is reliable. I was also able to demonstrate that when in complex with EF-Ts, although the nucleotides are being exchanged, no large-scale structural changes in EF-Tu occur. Combined, this now allows me to propose an alternate mode of nucleotide exchange where EF-Ts is accelerating the exchange by circumventing the slow conformational changes (to and from the *apo* state) of EF-Tu. The above-mentioned reporter systems will enable us to answer a number of questions regarding EF-Tu's conformation dynamics during its functional cycle and could be utilized to screen compound libraries for novel antibiotics that inhibit EF-Tu's conformational dynamics.

To investigate the previously reported structure-based prokaryotic translation initiation activity of the IGR IRES in *E. coli*, I combined several *in vivo* and *in vitro* approaches. When I began this project the paper from Colussi *et al.* had not been published and I thought I was the first person to think about the IRES working in bacteria. I was crestfallen when the publication came out in Nature but upon closer inspection it had been in review there for over a year before being published which did alleviate some of the sadness. I did have an original idea, I was just not the first to have it. I decided to compare IGR IRES translation between the domains of life and using random mutagenesis, flow cytometry, and next generation sequencing. However, my initial results never lined up with the previous report and I spent far too long doubting myself. Interestingly, I found no evidence of conserved RNA structure being critical for translation activity, suggesting that the IRES was simply being translated as a large structured 5' UTR via the canonical translation mechanism. However, after spending years checking and double checking, I was confident in my data and myself and ultimately, I was able to identify ribosomal protein S1 as the main driver for IGR IRES translation in *E. coli*. This story was originally submitted to Nature as a Brief Communication Arising in 2017, it spent 1 year in review, had 4 reviewers and a lot of back and forth with the editor and the authors of the original study reporting structure-based translation in bacteria (who were adamant in their data, but very fair overall in the process). The response to the comments from the corresponding author required many extra experiments/techniques some of which I had to learn or develop and ultimately addressed the concerns raise by the authors, supporting our finding. However, ultimately this ended in a rejection of my manuscript. It was then rewritten completely focusing on how bacteria translate DNA obtained from their environment and submitted to Nature Communications (desk rejected), and then Molecular Cell. The Molecular Cell editor rejected it based on the reviewer comments, which after reading I

realised I could not dispute, it was the wrong story. It was then rewritten a third time in the form you see now and submitted to Nucleic Acids Research. Again here we had four reviewers and was rejected but the editor (Dr. Eric Westhof) left open the possibility of a resubmission if we could address the concerns of the reviewers. As luck would have it, I was going to the RNA Society meeting where I knew Dr. Westhof would be in attendance. I spent the next two months furiously doing the experiments and making the suggested changes, I then tracked Dr. Westhof down on the first day of the conference, explained who I was and asked him to come to my poster and that I would show him I had listened and that the paper was ready for resubmission. It must have been a good poster presentation as the resubmission was accepted with no revisions, capping the wonderful underdog story of a smaller lab publishing data that disputed already published data from a much larger lab in a very important scientific journal.

Finally, to understand a different seemingly structure-based prokaryotic translation initiation mechanisms utilised by the *rpsA* TIR, I purified and characterized this molecule. Biophysical characterization of this RNA in complex with 30S, 50S, and 70S ribosomes demonstrated that it forms a stable complex suitable for high resolution structure determination using cryo-EM. A high-resolution structure would answer several pressing questions regarding the mechanism by which the *rpsA* TIR initiates translation and potentially shed light onto the autoregulation of this mRNA by its gene product ribosomal protein S1. While I was only able to begin a small part of this project, I am excited to see where it will continue as in my opinion this molecule is the best bet for *bona fide* structure based translation in prokaryotes.

This thesis is distillation of best portions of the work done during my time in graduate school, if you're a prospective student know this, science is as hard as everyone says it is, but it's 10X as fun as everyone says it is.

References

1. Hurwitz, J., A. Bresler, and R. Diringer, *The enzymic incorporation of ribonucleotides into polyribonucleotides and the effect of DNA.* BBRC, 1960. **3**(1): p. 15-19.
2. Weiss, S.B., *Enzymatic incorporation of ribonucleoside triphosphates into the interpolynucleotide linkages of ribonucleic acid.* Proc Natl Acad Sci U S A, 1960. **46**(8): p. 1020-1030.
3. Stephenson, M.L., K.V. Thimann, and P.C. Zamecnik, *Incorporation of C14-amino acids into proteins of leaf disks and cell-free fractions of tobacco leaves.* Arch Biochem Biophys, 1956. **65**(1): p. 194-209.
4. Bonner, J., P.O. Ts'O, and J. Vinograd, *Microsomal nucleoprotein particles from pea seedlings.* J Biophys Biochem Cytol, 1956. **2**(4): p. 451-466.
5. Schachman, H.K., A.B. Pardee, and R.Y. Stanier, *Studies on the macro-molecular organization of microbial cells.* Arch Biochem Biophys, 1952. **38**: p. 245-260.
6. Petermann, M.L. and M.G. Hamilton, *The purification and properties of cytoplasmic ribonucleoprotein from rat liver.* J Biol Chem, 1957. **224**(2): p. 725-736.
7. Roberts, L. and H.J. Wieden, *Viruses, IRESs, and a universal translation initiation mechanism.* Biotechnol Genet Eng Rev, 2018. **34**(1): p. 60-75.
8. Rheinberger, H.J., H. Sternbach, and K.H. Nierhaus, *Three tRNA binding sites on Escherichia coli ribosomes.* Proc Natl Acad Sci U S A, 1981. **78**(9): p. 5310-5314.
9. Subramanian, A.R. and B.D. Davis, *Activity of initiation factor F3 in dissociating Escherichia coli ribosomes.* Nature, 1970. **228**(5278): p. 1273-1275.

10. Wintermeyer, W. and C. Gualerzi, *Effect of Escherichia coli initiation factors on the kinetics of N-Acphe-tRNAPhe binding to 30S ribosomal subunits. A fluorescence stopped-flow study.* Biochemistry, 1983. **22**(3): p. 690-694.

11. Anderson, J.S., et al., *A GTP requirement for binding initiator tRNA to ribosomes.* Nature, 1967. **215**(5100): p. 490-492.

12. Antoun, A., et al., *The roles of initiation factor 2 and guanosine triphosphate in initiation of protein synthesis.* EMBO J, 2003. **22**(20): p. 5593-5601.

13. Gualerzi, C.O. and C.L. Pon, *Initiation of mRNA translation in bacteria: structural and dynamic aspects.* Cell Mol Life Sci, 2015. **72**(22): p. 4341-4367.

14. Lockwood, A.H., P.R. Chakraborty, and U. Maitra, *A complex between initiation factor IF2, guanosine triphosphate, and fMet-tRNA: an intermediate in initiation complex formation.* Proc Natl Acad Sci U S A, 1971. **68**(12): p. 3122-3126.

15. Milon, P., et al., *Kinetic checkpoint at a late step in translation initiation.* Mol Cell, 2008. **30**(6): p. 712-720.

16. Milon, P., et al., *Real-time assembly landscape of bacterial 30S translation initiation complex.* Nat Struct Mol Biol, 2012. **19**(6): p. 609-615.

17. Rodnina, M.V., et al., *GTPases mechanisms and functions of translation factors on the ribosome.* Biol Chem, 2000. **381**(5-6): p. 377-387.

18. Hartz, D., D.S. McPheeters, and L. Gold, *Selection of the initiator tRNA by Escherichia coli initiation factors.* Genes Dev, 1989. **3**(12A): p. 1899-1912.

19. Pape, T., W. Wintermeyer, and M.V. Rodnina, *Complete kinetic mechanism of elongation factor Tu-dependent binding of aminoacyl-tRNA to the A site of the E. coli ribosome.* EMBO J, 1998. **17**(24): p. 7490-7497.

20. Rodnina, M.V., et al., *Initial binding of the elongation factor Tu.GTP.aminoacyl-tRNA complex preceding codon recognition on the ribosome.* J Biol Chem, 1996. **271**(2): p. 646-652.

21. Eccleston, J.F., D.B. Dix, and R.C. Thompson, *The rate of cleavage of GTP on the binding of Phe-tRNA.elongation factor Tu.GTP to poly(U)-programmed ribosomes of Escherichia coli.* J Biol Chem, 1985. **260**(30): p. 16237-16241.

22. Rodnina, M.V., et al., *Codon-dependent conformational change of elongation factor Tu preceding GTP hydrolysis on the ribosome.* EMBO J, 1995. **14**(11): p. 2613-2619.

23. Kothe, U. and M.V. Rodnina, *Delayed release of inorganic phosphate from elongation factor Tu following GTP hydrolysis on the ribosome.* Biochemistry, 2006. **45**(42): p. 12767-12774.

24. Blanchard, S.C., et al., *tRNA selection and kinetic proofreading in translation.* Nat Struct Mol Biol, 2004. **11**(10): p. 1008-1014.

25. Noel, J.K. and P.C. Whitford, *How EF-Tu can contribute to efficient proofreading of aa-tRNA by the ribosome.* Nat Comm, 2016. **7**(1): p. 13314.

26. Girodat, D., et al., *Elongation factor Tu switch I element is a gate for aminoacyl-tRNA selection.* J Mol Biol, 2020. **432**(9): p. 3064-3077.

27. Rodnina, M.V., *Translation in Prokaryotes.* Cold Spring Harb Perspect Biol, 2018. **10**(9): p. a032664.

28. Epstein, R.H., A. Bolle, and C.M. Steinberg, *Amber mutants of bacteriophage T4D: their isolation and genetic characterization.* Genetics, 2012. **190**(3): p. 833-840.

29. Brenner, S., A.O.W. Stretton, and S. Kaplan, *Genetic code: The 'nonsense' triplets for chain termination and their suppression.* Nature, 1965. **206**(4988): p. 994-998.

30. Brenner, S., et al., *UGA: A Third Nonsense Triplet in the Genetic Code.* Nature, 1967. **213**(5075): p. 449-450.

31. Benzer, S. and S.P. Champe, *Ambivalent rII mutants of phage T4.* Proc Natl Acad Sci U S A, 1961. **47**(7): p. 1025-1038.

32. Capecchi, M.R., *Polypeptide chain termination in vitro: isolation of a release factor.* Proc Natl Acad Sci U S A, 1967. **58**(3): p. 1144-1151.

33. Scolnick, E., et al., *Release factors differing in specificity for terminator codons.* Proc Natl Acad Sci U S A, 1968. **61**(2): p. 768-774.

34. Freistroffer, D.V., et al., *Release factor RF3 in E.coli accelerates the dissociation of release factors RF1 and RF2 from the ribosome in a GTP-dependent manner.* EMBO J, 1997. **16**(13): p. 4126-4133.

35. Hirashima, A. and A. Kaji, *Factor dependent breakdown of polysomes.* BBRC, 1970. **41**(4): p. 877-883.

36. Hirashima, A. and A. Kaji, *Purification and properties of ribosome-releasing factor.* Biochemistry, 1972. **11**(22): p. 4037-4044.

37. Hirashima, A. and A. Kaji, *Factor-dependent release of ribosomes from messenger RNA. Requirement for two heat-stable factors.* J Mol Biol, 1972. **65**(1): p. 43-58.

38. Hirashima, A. and A. Kaji, *Role of elongation factor G and a protein factor on the release of ribosomes from messenger ribonucleic acid.* J Biol Chem, 1973. **248**(21): p. 7580-7587.

39. Verstraeten, N., et al., *The universally conserved prokaryotic GTPases.* MMBR, 2011. **75**(3): p. 507-542.

40. Bourne, H.R., D.A. Sanders, and F. McCormick, *The GTPase superfamily: a conserved switch for diverse cell functions.* Nature, 1990. **348**(6297): p. 125-132.

41. Song, H., et al., *Crystal structure of intact elongation factor EF-Tu from Escherichia coli in GDP conformation at 2.05 A resolution.* J Mol Biol, 1999. **285**(3): p. 1245-1256.

42. Kjeldgaard, M. and J. Nyborg, *Refined structure of elongation factor EF-Tu from Escherichia coli.* J Mol Biol, 1992. **223**(3): p. 721-742.

43. Nissen, P., et al., *Crystal structure of the ternary complex of Phe-tRNAPhe, EF-Tu, and a GTP analog.* Science, 1995. **270**(5241): p. 1464-1472.

44. Neal, S.E., et al., *Kinetic analysis of the hydrolysis of GTP by p21N-ras. The basal GTPase mechanism.* J Biol Chem, 1988. **263**(36): p. 19718-19722.

45. Gromadski, K.B., H.J. Wieden, and M.V. Rodnina, *Kinetic mechanism of elongation factor Ts-catalyzed nucleotide exchange in elongation factor Tu.* Biochemistry, 2002. **41**(1): p. 162-169.

46. Schummer, T., K.B. Gromadski, and M.V. Rodnina, *Mechanism of EF-Ts-catalyzed guanine nucleotide exchange in EF-Tu: contribution of interactions mediated by helix B of EF-Tu.* Biochemistry, 2007. **46**(17): p. 4977-4984.

47. Takai, Y., T. Sasaki, and T. Matozaki, *Small GTP-Binding Proteins.* Physiol Rev, 2001. **81**(1): p. 153-208.

48. Furano, A.V., *Content of elongation factor Tu in Escherichia coli.* Proc Natl Acad Sci U S A, 1975. **72**(12): p. 4780-4784.

49. Ishihama, Y., et al., *Protein abundance profiling of the Escherichia coli cytosol.* BMC Genomics, 2008. **9**: p. 102.

50. De Laurentiis, E.I., F. Mo, and H.J. Wieden, *Construction of a fully active Cys-less elongation factor Tu: functional role of conserved cysteine 81.* Biochim Biophys Acta, 2011. **1814**(5): p. 684-692.

51. Ivell, R., G. Sander, and A. Parmeggiani, *Modulation by monovalent and divalent cations of the guanosine-5'-triphosphatase activity dependent on elongation factor Tu.* Biochemistry, 1981. **20**(24): p. 6852-6859.

52. Gilman, A.G., *G proteins: transducers of receptor-generated signals.* Annu Rev Biochem, 1987. **56**: p. 615-649.

53. Parmeggiani, A., et al., *Properties of a genetically engineered G domain of elongation factor Tu.* Proc Natl Acad Sci U S A, 1987. **84**(10): p. 3141-3145.

