The Role of Initiation Factor Dynamics in Translation Initiation

Margaret Mary Elvekrog

Submitted in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy
in the Graduate School of Arts and Sciences

COLUMBIA UNIVERSITY

2011

Abstract

The Role of Initiation Factor Dynamics in Translation Initiation

Margaret Mary Elvekrog

Like most biological polymerization reactions, ribosome-catalyzed protein synthesis, or translation, can be divided into initiation, elongation, and termination stages. Initiation is the rate-limiting stage of translation and a critical site for translational control of gene expression. Throughout all stages of protein synthesis, the ribosome is aided by essential protein co-factors known as translation factors. I have studied the role that two translation initiation factors, IF1 and IF3, play in the mechanism and regulation of translation initiation in *Escherichia coli*. Specifically, I have used single-molecule fluorescence resonance energy transfer (smFRET) as a primary tool for investigating how the dynamics of IF1 and IF3 regulate the accuracy with which the translational machinery selects an initiator transfer RNA (tRNA) and the correct messenger RNA (mRNA) start codon during the initiation stage of protein synthesis.

In Chapter Two, I describe investigations into the role of IF3's conformational dynamics in substrate selection. *E. coli* IF3 is an essential protein that binds directly to the small, 30S, ribosomal subunit and plays crucial roles in controlling the accuracy of tRNA and mRNA codon selection during translation initiation. IF3 is comprised of globular N-terminal (NTD) and C-terminal (CTD) domains that are connected *via* a flexible, lysine-rich linker that is highly conserved across prokaryotes in both length and hydrophilicity. By developing a fluorescently labeled IF3 that retains full biochemical activity, I have used smFRET to demonstrate that 30S subunit-bound IF3 can dynamically interconvert between three interdomain configurations during translation initiation. It is shown that formation of a properly base-paired anticodon-codon interaction between an initiator *N*-formylmethionyl-tRNA^{fMet} (fMet-tRNA^{fMet}) and an authentic AUG start codon within a correctly formed 30S initiation complex (30S IC) results in an IF3 interdomain reconfiguration. Interpreted within the context of the available genetic, biochemical, and structural data, it is proposed that this IF3 configuration represents a conformation of the 30S IC that is

optimized for rapid docking of the large, 50S, ribosomal subunit and formation of a 70S initiation complex (70S IC) that is ready to enter the elongation stage of protein synthesis. Work toward testing this mechanistic model is presented in Chapter Three.

In Chapter Four, I describe investigations into the dynamics of 30S IC assembly. Assembly of a 30S IC involves the association of IF1, IF2(GTP), IF3, initiator tRNA, and mRNA with a 30S subunit. Because the efficiency of 30S IC assembly is likely to vary across mRNAs depending on the details of each mRNA's translation initiation region, 30S IC assembly may serve as a critical point of regulation for the translational control of gene expression. The molecular details of 30S IC assembly remain poorly defined, thus preventing a full understanding of the assembly pathway(s) and its regulation. Due to its role in ribosome recycling and the high affinity with which it binds to the 30S subunit, current models of 30S IC assembly posit that IF3 is the first initiation factor that binds to the 30S subunit and that this factor remains bound to the 30S subunit throughout the assembly process. The order and relative timing of IF1, IF2(GTP), mRNA, and tRNA binding to the IF3-bound 30S subunit, however, are currently unknown. As a first step toward developing a complete kinetic scheme describing 30S IC assembly, I have developed an smFRET signal between IF3 and IF1. Due to its stable and early association with the 30S subunit, IF3 is an appropriate static landmark from which to monitor the kinetics of IF1 binding to the 30S subunit during 30S IC assembly. IF1 has roles in initiator tRNA and start codon selection, as well as stabilizing IF2 on the 30S subunit. Thus, elucidating the timing of IF1's binding to the 30S subunit relative to tRNA, mRNA, and IF2 binding provides previously inaccessible insight into the function of IF1 in regulating substrate selection and stabilizing IF2 on the 30S subunit. My results reveal that IF1 binds to the 30S subunit reversibly in the presence of just IF3, as well as in the presence of IF3 and initiator tRNA. Analogous experiments performed in the presence of both IF2 and IF3 demonstrate that IF2 virtually traps IF1 onto the 30S subunit. Taken together, this work reveals IF1's 30S subunit binding kinetics and lays the foundation for further investigations into the 30S IC assembly pathway.

Table of Contents

List of Figures

Number **Page**

List of Tables

List of abbreviations and acronyms

A site	aminoacyl-tRNA binding site
aa-tRNA	aminoacyl transfer ribonucleic acid
aIF	archaeal initiation factor
APD	avalanche photodiode
Arg	arginine
APS	ammonium persulfate
ASD	anti-Shine-Dalgarno
Asn	asparagine
BLAST	basic local alignment search tool
BME	β-mercaptoethanol
BSA	bovine serum albumin
CCD	charge coupled device
cDNA	complementary DNA
Cryo-EM	cryogenic electron microscopy
CTD	Carboxy-terminal domain
Cys	cysteine
Da	Dalton
DNA	deoxyribonucleic acid
E site	deacylated-tRNA binding (Exit) site
EDTA	ethylenediaminetetraacetic acid
EF	elongation factor
eIF	eukaryotic initiation factor
EMCCD	electron multiplying charge coupled device
Fe(II)-BABE	Fe(II)-1-[*p*-(bromoacetamido) benzyl]-EDTA
fMet	*N*-formyl-methionine
fMet-tRNA^{fMet}	initiator tRNA: *N*-formylmethionyl-tRNA
FOV	field of view
FRET	fluorescence resonance energy transfer
FWHM	full width at half maximum
GDP	guanosine diphosphate
GDPNP	guanosine-5'-[(β,γ)-imido]triphosphate
GMPPCP	guanosine-5'-[(β,γ)-methyleno]triphosphate
GTP	guanosine triphosphate

GTPase	guanosine triphosphatase
h	16S rRNA helix
H	23S rRNA helix
HIC	hydrophobic interaction chromatography
His_6	hexa-histidine
IC	initiation complex
IF	prokaryotic initiation factor
$IF2_{mt}$	mammalian mitochondrial IF2
IF2α	alpha isoform of initiation factor 2
IF2β	beta isoform of initiation factor 2
IF2γ	gamma isoform of initiation factor 2
IPTG	β-D-1-thiogalactopyranoside
k_a	association rate constant
k_d	dissociation rate constant
K