54. Maracci, C., et al., *Ribosome-induced tuning of GTP hydrolysis by a translational GTPase.* Proc Natl Acad Sci U S A, 2014. **111**(40): p. 14418-14423.

55. Young, R. and H. Bremer, *Polypeptide-chain-elongation rate in Escherichia coli B/r as a function of growth rate.* Biochem J, 1976. **160**(2): p. 185-194.

56. Shi, X., et al., *Functional role of the sarcin-ricin loop of the 23S rRNA in the elongation cycle of protein synthesis.* J Mol Biol, 2012. **419**(3-4): p. 125-138.

57. Sanbonmatsu, K.Y., *Alignment/misalignment hypothesis for tRNA selection by the ribosome.* Biochimie, 2006. **88**(8): p. 1075-1089.

58. Schmeing, T.M., et al., *The crystal structure of the ribosome bound to EF-Tu and aminoacyl-tRNA.* Science, 2009. **326**(5953): p. 688-694.

59. Wilden, B., et al., *Role and timing of GTP binding and hydrolysis during EF-G-dependent tRNA translocation on the ribosome.* Proc Natl Acad Sci U S A, 2006. **103**(37): p. 13670-13675.

60. Peske, F., et al., *Timing of GTP binding and hydrolysis by translation termination factor RF3.* Nucleic Acids Res, 2014. **42**(3): p. 1812-1820.

61. Hauryliuk, V., et al., *Thermodynamics of GTP and GDP binding to bacterial initiation factor 2 suggests two types of structural transitions.* J Mol Biol, 2009. **394**(4): p. 621-626.

62. Girodat, D., et al., *Elongation factor Tu's nucleotide binding Is governed by a thermodynamic landscape unique among bacterial translation factors.* J Am Chem Soc, 2019. **141**(26): p. 10236-10246.

63. Buckstein, M.H., J. He, and H. Rubin, *Characterization of nucleotide pools as a function of physiological state in Escherichia coli.* J Bacteriol, 2008. **190**(2): p. 718-726.

64. Bochner Br Fau - Ames, B.N. and B.N. Ames, *Complete analysis of cellular nucleotides by two-dimensional thin layer chromatography.* J Biol Chem, 1989(0021-9258 (Print)): p. 9759-9769.

65. Loveland, A.B., et al., *Ensemble cryo-EM elucidates the mechanism of translation fidelity.* Nature, 2017(1476-4687 (Electronic)): p. 113-117.

66. Fislage, M., et al., *Cryo-EM shows stages of initial codon selection on the ribosome by aa-tRNA in ternary complex with GTP and the GTPase-deficient EF-TuH84A.* Nucleic Acids Res, 2018. **46**(11): p. 5861-5874.

67. Kavaliauskas, D., et al., *Structural dynamics of translation elongation factor Tu during aa-tRNA delivery to the ribosome.* Nucleic Acids Res, 2018. **46**(16): p. 8651-8661.

68. Johansen, J.S., et al., *E. coli elongation factor Tu bound to a GTP analogue displays an open conformation equivalent to the GDP-bound form.* Nucleic Acids Res, 2018. **46**(16): p. 8641-8650.

69. Thirup, S.S., et al., *Structural outline of the detailed mechanism for elongation factor Ts-mediated guanine nucleotide exchange on elongation factor Tu.* J Struct Biol, 2015. **191**(1): p. 10-21.

70. Garst, A.D., A.L. Edwards, and R.T. Batey, *Riboswitches: structures and mechanisms.* Cold Spring Harb Perspect Biol, 2011. **3**(6): p. a003533.

71. Stein, V. and K. Alexandrov, *Synthetic protein switches: design principles and applications.* Trends Biotechnol, 2015. **33**(2): p. 101-110.

72. Harris Jared, D., J. Moran Mark, and I. Aprahamian, *New molecular switch architectures.* Proc Natl Acad Sci U S A, 2018. **115**(38): p. 9414-9422.

73. Bos, J.L., *Ras oncogenes in human cancer: a review.* Cancer Res, 1989(0008-5472 (Print)): p. 4682-4689.

74. Wolf, H., G. Chinali, and A. Parmeggiani, *Mechanism of the inhibition of protein synthesis by kirromycin. Role of elongation factor Tu and ribosomes.* Eur J Biochem, 1977. **75**(1): p. 67-75.

75. Savelsbergh, A., M.V. Rodnina, and W. Wintermeyer, *Distinct functions of elongation factor G in ribosome recycling and translocation.* RNA, 2009. **15**(5): p. 772-780.

76. Buskiewicz, I., et al., *Conformations of the signal recognition particle protein Ffh from Escherichia coli as determined by FRET.* J Mol Biol, 2005. **351**(2): p. 417-30.

77. Blanchard, S.C., et al., *tRNA selection and kinetic proofreading in translation.* 2004(1545-9993 (Print)).

78. Ogle, J.M., et al., *Selection of tRNA by the ribosome requires a transition from an open to a closed form.* Cell, 2002. **111**(5): p. 721-732.

79. Ogle, J.M., et al., *Recognition of cognate transfer RNA by the 30S ribosomal subunit.* Science, 2001. **292**(5518): p. 897-902.

80. Weissbach, H., D. Lee Miller, and J. Hachmann, *Studies on the role of factor Ts in polypeptide synthesis.* Arch Biochem Biophys, 1970. **137**(1): p. 262-269.

81. Kawashima, T., et al., *The structure of the Escherichia coli EF-Tu.EF-Ts complex at 2.5 A resolution.* Nature, 1996. **379**(6565): p. 511-518.

82. MATLAB, *Version R2019b*. 2010: The Mathworks Inc.

83. Mercier, E., D. Girodat, and H.J. Wieden, *A conserved P-loop anchor limits the structural dynamics that mediate nucleotide dissociation in EF-Tu.* Sci Rep, 2015. **5**: p. 7677.

84. Milon, P., et al., *Transient kinetics, fluorescence, and FRET in studies of initiation of translation in bacteria.* Methods Enzymol, 2007. **430**: p. 1-30.

85. Phillips, J.C., et al., *Scalable molecular dynamics with NAMD.* J Comput Chem, 2005. **26**(16): p. 1781-1802.

86. Humphrey, W., A. Dalke, and K. Schulten, *VMD: visual molecular dynamics.* J Mol Graph, 1996. **14**(1): p. 33-8, 27-8.

87. Smith, D.D., et al., *Construction of a highly selective and sensitive carbohydrate-detecting biosensor utilizing computational identification of non-disruptive conjugation sites (CINC) for flexible and streamlined biosensor design.* Biosens Bioelectron, 2022. **200**: p. 113899.

88. Glaser, F., et al., *ConSurf: identification of functional regions in proteins by surface-mapping of phylogenetic information.* Bioinformatics, 2003. **19**(1): p. 163-164.

89. Landau, M., et al., *ConSurf 2005: the projection of evolutionary conservation scores of residues on protein structures.* Nucleic Acids Res, 2005. **33**(Web Server issue): p. W299-302.

90. Pope, J.R., et al., *Association of fluorescent protein pairs and its significant impact on fluorescence and energy transfer.* Adv Sci (Weinh), 2020. **8**(1): p. 2003167.

91. Fernandes, D.D., et al., *Characterization of fluorescein arsenical hairpin (FlAsH) as a probe for single-molecule fluorescence spectroscopy.* Sci Rep, 2017. **7**(1): p. 13063.

92. Wieden, H.-J., et al., *Mechanism of elongation factor (EF)-Ts-catalyzed nucleotide exchange in EF-Tu: contribution of contacts at the guanine base.* J Biol Chem, 2002. **277**(8): p. 6032-6036.

93. Smith, D.D., et al., *Development of a real-time pectic oligosaccharide-detecting biosensor using the rapid and flexible computational identification of non-disruptive conjugation sites (CINC) biosensor design platform.* Sensors (Basel), 2022. **22**(3): p. 948.

94. Wang, Y., et al., *Crystal structure of the EF-Tu˙EF-Ts complex from Thermus thermophilus.* Nat Struct Biol 1997. **4**(8): p. 650-656.

95. Warias, M., H. Grubmüller, and L.V. Bock, *tRNA Dissociation from EF-Tu after GTP Hydrolysis: Primary Steps and Antibiotic Inhibition.* Biophys J, 2020. **118**(1): p. 151-161.

96. Burnett, B.J., et al., *Direct evidence of an elongation factor-Tu/Ts·GTP·Aminoacyl-tRNA quaternary complex.* J Biol Chem, 2014. **289**(34): p. 23917-23927.

97. Ott, G., et al., *Ternary complexes of Escherichia coli aminoacyl-tRNAs with the elongation factor Tu and GTP: thermodynamic and structural studies.* Biochim Biophys Acta, 1990. **1050**(1-3): p. 222-225.

98. Bretscher, M.S. and K.A. Marcker, *Polypeptidyl-sigma-ribonucleic acid and amino-acyl-sigma-ribonucleic acid binding sites on ribosomes.* Nature, 1966. **211**(5047): p. 380-384.

99. Zamir, A., P. Leder, and D. Elson, *A ribosome-catalyzed reaction between N-formylmethionyl-trna and puromycin.* Proc Natl Acad Sci U S A, 1966. **56**(6): p. 1794-1801.

100. Kennell, D. and H. Riezman, *Transcription and translation initiation frequencies of the Escherichia coli lac operon.* J Mol Biol, 1977. **114**(1): p. 1-21.

101. Hoagland, M.B., et al., *A soluble ribonucleic acid intermediate in protein synthesis.* J Biol Chem, 1958. **231**(1): p. 241-257.

102. Gordon, J., *Interaction of guanosine 5'-triphosphate with a supernatant fraction from E. coli and aminoacyl-sRNA.* Proc Natl Acad Sci U S A, 1967. **58**(4): p. 1574-1578.

103. Gordon, J., *Hydrolysis of guanosine 5'-triphosphate associated wh binding of aminoacyl transfer ribonucleic acid to ribosomes.* J Biol Chem, 1969. **244**(20): p. 5680-5686.

104. Pape, T., W. Wintermeyer, and M. Rodnina, *Induced fit in initial selection and proofreading of aminoacyl-tRNA on the ribosome.* EMBO J, 1999. **18**(13): p. 3800-3807.

105. Dintzis, H.M., *Assembly of the peptide chains of hemoglobin.* Proc Natl Acad Sci U S A, 1961. **47**: p. 247-261.

106. Nishizuka, Y. and F. Lipmann, *The interrelationship between guanosine triphosphatase and amino acid polymerization.* Arch Biochem Biophys, 1966. **116**(1): p. 344-351.

107. Nishizuka, Y. and F. Lipmann, *Comparison of guanosine triphosphate split and polypeptide synthesis with a purified E. coli system.* Proc Natl Acad Sci U S A, 1966. **55**(1): p. 212-219.

108. Shine, J. and L. Dalgarno, *The 3'-terminal sequence of Escherichia coli 16S ribosomal RNA: complementarity to nonsense triplets and ribosome binding sites.* Proc Natl Acad Sci U S A, 1974. **71**(4): p. 1342-1346.

109. Jackson, R.J., C.U. Hellen, and T.V. Pestova, *The mechanism of eukaryotic translation initiation and principles of its regulation.* Nat Rev Mol Cell Biol, 2010. **11**(2): p. 113-127.

110. Kozak, M., *Point mutations define a sequence flanking the AUG initiator codon that modulates translation by eukaryotic ribosomes.* Cell, 1986. **44**(2): p. 283-292.

111. Jan, E., *Divergent IRES elements in invertebrates.* Virus Res, 2006. **119**(1): p. 16-28.

112. Pelletier, J., et al., *Cap-independent translation of poliovirus mRNA is conferred by sequence elements within the 5' noncoding region.* Mol Cell Biol, 1988. **8**(3): p. 1103-1112.

113. Pelletier, J. and N. Sonenberg, *Internal initiation of translation of eukaryotic mRNA directed by a sequence derived from poliovirus RNA.* Nature, 1988. **334**(6180): p. 320-325.

114. Wilson, J.E., et al., *Naturally occurring dicistronic cricket paralysis virus RNA is regulated by two internal ribosome entry sites.* Mol Cell Biol, 2000. **20**(14): p. 4990-4999.

115. Kieft, J.S., *Viral IRES RNA structures and ribosome interactions.* Trends Biochem Sci, 2008. **33**(6): p. 274-283.

116. Filbin, M.E. and J.S. Kieft, *Toward a structural understanding of IRES RNA function.* Curr Opin Struct Biol, 2009. **19**(3): p. 267-276.

117. Martinez-Salas, E., et al., *Picornavirus IRES elements: RNA structure and host protein interactions.* Virus Res, 2015. **206**: p. 62-73.

118. Gradi, A., et al., *Proteolysis of human eukaryotic translation initiation factor eIF4GII, but not eIF4GI, coincides with the shutoff of host protein synthesis after poliovirus infection.* Proc Natl Acad Sci U S A, 1998. **95**(19): p. 11089-11094.

119. Sweeney, T.R., et al., *The mechanism of translation initiation on Type 1 picornavirus IRESs.* EMBO J, 2014. **33**(1): p. 76-92.

120. Kafasla, P., et al., *Polypyrimidine tract binding protein stabilizes the encephalomyocarditis virus IRES structure via binding multiple sites in a unique orientation.* Mol Cell, 2009. **34**(5): p. 556-568.

121. Yu, Y., et al., *Common conformational changes induced in type 2 picornavirus IRESs by cognate trans-acting factors.* Nucleic Acids Res, 2011. **39**(11): p. 4851-4865.

122. Jaafar, Z.A., et al., *Translation initiation by the hepatitis C virus IRES requires eIF1A and ribosomal complex remodeling.* eLife, 2016. **5**: p. e21198.

123. Pestova, T.V., et al., *eIF2-dependent and eIF2-independent modes of initiation on the CSFV IRES: a common role of domain II.* EMBO J, 2008. **27**(7): p. 1060-1072.

124. Pestova, T.V., et al., *A prokaryotic-like mode of cytoplasmic eukaryotic ribosome binding to the initiation codon during internal translation initiation of hepatitis C and classical swine fever virus RNAs.* Genes Dev, 1998. **12**(1): p. 67-83.

125. Kim, J.H., et al., *eIF2A mediates translation of hepatitis C viral mRNA under stress conditions.* EMBO J, 2011. **30**(12): p. 2454-2464.

126. Locker, N., L.E. Easton, and P.J. Lukavsky, *HCV and CSFV IRES domain II mediate eIF2 release during 80S ribosome assembly.* EMBO J, 2007. **26**(3): p. 795-805.

127. Skabkin, M.A., et al., *Activities of Ligatin and MCT-1/DENR in eukaryotic translation initiation and ribosomal recycling.* Genes Dev, 2010. **24**(16): p. 1787-1801.

128. Terenin, I.M., et al., *Eukaryotic translation initiation machinery can operate in a bacterial-like mode without eIF2.* Nat Struct Mol Biol, 2008. **15**(8): p. 836-841.

129. Lancaster, A.M., E. Jan, and P. Sarnow, *Initiation factor-independent translation mediated by the hepatitis C virus internal ribosome entry site.* RNA, 2006. **12**(5): p. 894-902.

130. Jan, E., T.G. Kinzy, and P. Sarnow, *Divergent tRNA-like element supports initiation, elongation, and termination of protein biosynthesis.* Proc Natl Acad Sci U S A, 2003. **100**(26): p. 15410-15415.

131. Kanamori, Y. and N. Nakashima, *A tertiary structure model of the internal ribosome entry site (IRES) for methionine-independent initiation of translation.* RNA, 2001. **7**(2): p. 266-274.

132. Jan, E. and P. Sarnow, *Factorless ribosome assembly on the internal ribosome entry site of cricket paralysis virus.* J Mol Biol, 2002. **324**(5): p. 889-902.

133. Pfingsten, J.S., D.A. Costantino, and J.S. Kieft, *Structural basis for ribosome recruitment and manipulation by a viral IRES RNA.* Science, 2006. **314**(5804): p. 1450-1454.

134. Colussi, T.M., et al., *Initiation of translation in bacteria by a structured eukaryotic IRES RNA.* Nature, 2015. **519**(7541): p. 110-113.

135. Fernandez, I.S., et al., *Initiation of translation by cricket paralysis virus IRES requires its translocation in the ribosome.* Cell, 2014. **157**(4): p. 823-831.

136. Koh, C.S., et al., *Taura syndrome virus IRES initiates translation by binding its tRNA-mRNA-like structural element in the ribosomal decoding center.* Proc Natl Acad Sci U S A, 2014. **111**(25): p. 9139-9144.

137. Muhs, M., et al., *Cryo-EM of ribosomal 80S complexes with termination factors reveals the translocated cricket paralysis virus IRES.* Mol Cell, 2015. **57**(3): p. 422-432.

138. Schuler, M., et al., *Structure of the ribosome-bound cricket paralysis virus IRES RNA.* Nat Struct Mol Biol, 2006. **13**(12): p. 1092-1096.

139. Spahn, C.M., et al., *Cryo-EM visualization of a viral internal ribosome entry site bound to human ribosomes: the IRES functions as an RNA-based translation factor.* Cell, 2004. **118**(4): p. 465-475.

140. Yamamoto, H., A. Unbehaun, and C.M.T. Spahn, *Ribosomal Chamber Music: Toward an Understanding of IRES Mechanisms.* Trends Biochem Sci, 2017. **42**(8): p. 655-668.

141. Jackson, R.J., *The current status of vertebrate cellular mRNA IRESs.* Cold Spring Harb Perspect Biol, 2013. **5**(2): p. a011569.

142. Komar, A.A. and M. Hatzoglou, *Internal ribosome entry sites in cellular mRNAs: mystery of their existence.* J Biol Chem, 2005. **280**(25): p. 23425-23428.

143. Komar, A.A. and M. Hatzoglou, *Cellular IRES-mediated translation: the war of ITAFs in pathophysiological states.* Cell Cycle, 2011. **10**(2): p. 229-240.

144. Weingarten-Gabbay, S., et al., *Comparative genetics. Systematic discovery of cap-independent translation sequences in human and viral genomes.* Science, 2016. **351**(6270): p. 240.

145. Thakor, N., et al., *Cellular mRNA recruits the ribosome via eIF3-PABP bridge to initiate internal translation.* RNA Biol, 2017. **14**(5): p. 553-567.

146. Kondrashov, N., et al., *Ribosome-mediated specificity in Hox mRNA translation and vertebrate tissue patterning.* Cell, 2011. **145**(3): p. 383-397.

147. Xue, S., et al., *RNA regulons in Hox 5′ UTRs confer ribosome specificity to gene regulation.* Nature, 2015. **517**(7532): p. 33-38.

148. Meyer, K.D., et al., *5′ UTR m(6)A promotes cap-independent translation.* Cell, 2015. **163**(4): p. 999-1010.

149. Sharma, D.K., et al., *Role of Eukaryotic Initiation Factors during Cellular Stress and Cancer Progression.* J Nucleic Acids, 2016. **2016**: p. 8235121.

150. Melnikov, S., et al., *One core, two shells: bacterial and eukaryotic ribosomes.* Nat Struct Mol Biol, 2012. **19**(6): p. 560-567.

151. Telpalo-Carpio, S.A., et al., *Internal ribosome entry site (IRES) from Encephalomyocarditis virus (EMCV) as a tool for shuttle expression plasmids.* Biochem Biophys Res Commun, 2015. **468**(4): p. 548-553.

152. Salis, H.M., *The ribosome binding site calculator.* Methods Enzymol, 2011. **498**: p. 19-42.

153. Chappell, S.A., G.M. Edelman, and V.P. Mauro, *A 9-nt segment of a cellular mRNA can function as an internal ribosome entry site (IRES) and when present in linked multiple copies greatly enhances IRES activity.* Proc Natl Acad Sci U S A, 2000. **97**(4): p. 1536-1541.

154. Benelli, D. and P. Londei, *Translation initiation in Archaea: conserved and domain-specific features.* Biochem Soc Trans, 2011. **39**(1): p. 89-93.

155. Temmel, H., et al., *The RNA ligase RtcB reverses MazF-induced ribosome heterogeneity in Escherichia coli.* Nucleic Acids Res, 2017. **45**(8): p. 4708-4721.

156. Owens, G.C., et al., *Identification of two short internal ribosome entry sites selected from libraries of random oligonucleotides.* Proc Natl Acad Sci U S A, 2001. **98**(4): p. 1471-1476.

157. Venkatesan, A. and A. Dasgupta, *Novel fluorescence-based screen to identify small synthetic internal ribosome entry site elements.* Mol Cell Biol, 2001. **21**(8): p. 2826-2837.

158. Zhou, W., G.M. Edelman, and V.P. Mauro, *Isolation and identification of short nucleotide sequences that affect translation initiation in Saccharomyces cerevisiae.* Proc Natl Acad Sci U S A, 2003. **100**(8): p. 4457-4462.

159. Li, J., et al., *Analysis of IgG heavy chain to light chain ratio with mutant Encephalomyocarditis virus internal ribosome entry site.* Protein Eng Des Sel, 2007. **20**(10): p. 491-496.

160. Koh, E.Y., et al., *An internal ribosome entry site (IRES) mutant library for tuning expression level of multiple genes in mammalian cells.* PLoS One, 2013. **8**(12): p. e82100.

161. Garcia-Martin, J.A., et al., *RNAiFold2T: Constraint Programming design of thermo-IRES switches.* Bioinformatics, 2016. **32**(12): p. 360-368.

162. Ogawa, A., H. Masuoka, and T. Ota, *Artificial OFF-Riboswitches that downregulate internal ribosome entry without hybridization switches in a eukaryotic cell-free translation system.* ACS Synth Biol, 2017. **6**(9): p. 1656-1662.

163. Ogawa, A., *Rational design of artificial riboswitches based on ligand-dependent modulation of internal ribosome entry in wheat germ extract and their applications as label-free biosensors.* RNA, 2011. **17**(3): p. 478-488.

164. Zhao, E.M., et al., *RNA-responsive elements for eukaryotic translational control.* Nat Biotechnol, 2022. **40**(4): p. 539-545.

165. Thompson, S.R., K.D. Gulyas, and P. Sarnow, *Internal initiation in Saccharomyces cerevisiae mediated by an initiator tRNA/eIF2-independent internal ribosome entry site element.* Proc Natl Acad Sci U S A, 2001. **98**(23): p. 12972-12977.

166. Deforges, J., N. Locker, and B. Sargueil, *mRNAs that specifically interact with eukaryotic ribosomal subunits.* Biochimie, 2015. **114**: p. 48-57.

167. Kwan, T. and S.R. Thompson, *Noncanonical Translation Initiation in Eukaryotes.* Cold Spring Harb Perspect Biol, 2019. **11**(4): p. a032672.

168. Wilson, J.E., et al., *Initiation of protein synthesis from the A site of the ribosome.* Cell, 2000. **102**(4): p. 511-520.

169. Sasaki, J. and N. Nakashima, *Methionine-independent initiation of translation in the capsid protein of an insect RNA virus.* Proc Natl Acad Sci U S A, 2000. **97**(4): p. 1512-1515.

170. Muhs, M., et al., *Structural basis for the binding of IRES RNAs to the head of the ribosomal 40S subunit.* Nucleic Acids Res, 2011. **39**(12): p. 5264-5275.

171. Costantino, D.A., et al., *tRNA-mRNA mimicry drives translation initiation from a viral IRES.* Nat Struct Mol Biol, 2008. **15**(1): p. 57-64.

172. Murray, J., et al., *Structural characterization of ribosome recruitment and translocation by type IV IRES.* eLife, 2016. **5**: p. **e13567**.

173. Nakashima, N. and T. Uchiumi, *Functional analysis of structural motifs in dicistroviruses.* Virus Res, 2009. **139**(2): p. 137-147.

174. Boros, Á., et al., *Novel positive-sense, single-stranded RNA (+ssRNA) virus with di-cistronic genome from intestinal content of freshwater carp (Cyprinus carpio).* PLoS One, 2011. **6**(12): p. e29145.

175. Abaeva, I.S., T.V. Pestova, and C.U. Hellen, *Attachment of ribosomal complexes and retrograde scanning during initiation on the Halastavi árva virus IRES.* Nucleic Acids Res, 2016. **44**(5): p. 2362-2377.

176. Abaeva, I.S., et al., *The Halastavi árva virus intergenic region IRES promotes translation by the simplest possible initiation mechanism.* Cell Rep, 2020. **33**(10): p. 108476.

177. Jan, E., *Molecular biology: signals across domains of life.* Nature, 2015. **519**(7541): p. 40-41.

178. Shetty, R.P., D. Endy, and T.F. Knight, Jr., *Engineering BioBrick vectors from BioBrick parts.* J Biol Eng, 2008. **2**: p. 5.

179. Andersen, J.B., et al., *New unstable variants of green fluorescent protein for studies of transient gene expression in bacteria.* Appl Environ Microbiol, 1998. **64**(6): p. 2240-2246.

180. Becker, M., et al., *The 70S ribosome modulates the ATPase activity of Escherichia coli YchF.* RNA Biol, 2012. **9**(10): p. 1288-1301.

181. Petrov, A., et al., *Multiple Parallel Pathways of Translation Initiation on the CrPV IRES.* Mol Cell, 2016. **62**(1): p. 92-103.

182. Heller, J.L.E., *THE RPSA GENE AND ITS PRODUCT RIBOSOMAL PROTEIN S1: INVESTIGATING THE ROLE OF RNA STRUCTURE IN TRANSLATION*, in *Chemistry and Biochemistry*. 2019, University of Lethbridge.

183. Vigar, J.R.J., *RPSA AND RIBOSOMAL PROTEIN S1: INVESTIGATING A NONCANONICAL TRANSLATION INITIATION ELEMENT*, in *Chemistry and Biochemistry*. 2019, University of Lethbridge. p. 170.

184. Ranjbar, B. and P. Gill, *Circular dichroism techniques: biomolecular and nanostructural analyses- a review.* Chem Biol Drug Des, 2009. **74**(2): p. 101-120.

185. Schneider, C.A., W.S. Rasband, and K.W. Eliceiri, *NIH Image to ImageJ: 25 years of image analysis.* Nat Methods, 2012. **9**(7): p. 671-675.

186. *Registry of Standard Biological Parts.* 2003.

187. Vigar, J.R.J. and H.J. Wieden, *Engineering bacterial translation initiation - Do we have all the tools we need?* Biochim Biophys Acta Gen Subj, 2017. **1861**(11 Pt B): p. 3060-3069.

188. Oppenheim, D.S. and C. Yanofsky, *Translational coupling during expression of the tryptophan operon of Escherichia coli.* Genetics, 1980. **95**(4): p. 785-795.

189. Spanjaard, R.A. and J. van Duin, *Translational reinitiation in the presence and absence of a Shine and Dalgarno sequence.* Nucleic Acids Res, 1989. **17**(14): p. 5501-5507.

190. Hertz, M.I. and S.R. Thompson, *In vivo functional analysis of the Dicistroviridae intergenic region internal ribosome entry sites.* Nucleic Acids Res, 2011. **39**(16): p. 7276-7288.

191. Jang, C.J. and E. Jan, *Modular domains of the Dicistroviridae intergenic internal ribosome entry site.* RNA, 2010. **16**(6): p. 1182-1195.

192. Saito, K., R. Green, and A.R. Buskirk, *Translational initiation in E. coli occurs at the correct sites genome-wide in the absence of mRNA-rRNA base-pairing.* eLife, 2020. **9**: p. e55002.

193. de Smit, M.H. and J. van Duin, *Translational standby sites: how ribosomes may deal with the rapid folding kinetics of mRNA.* J Mol Biol, 2003. **331**(4): p. 737-743.

194. Romilly, C., S. Deindl, and E.G.H. Wagner, *The ribosomal protein S1-dependent standby site in tisB mRNA consists of a single-stranded region and a 5′ structure element.* Proc Natl Acad Sci U S A, 2019. **116**(32): p. 15901-15906.

195. Shimizu, Y., et al., *Cell-free translation reconstituted with purified components.* Nat Biotechnol, 2001. **19**(8): p. 751-755.

196. Duval, M., et al., *Escherichia coli ribosomal protein S1 unfolds structured mRNAs onto the ribosome for active translation initiation.* PLoS Biol, 2013. **11**(12): p. e1001731.

197. Lund, P.E., et al., *Protein unties the pseudoknot: S1-mediated unfolding of RNA higher order structure.* Nucleic Acids Res, 2020. **48**(4): p. 2107-2125.

198. Nishiyama, T., et al., *Eukaryotic ribosomal protein RPS25 interacts with the conserved loop region in a dicistroviral intergenic internal ribosome entry site.* Nucleic Acids Res, 2007. **35**(5): p. 1514-1521.

199. Landry, D.M., M.I. Hertz, and S.R. Thompson, *RPS25 is essential for translation initiation by the Dicistroviridae and hepatitis C viral IRESs.* Genes Dev, 2009. **23**(23): p. 2753-2764.

200. Moll, I., et al., *Leaderless mRNAs in bacteria: surprises in ribosomal recruitment and translational control.* Mol Microbiol, 2002. **43**(1): p. 239-246.

201. Hertz, M.I. and S.R. Thompson, *Mechanism of translation initiation by Dicistroviridae IGR IRESs.* Virology, 2011. **411**(2): p. 355-361.

202. Zuker, M., *Mfold web server for nucleic acid folding and hybridization prediction.* Nucleic Acids Res, 2003. **31**(13): p. 3406-3415.

203. Del Campo, C., et al., *Secondary Structure across the Bacterial Transcriptome Reveals Versatile Roles in mRNA Regulation and Function.* PLoS Genet, 2015. **11**(10): p. e1005613.

204. Yusupov, M.M., et al., *Crystal structure of the ribosome at 5.5 A resolution.* Science, 2001. **292**(5518): p. 883-896.

205. Borovinskaya, M.A., et al., *Structural basis for hygromycin B inhibition of protein biosynthesis.* RNA, 2008. **14**(8): p. 1590-1599.

206. Nishiyama, T., et al., *Structural elements in the internal ribosome entry site of Plautia stali intestine virus responsible for binding with ribosomes.* Nucleic Acids Res, 2003. **31**(9): p. 2434-2442.

207. Sørensen, M.A., J. Fricke, and S. Pedersen, *Ribosomal protein S1 is required for translation of most, if not all, natural mRNAs in Escherichia coli in vivo.* J Mol Biol, 1998. **280**(4): p. 561-569.

208. Sengupta, J., R.K. Agrawal, and J. Frank, *Visualization of protein S1 within the 30S ribosomal subunit and its interaction with messenger RNA.* Proc Natl Acad Sci U S A, 2001. **98**(21): p. 11991-11996.

209. Boni, I.V., et al., *Ribosome-messenger recognition: mRNA target sites for ribosomal protein S1.* Nucleic Acids Res, 1991. **19**(1): p. 155-162.

210. Tzareva, N.V., V.I. Makhno, and I.V. Boni, *Ribosome-messenger recognition in the absence of the Shine-Dalgarno interactions.* FEBS Lett, 1994. **337**(2): p. 189-194.

211. Chang, B., S. Halgamuge, and S.L. Tang, *Analysis of SD sequences in completed microbial genomes: non-SD-led genes are as common as SD-led genes.* Gene, 2006. **373**: p. 90-99.

212. Ringquist, S., et al., *High-affinity RNA ligands to Escherichia coli ribosomes and ribosomal protein S1: comparison of natural and unnatural binding sites.* Biochemistry, 1995. **34**(11): p. 3640-3648.

213. Diaconu, M., et al., *Structural basis for the function of the ribosomal L7/12 stalk in factor binding and GTPase activation.* Cell, 2005. **121**(7): p. 991-1004.

214. Sheahan, T. and H.J. Wieden, *Ribosomal Protein S1 Improves the Protein Yield of an In Vitro Reconstituted Cell-Free Translation System.* ACS Synth Biol, 2022. **11**(2): p. 1004-1008.

215. Skorski, P., et al., *The highly efficient translation initiation region from the Escherichia coli rpsA gene lacks a shine-dalgarno element.* J Bacteriol, 2006(0021-9193 (Print)): p. 6277-6285.

216. Schurr, T., E. Nadir, and H. Margalit, *Identification and characterization of E.coli ribosomal binding sites by free energy computation.* Nucleic Acids Res, 1993. **21**(17): p. 4019-4023.

217. Chen, H., et al., *Determination of the optimal aligned spacing between the Shine-Dalgarno sequence and the translation initiation codon of Escherichia coli mRNAs.* Nucleic Acids Res, 1994. **22**(23): p. 4953-4957.

218. Ringquist, S., et al., *Translation initiation in Escherichia coli: sequences within the ribosome-binding site.* Mol Microbiol, 1992. **6**(9): p. 1219-1229.

219. de Smit, M.H. and J. van Duin, *Secondary structure of the ribosome binding site determines translational efficiency: a quantitative analysis.* Proc Natl Acad Sci U S A, 1990. **87**(19): p. 7668-7672.

220. Schauder, B. and J.E. McCarthy, *The role of bases upstream of the Shine-Dalgarno region and in the coding sequence in the control of gene expression in Escherichia coli: translation and stability of mRNAs in vivo.* Gene, 1989. **78**(1): p. 59-72.

221. Studer, S.M. and S. Joseph, *Unfolding of mRNA Secondary Structure by the Bacterial Translation Initiation Complex.* Mol Cell, 2006. **22**(1): p. 105-115.

222. Cetnar, D.P. and H.M. Salis, *Systematic Quantification of Sequence and Structural Determinants Controlling mRNA stability in Bacterial Operons.* ACS Synth Biol, 2021. **10**(2): p. 318-332.

223. Tian, T. and H.M. Salis, *A predictive biophysical model of translational coupling to coordinate and control protein expression in bacterial operons.* Nucleic Acids Res, 2015. **43**(14): p. 7137-7151.

224. Li, G.W., et al., *Quantifying absolute protein synthesis rates reveals principles underlying allocation of cellular resources.* Cell, 2014. **157**(3): p. 624-635.

225. Green, Alexander A., et al., *Toehold Switches: De-Novo-Designed Regulators of Gene Expression.* Cell, 2014. **159**(4): p. 925-939.

226. Breaker, R.R., *Riboswitches and the RNA world.* Cold Spring Harb Perspect Biol, 2012. **4**(2): p. a003566.

227. Dixon, N., et al., *Reengineering orthogonally selective riboswitches.* Proc Natl Acad Sci U S A, 2010. **107**(7): p. 2830-2835.

228. Lynch, S.A., et al., *A high-throughput screen for synthetic riboswitches reveals mechanistic insights into their function.* Chem Biol, 2007. **14**(2): p. 173-184.

229. Salah, P., et al., *Probing the relationship between Gram-negative and Gram-positive S1 proteins by sequence analysis.* Nucleic Acids Res, 2009. **37**(16): p. 5578-5588.

230. Subramanian, A.R., *Structure and functions of ribosomal protein S1.* Prog Nucleic Acid Res Mol Biol, 1983. **28**: p. 101-142.

231. Byrgazov, K., et al., *Structural basis for the interaction of protein S1 with the Escherichia coli ribosome.* Nucleic Acids Res, 2015. **43**(1): p. 661-673.

232. Kitakawa, M. and K. Isono, *An amber mutation in the gene rpsA for ribosomal protein S1 in Escherichia coli.* Mol Gen Genet, 1982. **185**(3): p. 445-447.

233. Boni, I.V., V.S. Artamonova, and M. Dreyfus, *The last RNA-binding repeat of the Escherichia coli ribosomal protein S1 is specifically involved in autogenous control.* J Bacteriol, 2000. **182**(20): p. 5872-5879.

234. Bear, D.G., et al., *Alteration of polynucleotide secondary structure by ribosomal protein S1.* Proc Natl Acad Sci U S A, 1976. **73**(6): p. 1824-1828.

235. Kolb, A., et al., *Nucleic acid helix-unwinding properties of ribosomal protein S1 and the role of S1 in mRNA binding to ribosomes.* Proc Natl Acad Sci U S A, 1977. **74**(6): p. 2379-2383.

236. Qu, X., et al., *Ribosomal protein S1 unwinds double-stranded RNA in multiple steps.* Proc Natl Acad Sci U S A, 2012. **109**(36): p. 14458-14463.

237. Potapov, A.P. and A.R. Subramanian, *Effect of E. coli ribosomal protein S1 on the fidelity of the translational elongation step: reading and misreading of poly(U) and poly(dT).* Biochem Int, 1992. **27**(4): p. 745-753.

238. Sukhodolets, M.V. and S. Garges, *Interaction of Escherichia coli RNA polymerase with the ribosomal protein S1 and the Sm-like ATPase Hfq.* Biochemistry, 2003. **42**(26): p. 8022-8034.

239. Demo, G., et al., *Structure of RNA polymerase bound to ribosomal 30S subunit.* eLife, 2017. **6**: p. e28560.

240. Blumenthal, T., T.A. Landers, and K. Weber, *Bacteriophage Q replicase contains the protein biosynthesis elongation factors EF Tu and EF Ts.* Proc Natl Acad Sci U S A, 1972. **69**(5): p. 1313-1317.

241. Fan, Y., et al., *Structural basis for ribosome protein S1 interaction with RNA in trans-translation of Mycobacterium tuberculosis.* Biochem Biophys Res Commun, 2017. **487**(2): p. 268-273.

242. Vallejos-Sánchez, K., et al., *Mycobacterium tuberculosis ribosomal protein S1 (RpsA) and variants with truncated C-terminal end show absence of interaction with pyrazinoic acid.* Sci Rep, 2020. **10**(1): p. 8356.

243. Kim, D., et al., *Comparative analysis of regulatory elements between Escherichia coli and Klebsiella pneumoniae by genome-wide transcription start site profiling.* PLoS Genet, 2012. **8**(8): p. e1002867.

244. Boni, I.V., et al., *Non-canonical mechanism for translational control in bacteria: synthesis of ribosomal protein S1.* EMBO J, 2001. **20**(15): p. 4222-4232.

245. Tchufistova, L.S., A.V. Komarova, and I.V. Boni, *A key role for the mRNA leader structure in translational control of ribosomal protein S1 synthesis in γ-proteobacteria.* Nucleic Acids Res, 2003. **31**(23): p. 6996-7002.

246. Mustoe, A.M., et al., *Pervasive Regulatory Functions of mRNA Structure Revealed by High-Resolution SHAPE Probing.* Cell, 2018. **173**(1): p. 181-195.

247. D'Souza, M.H., et al., *Biophysical characterisation of human LincRNA-p21 sense and antisense Alu inverted repeats.* Nucleic Acids Res, 2022. **50**(10): p. 5881-5898.

248. Keffer-Wilkes, L.C., G.R. Veerareddygari, and U. Kothe, *RNA modification enzyme TruB is a tRNA chaperone.* Proc Natl Acad Sci U S A, 2016. **113**(50): p. 14306-14311.

Appendix 1.

Supplemental Material to Chapter 2

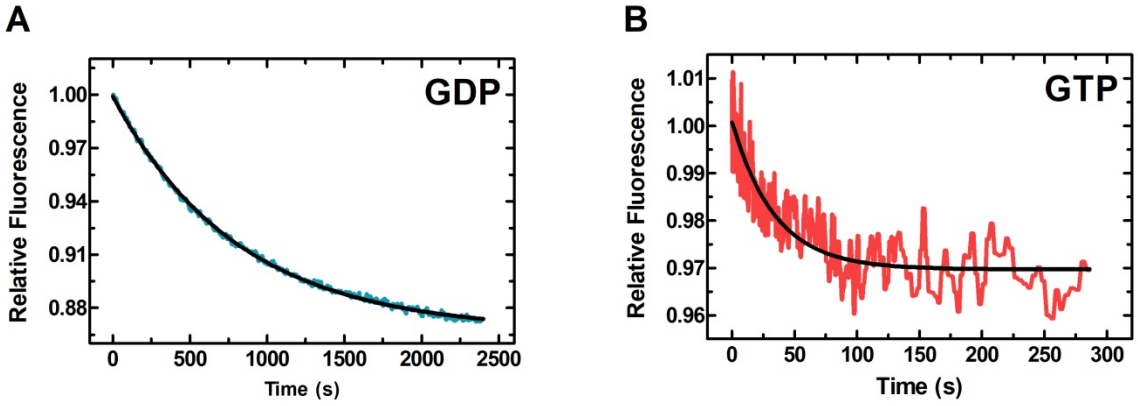

Figure S2.1. EF-Tu mutations T34C and L265C do not affect nucleotide. Rapid mixing of MANT-GDP•EF-Tu T34C/L265C (A) and MANT-GTP●EF-Tu T34C/L265C (B) against excess unlabelled nucleotide. Average of three traces are plotted (blue and red) and a single exponential decay function is fit to the data (black line).

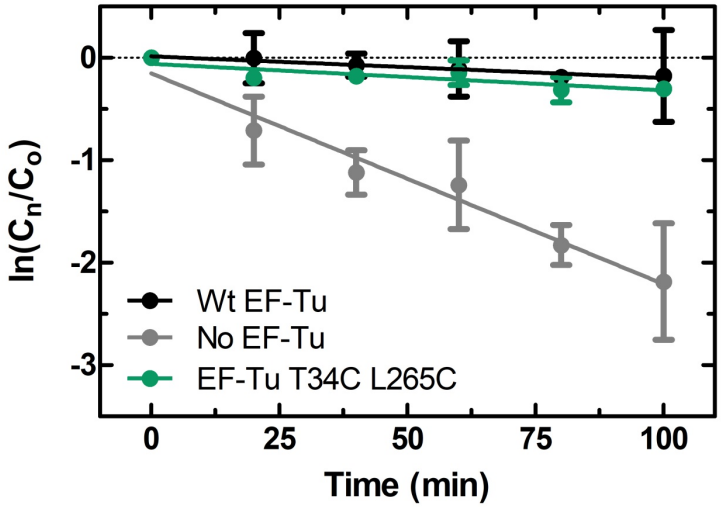

Figure S2.2. EF-Tu mutations T34C and L265C do not hinder ternary complex formation. Hydrolysis of [^{14}C]Phe-tRNA^{Phe} incubated in the presence or absence of EF-Tu. The slope indicates the rate of aminoacyl-ester bond hydrolysis. C_n is the concentration of [^{14}C]Phe-tRNA^{Phe} at any time point and C_o is the initial concentration of [^{14}C]Phe-tRNA^{Phe}. Mean of three replicates is plotted, error bars indicate one standard deviation.

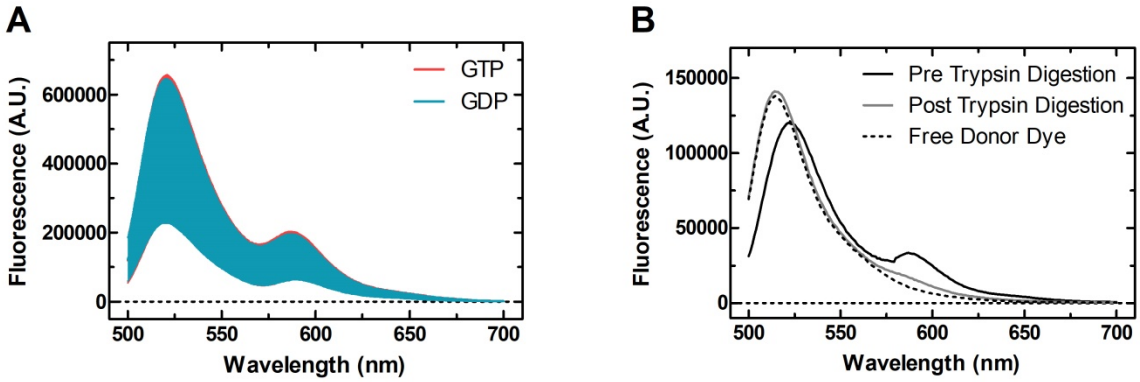

Figure S2.3. Intramolecular hetero-FRET is occurring in EF-Tu^{FR}. Fluorescence emission scans of double-labeled EF-Tu^{FR} upon excitation at 493 nm. (A) Samples were incubated with either a 10-fold excess of GTP (red) or GDP (blue) prior to measurement, filled area denotes one standard deviation. (B) EF-Tu^{FR} before Trypsin digestion (black line), post trypsin digestion (grey line) and free donor dye (fluorescein, dashed line).

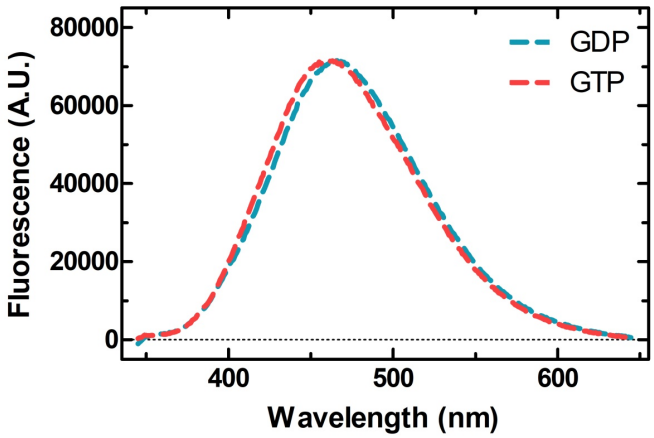

Figure S2.4. Trypsin digestion confirms intramolecular hetero-FRET in EF-Tu^{ID}. Fluorescence emission scans of double-labeled EF-Tu^{ID} upon excitation at 336 nm. Samples were incubated with either a 10-fold excess of GTP (red) or GDP (blue). Samples were then digested with trypsin overnight before being measured.

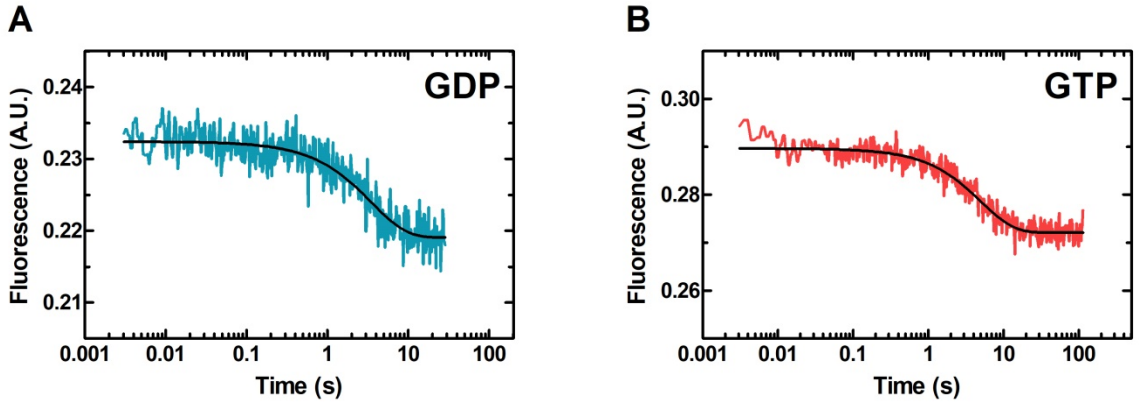

Figure S2.5. EDTA accelerates nucleotide dissociation from EF-Tu^{Da}. Rapid mixing of (A) MANT-GDP●EF-Tu^{Da} and (B) MANT-GTP●EF-Tu^{Da} complexes with 10 mM EDTA. Example traces are plotted (blue and red) and a single exponential decay function is fit to the data (black line).

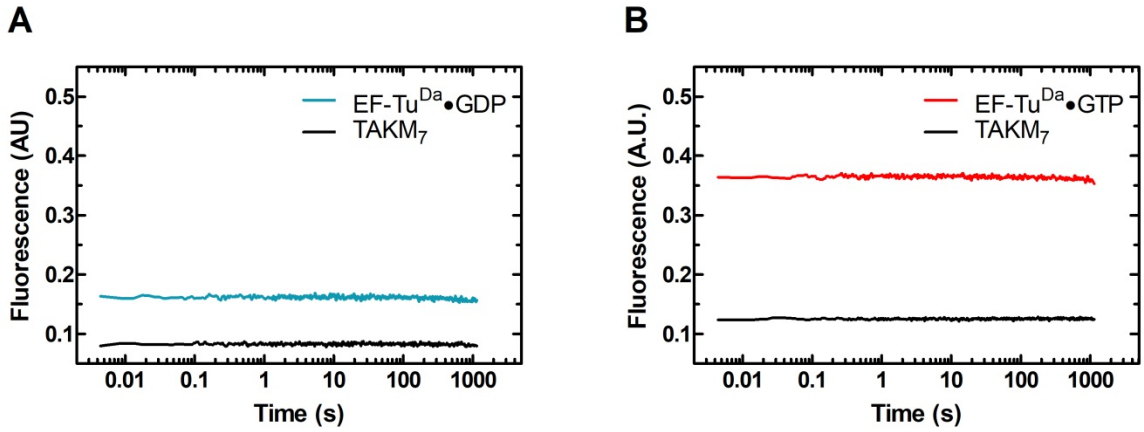

Figure S2.6. EF-Tu^{Da} signal change over time is not due to dilution or rapid mixing. Rapidly mixing (A) EF-Tu^{Da}●GDP or (B) EF-Tu^{Da}●GTP with themselves or TAKM$_7$ buffer shows no change in EF-Tu^{Da} fluorescence.

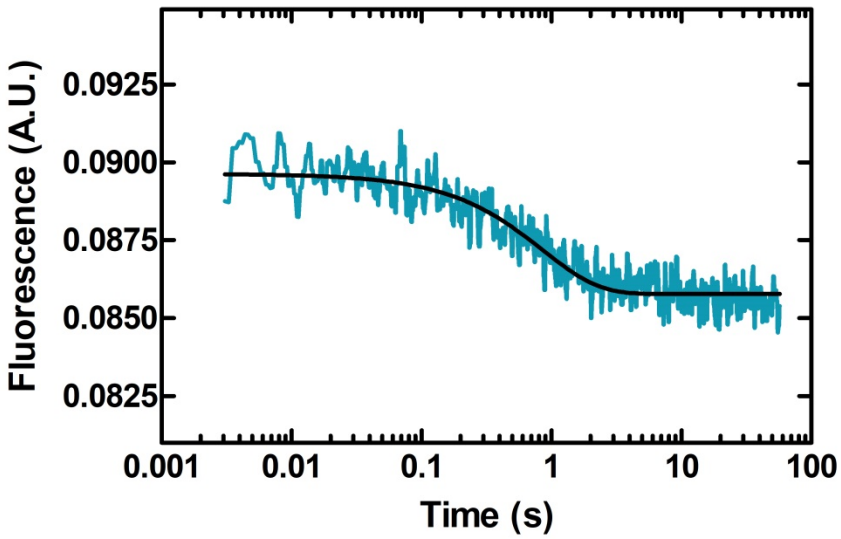

Figure S2.7. EF-Ts stimulates EF-Tu^{Da} nucleotide release as efficiently as wild type EF-Tu. Rapidly mixing EF-Tu^{Da}•MANT-GDP (0.15 μM) with EF-Ts (0.1 μM). Example trace is plotted (blue) and a single exponential decay function is fit to the data (black line).

Appendix 2.

Supplemental Material to Chapter 4

Reprinted with permission from “The prokaryotic activity of the IGR IRESs is mediated by ribosomal protein S1”

Roberts L., Wieden H.-J. (2022). The prokaryotic activity of the IGR IRESs is mediated by ribosomal protein S1. Nucleic Acids Research, Volume 50, Issue 16, Pages 9355-9367.

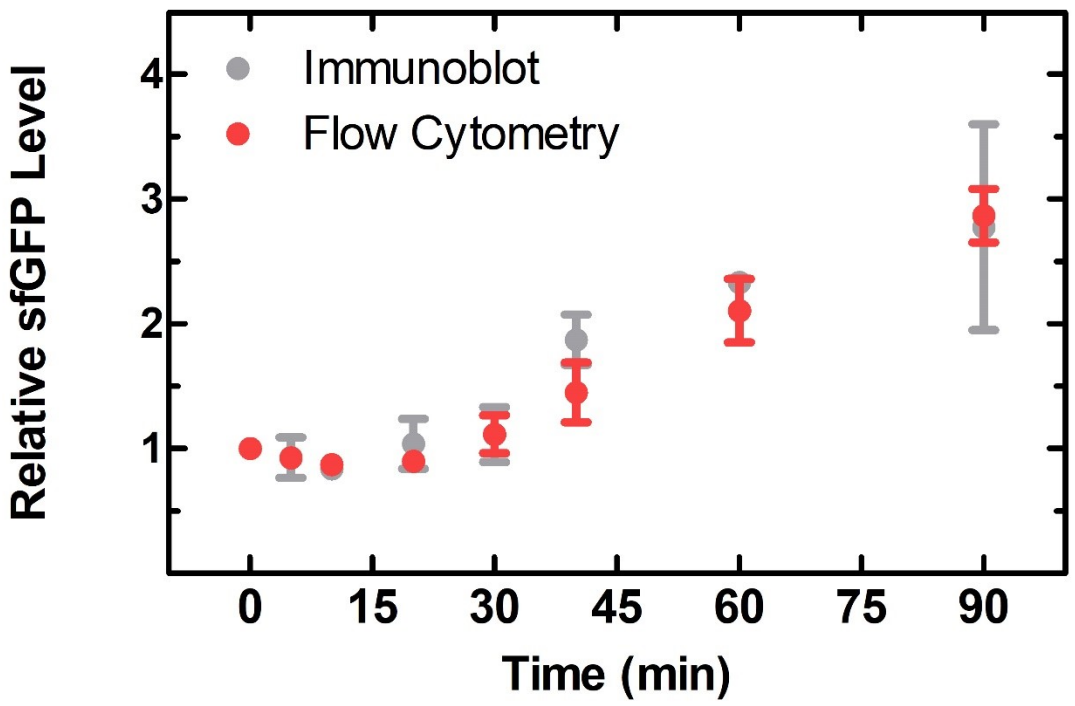

Figure S4.1. Flow cytometry accurately reports the amount of sfGFP protein *in vivo*. Relative fluorescence time course of *E. coli* containing PSIV IGR IRES constructs. Comparison between PK2_K/O flow cytometry data (fluorescence, red) and PK2_K/O immunoblot data (protein level, grey). Mean values of three biological replicates are plotted; error bars indicate one standard deviation.

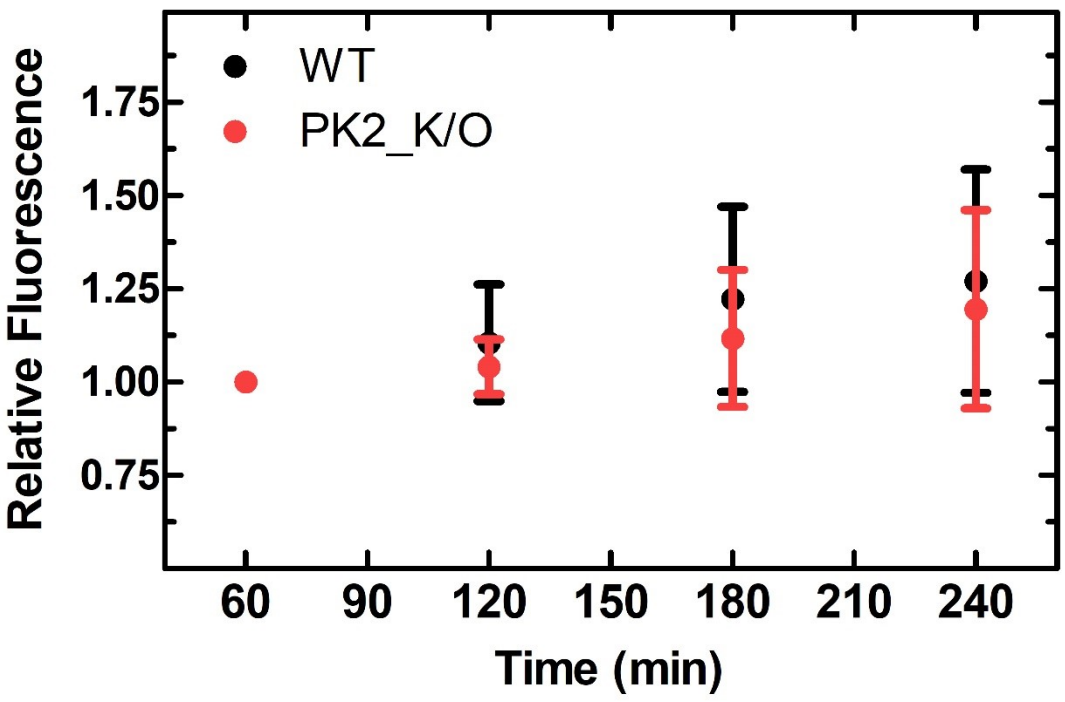

Figure S4.2. **Superfolder green fluorescent protein (sfGFP) is stable over multiple hours *in vivo***. Relative fluorescence time course of *E. coli* containing PSIV IGR IRES constructs post shift to minimal media measured by flow cytometry. Mean values of three biological replicates are plotted; error bars indicate one standard deviation.

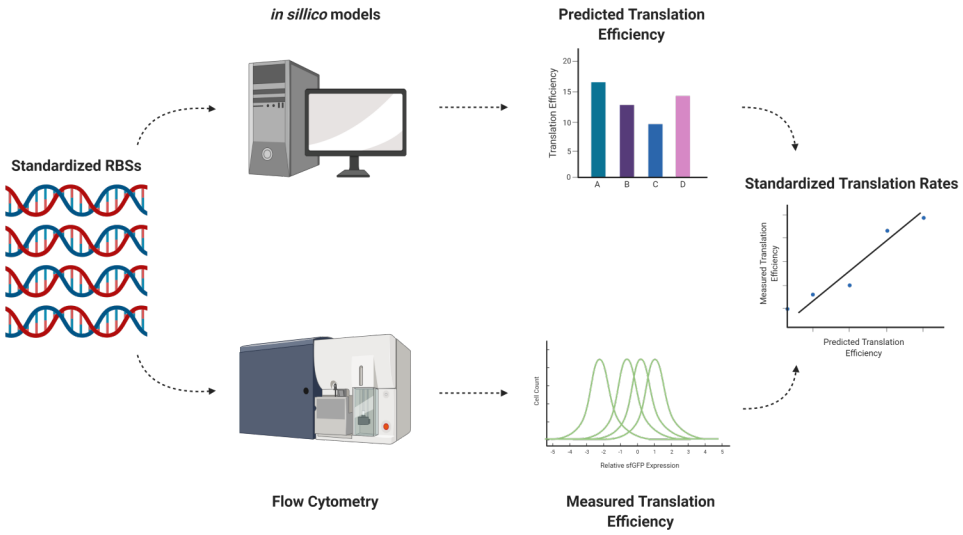

Figure S4.3. Novel real-time fluorescence-based single-cell translation assay for benchmarking IRES translational efficiency *in vivo*. Using standardized ribosome binding sites, the translation efficiency of foreign translation elements can be benchmarked against canonical bacterial translation.

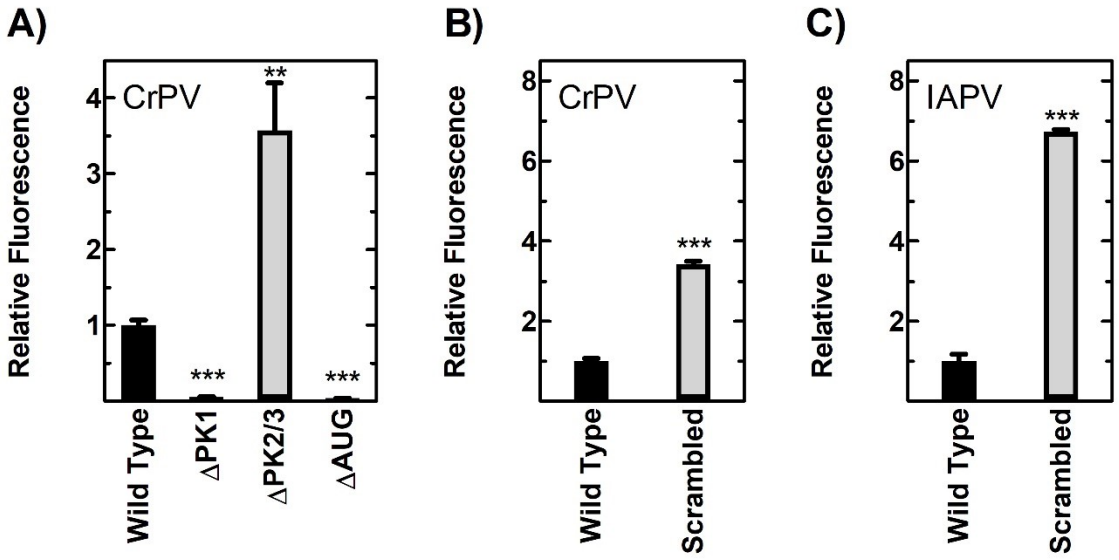

Figure S4.4. Pseudoknot deletion and Scrambled IGR IRES variants have translation efficiency inconsistent with a structure-based mechanism. Relative fluorescence of *E. coli* containing IGR IRES constructs with pseudoknot deletions (A), or scrambled IRES sequences (B) and (C). Translation efficiency measured by flow cytometry, mean values of three biological replicates are plotted relative to the respective WT IRES; error bars indicate one standard deviation. Constructs with statistically significant differences from WT are indicated (* = $P < 0.05$, ** = $P < 0.01$, *** = $P < 0.001$, **** = $P < 0.0001$).

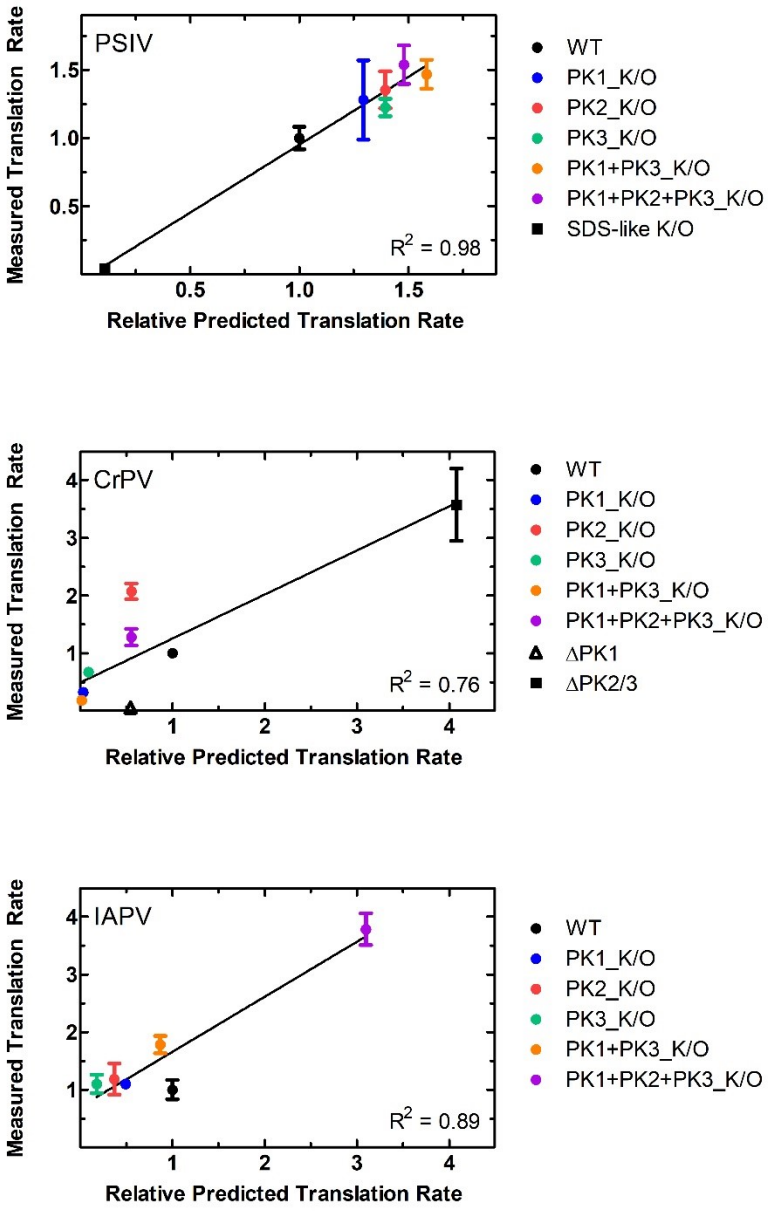

Figure S4.5. Correlation of predicted and measured translation efficiency of IGR IRES constructs. Translation efficiency predicted using the Salis lab RBS calculator, translation efficiency measured by flow cytometry. Mean values of three biological replicates are plotted; error bars indicate one standard deviation.

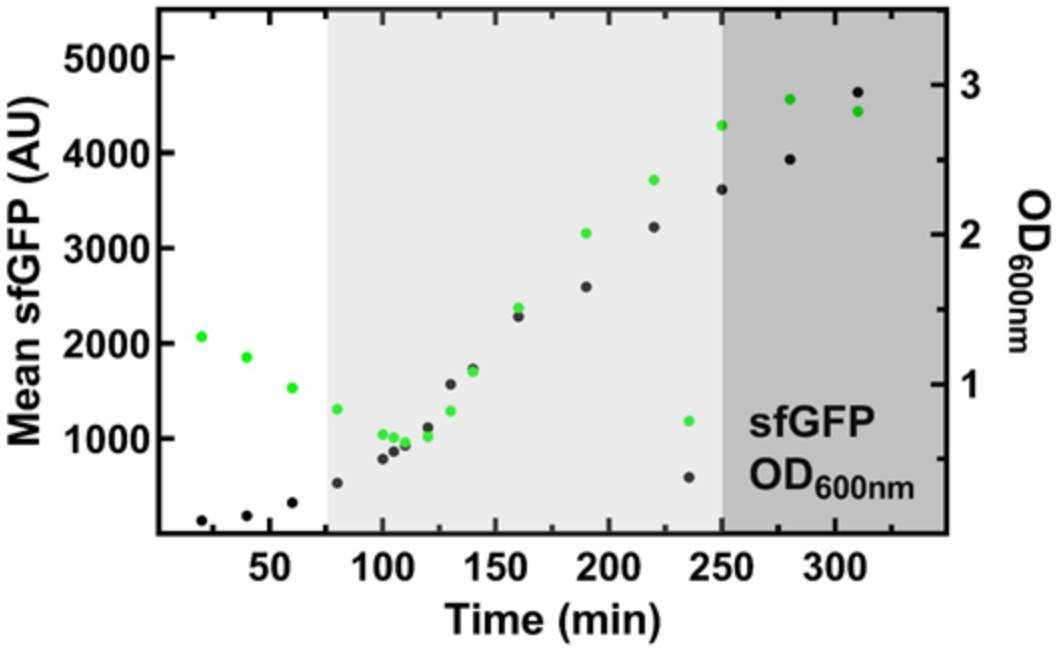

Figure S4.6. Time course of sfGFP expression over several phases of cell growth. Relative fluorescence *in vivo* time course (green) of *E. coli* containing the WT PSIV IGR IRES construct as measured by flow cytometry compared to the optical density at 600 nm (black). Growth phases are indicated by the white (lag), light grey (exponential), and dark grey (early stationary) shading. Fluorescence decreased over the first 120 min as cells divide and dilute the sfGFP before expression is induced at 120 min (0.6 OD 600 nm) with IPTG.

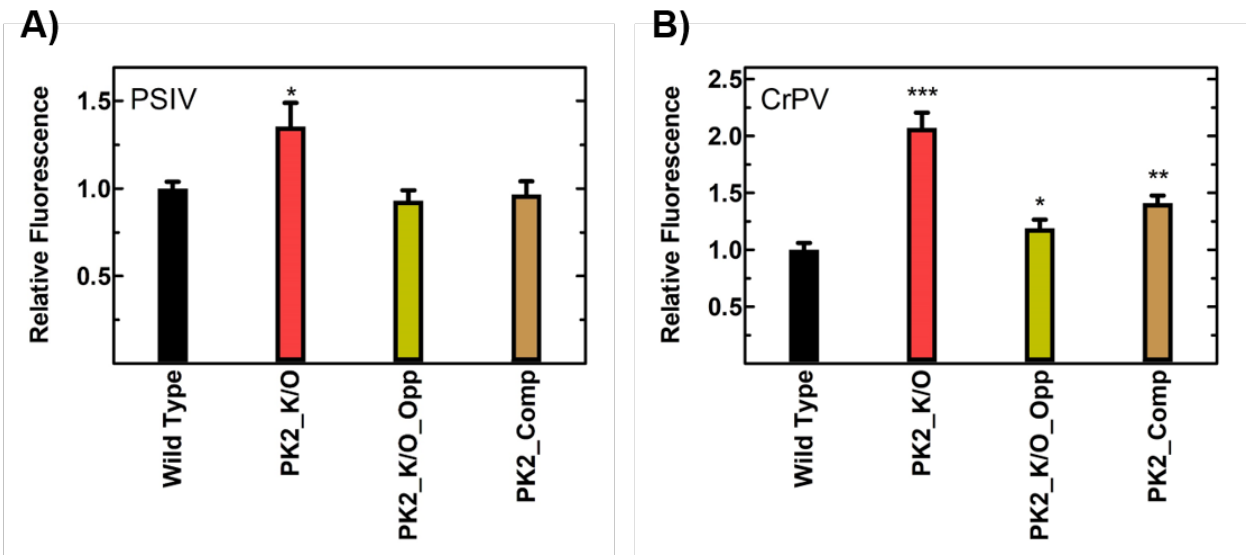

Figure S4.7. Compensatory mutation reduces PK2_K/O IRES translation efficiency. PSIV (A) and CrPV (B) IGR IRES translation efficiency. Mean fluorescent values of three biological replicates obtained by flow cytometry are plotted relative to the respective WT IRES; error bars indicate one standard deviation. Constructs with statistically significant differences from WT are indicated (* = $P < 0.05$)

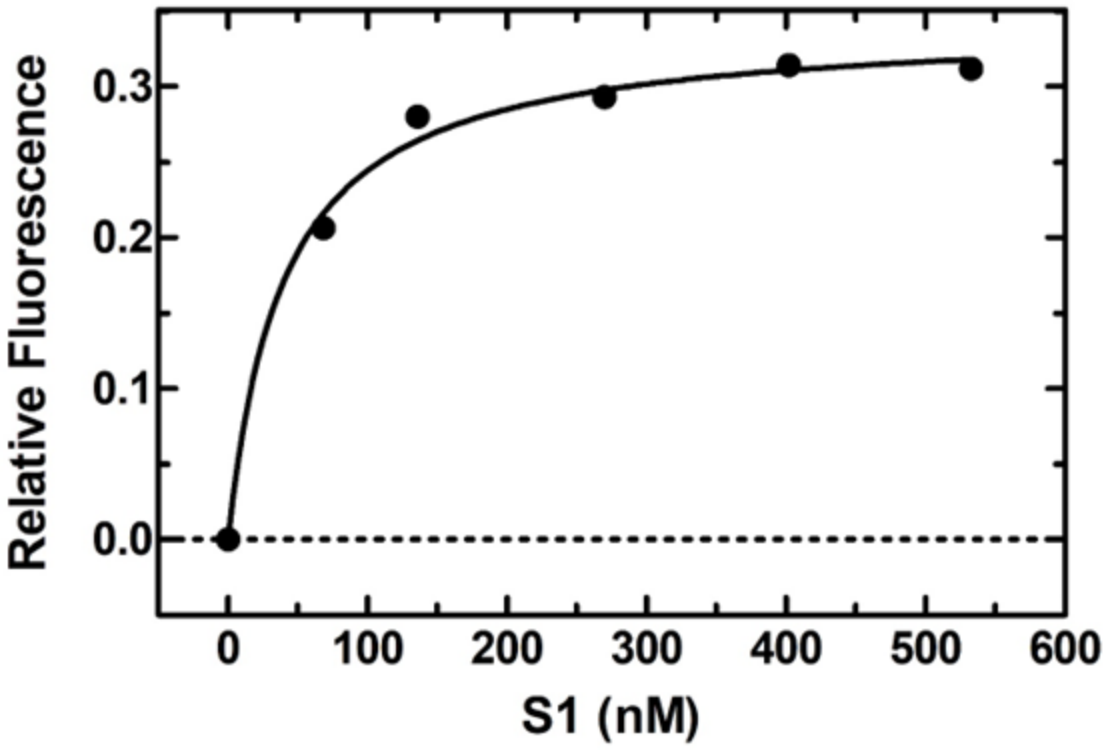

Figure S4.8. **The CrPV IGR IRES binds the ribosomal protein S1 with nanomolar affinity.** Titration of fluorescently labelled (pyrene) WT CrPV IGR IRES with ribosomal protein S1. Fluorescently labeled WT CrPV IGR IRES RNA was incubated with increasing amounts of ribosomal protein S1. Relative fluorescence emission at 391 nm is shown (λ_{ex} = 341 nm).

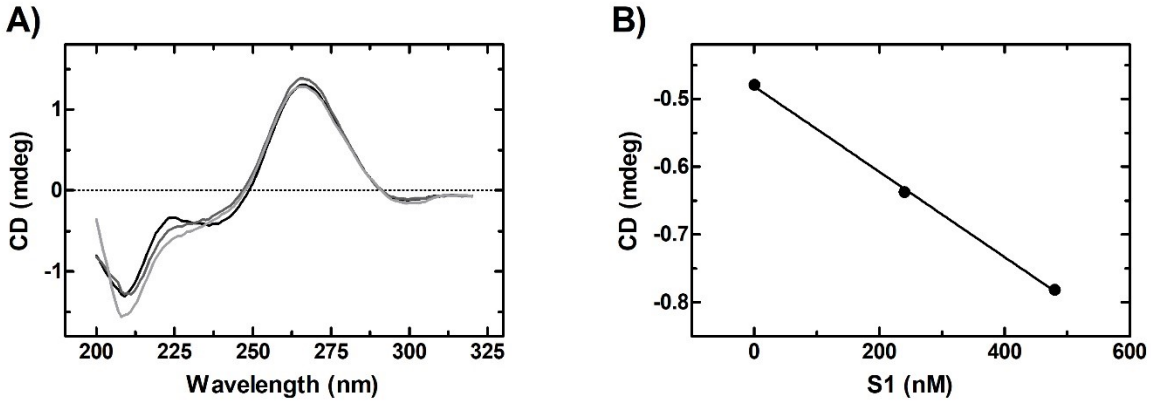

Figure S4.9. **Ribosomal protein S1 alters IRES RNA structure.** (A) Titration of WT CrPV IGR IRES with ribosomal protein S1 while measuring RNA structure via circular dichroism spectroscopy. Wild type CrPV IGR IRES RNA was incubated with increasing amounts of ribosomal protein S1, black trace (0 nM S1), dark grey trace (240 nM S1), light grey trace (480 nM S1). (B) Change in helical RNA structure at 220 nm is plotted against S1 concentration.

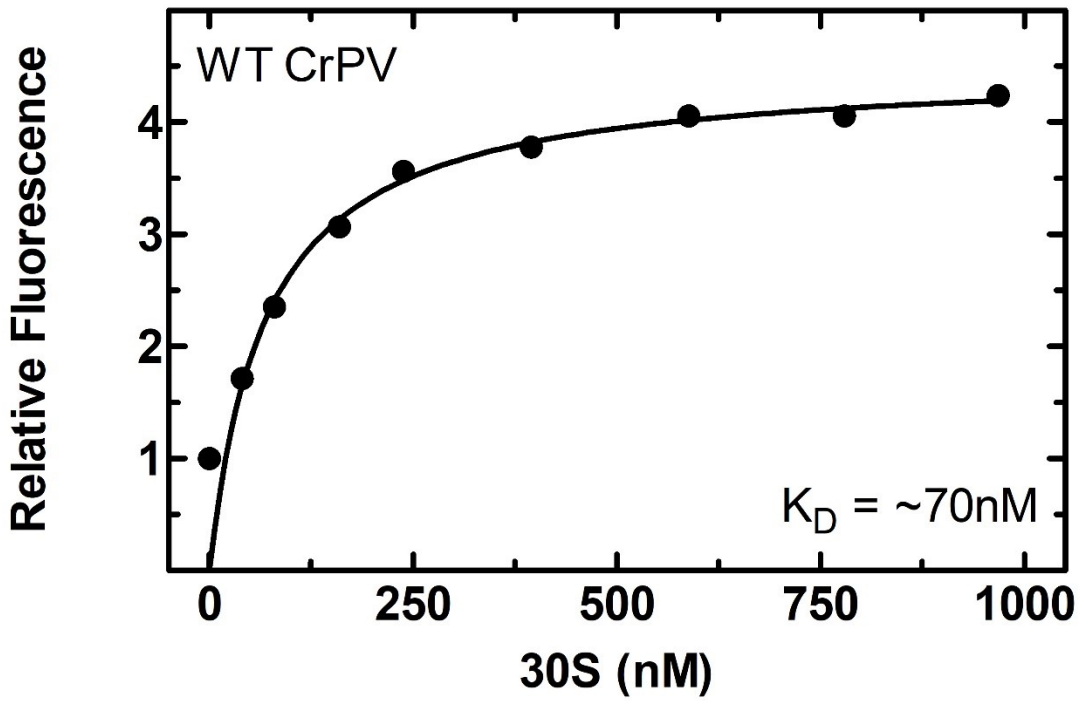

Figure S4.10. **The CrPV IGR IRES binds the 30S subunit with a nanomolar affinity.** Titration of fluorescently labelled (pyrene) WT CrPV IGR IRES with the 30S ribosomal subunit. Fluorescently labeled WT CrPV IGR IRES RNA was incubated with increasing amounts of 30S subunits. Relative fluorescence emission at 391 nm is shown (λ_{ex} = 341 nm).

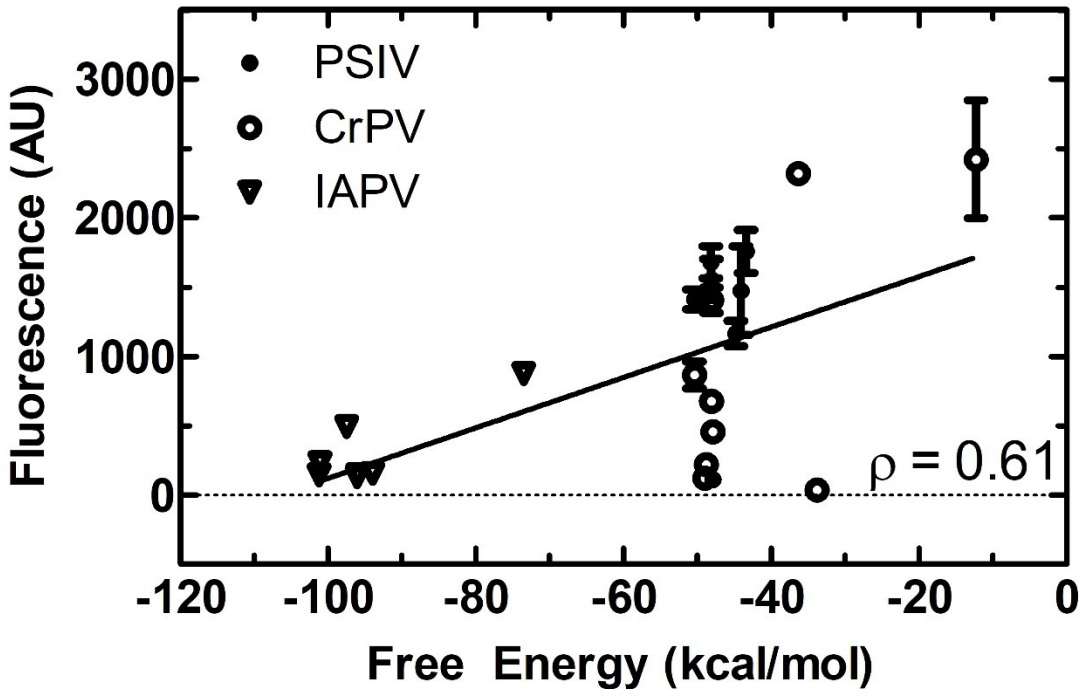

Figure S4.11. Predicted free energy and translation efficiency of IGR IRES constructs are correlated. Free energy predicted using mFOLD, translation efficiency measured by flow cytometry. Mean values of three biological replicates are plotted, error bars indicate one standard deviation.

Table S4.1. DNA sequences of monocistronic and bicistronic* constructs.

CrPV	Sequence Upstream from sfGFP
WT	5'- AAAGCAAAAATGTGATCTTGCTTGTAAATACAATTTTGAGAG GTTAATAAATTACAAGTAGTGCTATTTTTGTATTTAGGTTAGC TATTTAGCTTTACGTTCCAGGATGCCTAGTGGCAGCCCCACA ATATCCAGGAAGCCCTCTCTGCGGTTTTTCAGATTAGGTAGT CGAAAAACCTAAGAAATTTACCTGCTACATTTCAAGATAAA- 3'
PK1_K/O	5'- AAAGCAAAAATGTGATCTTGCTTGTAAATACAATTTTGAGAG GTTAATAAATTACAAGTAGTGCTATTTTTGTATTTAGGTTAGC TATTTAGCTTTACGTTCCAGGATGCCTAGTGGCAGCCCCACA ATATCCAGGAAGCCCTCTCTGCGGTTTTTCAGATTAGGTAGT CGAAAAACCTAAGAAATTTAGGTGCTACATTTCAAGATAAA- 3'
PK2_K/O	5'- AAAGCAAAAATGTGATCTTGCTTGTAAATACAATTTTGAGAG GTTAATAAATTACAAGTAGTGCTATTTTTGTATTTAGGTTAGC TATTTAGCTTTACGTTCCAGGATGCCTAGTGGCAGCCCCACA ATATCCAGGAAGCGGAGAGTGCGGTTTTTCAGATTAGGTAGT

	CGAAAAACCTAAGAAATTTACCTGCTACATTTCAAGATAAA-3'
PK3_K/O	5'-AAAGCAAAAATGTGATCTTGCTTGTAAATACAATTTTGAGAGGTTAATAAATTACAAGTAGTGCTATTTTTGTATTTAGGTTAGCTATTTAGCTTTACGTTCCAGGATGCCTAGTGGCAGCCCGTGAATATCCAGGAAGCCCTCTCTGCGGTTTTTCAGATTAGGTAGTCGAAAAACCTAAGAAATTTACCTGCTACATTTCAAGATAAA-3'
PK1+PK3_K/O	5'-AAAGCAAAAATGTGATCTTGCTTGTAAATACAATTTTGAGAGGTTAATAAATTACAAGTAGTGCTATTTTTGTATTTAGGTTAGCTATTTAGCTTTACGTTCCAGGATGCCTAGTGGCAGCCGGTGAATATCCAGGAAGCCCTCTCTGCGGTTTTTCAGATTAGGTAGTCGAAAAACCTAAGAAATTTAGGTGCTACATTTCAAGATAAA
PK1+PK2+PK3_K/O	5'-AAAGCAAAAATGTGATCTTGCTTGTAAATACAATTTTGAGAGGTTAATAAATTACAAGTAGTGCTATTTTTGTATTTAGGTTAGCTATTTAGCTTTACGTTCCAGGATGCCTAGTGGCAGCCGGTGAATATCCAGGAAGCGGAGAGTGCGGTTTTTCAGATTAGGTAGT

	CGAAAAACCTAAGAAATTTAGGTGCTACATTTCAAGATAAA-3'
PK1 Deletion	5'-AAAGCAAAAATGTGATCTTGCTTGTAAATACAATTTTGAGAGGTTAATAAATTACAAGTAGTGCTATTTTTGTATTTAGGTTAGCTATTTAGCTTTACGTTCCAGGATGCCTAGTGGCAGCCCCACAATATCCAGGAAGCCCTCTCTGCCTACATTTCAAGATAAA-3'
PK2+PK3 Deletion	5'-GCGGTTTTTCAGATTAGGTAGTCGAAAAACCTAAGAAATTTACCTGCTACATTTCAAGATAAA-3'
Scrambled	5'-GTAAGATGTTTCATGACCAACCCTAAATCTAAGTATATAAGTTGCACTTGAAAATGGTAATTCTTAATAAGCTGGATGAGTAACGAGTATTATTATCGAACGGAATTTTCCTGGTGAAACACTCTTTCATGACACGGAAATAAGGTGCGTCTTCTTTCAATATATCTCACGATGGCTTGAACAGACGGTTTCATAACTTTCTTAAATGCG-3'
PK2_Opp	5'-AAAGCAAAAATGTGATCTTGCTTGTAAATACAATTTTCTCTCCTTAATAAATTACAAGTAGTGCTATTTTTGTATTTAGGTTAGCTATTTAGCTTTACGTTCCAGGATGCCTAGTGGCAGCCCCACA

	ATATCCAGGAAGCCCTCTCTGCGGTTTTTCAGATTAGGTAGT CGAAAAACCTAAGAAATTTACCTGCTACATTTCAAGATAAA- 3'
PK2_Comp	5'- AAAGCAAAAATGTGATCTTGCTTGTAAATACAATTTTCTCTC CTTAATAAATTACAAGTAGTGCTATTTTTGTATTTAGGTTAGC TATTTAGCTTTACGTTCCAGGATGCCTAGTGGCAGCCCCACA ATATCCAGGAAGCGGAGAGTGCGGTTTTTCAGATTAGGTAGT CGAAAAACCTAAGAAATTTACCTGCTACATTTCAAGATAAA- 3'
IAPV	Sequence Upstream from sfGFP
WT	5'- GAGCGGTTTCTGGAATACTATATGTAAGTATAGTGTTCTGGA GGCATCATTCTATGGTTACCCATCATTAGAGGAAATTTCCAA TAAACTCTGGTGTAAGGCTTAGAGTGATGGTCGAGGTGCCCT ATTTAGGGTGAGGAGCCTCGGTGGCAGCCCCACCAAATCCTC TATTGGATAGGAACAGCTGTACTGGGCAGTTACAGCAGTCGT ATGGTAACACATGCGGCGTTCCGAAATACCATGCCTGGCGAT TCACAACAAGAA-3'
PK1_K/O	5'- GAGCGGTTTCTGGAATACTATATGTAAGTATAGTGTTCTGGA

	GGCATCATTCTATGGTTACCCATCATTAGAGGAAATTTCCAA TAAACTCTGGTGTAAGGCTTAGAGTGATGGTCGAGGTGCCCT ATTTAGGGTGAGGAGCCTCGGTGGCAGCCCCACCAAATCCTC TATTGGATAGGAACAGCTGTACTGGGCAGTTACAGCAGTCGT ATGGTAACACATGCGGCGTTCCGAAATACCATGGGTGGCGAT TCACAACAAGAA-3'
PK2_K/O	5'- GAGCGGTTTCTGGAATACTATATGTAAGTATAGTGTTCTGGA GGCATCATTCTATGGTTACCCATCATTAGAGGAAATTTCCAA TAAACTCTGGTGTAAGGCTTAGAGTGATGGTCGAGGTGCCCT ATTTAGGGTGAGGAGCCTCGGTGGCAGCCCCACCAAATCCTC TTAACCATAGGAACAGCTGTACTGGGCAGTTACAGCAGTCGT ATGGTAACACATGCGGCGTTCCGAAATACCATGCCTGGCGAT TCACAACAAGAA-3'
PK3_K/O	5'- GAGCGGTTTCTGGAATACTATATGTAAGTATAGTGTTCTGGA GGCATCATTCTATGGTTACCCATCATTAGAGGAAATTTCCAA TAAACTCTGGTGTAAGGCTTAGAGTGATGGTCGAGGTGCCCT ATTTAGGGTGAGGAGCCTCGGTGGCAGCCCCTGGAAATCCTC TATTGGATAGGAACAGCTGTACTGGGCAGTTACAGCAGTCGT ATGGTAACACATGCGGCGTTCCGAAATACCATGCCTGGCGAT TCACAACAAGAA-3'

PK1+PK3_K/O	5'- GAGCGGTTTCTGGAATACTATATGTAAGTATAGTGTTCTGGA GGCATCATTCTATGGTTACCCATCATTAGAGGAAATTTCCAA TAAACTCTGGTGTAAGGCTTAGAGTGATGGTCGAGGTGCCCT ATTTAGGGTGAGGAGCCTCGGTGGCAGCCCCTGGAAATCCTC TATTGGATAGGAACAGCTGTACTGGGCAGTTACAGCAGTCGT ATGGTAACACATGCGGCGTTCCGAAATACCATGGGTGGCGAT TCACAACAAGAA-3'
PK1+PK2+PK3_K/O	5'- GAGCGGTTTCTGGAATACTATATGTAAGTATAGTGTTCTGGA GGCATCATTCTATGGTTACCCATCATTAGAGGAAATTTCCAA TAAACTCTGGTGTAAGGCTTAGAGTGATGGTCGAGGTGCCCT ATTTAGGGTGAGGAGCCTCGGTGGCAGCCCCTGGAAATCCTC TTAACCATAGGAACAGCTGTACTGGGCAGTTACAGCAGTCGT ATGGTAACACATGCGGCGTTCCGAAATACCATGGGTGGCGAT TCACAACAAGAA-3'
Scrambled	5'- CCCGTATGGGGATCGGACCGTGTTGCGCGCGATTGGATATAC AAGCATAATTCTAAAAGGGACCTTGTCTGGGTTATCACTCCG ATCTTGCGTTAACCATTATGTTATAACGCAGACTATTGAGCC GCGAGACAAAGCCCCTTTGATTTTAAGTGACATCACTAGGTC AAGCAGCAAGGTCTGGCAGACTAGGTATTAAGTAAAGTGTG

	CATACGGTGAGTGTTGAGGCGCCGCCAGTCAGCGGGATAAA AGATCTTAGCTTAT-3'
PSIV	Sequence Upstream from sfGFP
WT	5'- AAGCTGACTATGTGATCTTATTAAAATTAGGTTAAATTTCGA GGTTAAAAATAGTTTTAATATTGCTATAGTCTTAGAGGTCTT GTATATTTATACTTACCACACAAGATGGACCGGAGCAGCCCT CCAATATCTAGTGTACCCTCGTGCTCGCTCAAACATTAAGTG GTGTTGTGCGAAAAGAATCTCACTTCAAGAAAAAGAATTTAC C-3'
PK1_K/O	AAGCTGACTATGTGATCTTATTAAAATTAGGTTAAATTTCGA GGTTAAAAATAGTTTTAATATTGCTATAGTCTTAGAGGTCTT GTATATTTATACTTACCACACAAGATGGACCGGAGCAGCCCT CCAATATCTAGTGTACCCTCGTGCTCGCTCAAACATTTTCACG TGTTGTGCGAAAAGAATCTCACTTCAAGAAAAAGAATTTACC -3'
PK2_K/O	5'- AAGCTGACTATGTGATCTTATTAAAATTAGGTTAAATTTCGA GGTTAAAAATAGTTTTAATATTGCTATAGTCTTAGAGGTCTT GTATATTTATACTTACCACACAAGATGGACCGGAGCAGCCCT CCAATATCTAGTGTACGGAGCTGCTCGCTCAAACATTAAGTG

	GTGTTGTGCGAAAAGAATCTCACTTCAAGAAAAAGAATTTAC C-3'
PK2_Opp	5'- AAGCTGACTATGTGATCTTATTAAAATTAGGTTAAATTTGCT CCTTAAAAATAGTTTTAATATTGCTATAGTCTTAGAGGTCTTG TATATTTATACTTACCACACAAGATGGACCGGAGCAGCCCTC CAATATCTAGTGTACCCTCGTGCTCGCTCAAACATTAAGTGG TGTTGTGCGAAAAGAATCTCACTTCAAGAAAAAGAATTTACC -3'
PK2_Comp	5'- AAGCTGACTATGTGATCTTATTAAAATTAGGTTAAATTTGCT CCTTAAAAATAGTTTTAATATTGCTATAGTCTTAGAGGTCTTG TATATTTATACTTACCACACAAGATGGACCGGAGCAGCCCTC CAATATCTAGTGTACGGAGCTGCTCGCTCAAACATTAAGTGG TGTTGTGCGAAAAGAATCTCACTTCAAGAAAAAGAATTTACC -3'
PK3_K/O	5'- AAGCTGACTATGTGATCTTATTAAAATTAGGTTAAATTTCGA GGTTAAAAATAGTTTTAATATTGCTATAGTCTTAGAGGTCTT GTATATTTATACTTACCACACAAGATGGACCGGAGCAGCCGA GGAATATCTAGTGTACCCTCGTGCTCGCTCAAACATTAAGTG

	GTGTTGTGCGAAAAGAATCTCACTTCAAGAAAAAGAATTTAC C-3'
PK1+PK3_K/O	5'- AAGCTGACTATGTGATCTTATTAAAATTAGGTTAAATTTCGA GGTTAAAAATAGTTTTAATATTGCTATAGTCTTAGAGGTCTT GTATATTTATACTTACCACACAAGATGGACCGGAGCAGCCGA GGAATATCTAGTGTACCCTCGTGCTCGCTCAAACATTTTCAC GTGTTGTGCGAAAAGAATCTCACTTCAAGAAAAAGAATTTAC C-3'
PK1+PK2+PK3_K/O	5'- AAGCTGACTATGTGATCTTATTAAAATTAGGTTAAATTTCGA GGTTAAAAATAGTTTTAATATTGCTATAGTCTTAGAGGTCTT GTATATTTATACTTACCACACAAGATGGACCGGAGCAGCCGA GGAATATCTAGTGTACGGAGCTGCTCGCTCAAACATTTTCAC GTGTTGTGCGAAAAGAATCTCACTTCAAGAAAAAGAATTTAC C-3'
SDS-like_K/O	5'- AAGCTGACTATGTGATCTTATTAAAATTAGGTTAAATTTCGA GGTTAAAAATAGTTTTAATATTGCTATAGTCTTAGAGGTCTT GTATATTTATACTTACCACACAAGATGGACCGGAGCAGCCCT CCAATATCTAGTGTACCCTCGTGCTCGCTCAAACATTAAGTG

	GTGTTGTGCGAAAAGAATCTCACTTCAACTGGTCTTGACGGC C-3'
RBS	Sequence Upstream from sfGFP
Strong (BBa_B0034)	5'-AAAGAGGAGAAATACTAG-3'
Medium (BBa_B0032)	5'-TCACACAGGAAAGTACTAG-3'
Weak (BBa_B0033)	5'-TCACACAGGACTACTAG-3'
Dead (BBa- B0034 Inv)	5'-TTTCTCCTCTTTACTAG-3'

*All PSIV bicistronic constructs are identical to the monocistronic constructs, except they have the following sequence (coding for T7-RBS-mRFP) upstream of the IGR IRES: 5'-TAATACGACTCACTATAGGGAGATACTAGAGTCACACAGGACTACTAGATGGCTTCC TCCGAAGACGTTATCAAAGAGTTCATGCGTTTCAAAGTTCGTATGGAAGGTTCCGTT AACGGTCACGAGTTCGAAATCGAAGGTGAAGGTGAAGGTCGTCCGTACGAAGGTAC CCAGACCGCTAAACTGAAAGTTACCAAAGGTGGTCCGCTGCCGTTCGCTTGGGACAT CCTGTCCCCGCAGTTCCAGTACGGTTCCAAAGCTTACGTTAAACACCCGGCTGACAT CCCGGACTACCTGAAACTGTCCTTCCCGGAAGGTTTCAAATGGGAACGTGTTATGAA CTTCGAAGACGGTGGTGTTGTTACCGTTACCCAGGACTCCTCCCTGCAAGACGGTGA GTTCATCTACAAAGTTAAACTGCGTGGTACCAACTTCCCGTCCGACGGTCCGGTTAT GCAGAAAAAAACCATGGGTTGGGAAGCTTCCACCGAACGTATGTACCCGGAAGACG GTGCTCTGAAAGGTGAAATCAAAATGCGTCTGAAACTGAAAGACGGTGGTCACTAC GACGCTGAAGTTAAAACCACCTACATGGCTAAAAAACCGGTTCAGCTGCCGGGTGC

TTACAAAACCGACATCAAACTGGACATCACCTCCCACAACGAAGACTACACCATCG

TTGAACAGTACGAACGTGCTGAAGGTCGTCACTCCACCGGTGCTTAATAACGCTGAT

AGTGCTAGTGTAGATCGCTACTAGAG-3’

Table S4.2. Sequences for control RNAs used in filter binding.

Control RNAs	Sequences used in filter binding
rpsO	5'-GGAUCCTAAUACGACUCACUAUAGGUACGAGUAGAAUACUGCCGCUUAACGUCGCGUAAAUUGUUUAACACUUUGCGUAACGUACACUGGGAUCGCUGAAUUAGAGAUCGGCGUCCUUUCAUUCUAUAUACUUUGGAGUUUUAAAAUGUCUCAAGUACUGAAGCAACAGCUAAAAUCGUUUCUGAGUUUGGUCGUGACGCAAACGACACC-3'
sodB	5'-GGAUCCUAAUACGACUCACUAUAGGAUACGCACAAUAAGGCUAUUGUACGUAUGCAAAUUAAUAAUAAAGGAGAGUAGCAAUGUCAUUCGAAUUACCUGCACUACCAUAUGCUAAAGAUGCUCUGGCACCGCACAUUUCUGCGG-3'
$tRNA^{Phe}$	5'-GCGCGGAUGCUCAGUCGGUAGAGCAGGGGAUUGAAAAUCCCCGUGUCCUUGGUUCGAUUCCGAGUCCGCGCACCA-3'

Table S4.3. Equilibrium dissociation constants and percent of RNA bound for CrPV IGR IRES variants and the 40S (HeLa) ribosomal subunit. Mean values of three replicates are reported; error indicates one standard deviation.

CrPV Construct	**K_D (nM)**	**Percent Bound (%)**
Wild type	14 ± 8	97 ± 7
PK1_K/O	4 ± 2	34 ± 2
PK1+PK3_K/O	No Binding	No Binding

Table S4.4. Equilibrium dissociation constants and percent of RNA bound for control RNAs and the 30S ribosome. Mean values of three replicates are reported; error indicates one standard deviation.

Construct	K_D (nM)	Percent Bound (%)
rpsO	10 ± 5	39 ± 2
sodB	13 ± 2	41 ± 3
$tRNA^{Phe}$	35 ± 10	10 ± 1

Table S4.5. Equilibrium dissociation constants and percent of RNA bound for control RNAs and the 70S ribosome. Mean values of three replicates are reported; error indicates one standard deviation.

Construct	K_D (nM)	Percent Bound (%)
rpsO	94 ± 14	44 ± 10
sodB	47 ± 13	56 ± 8
$tRNA^{Phe}$	No Binding	No Binding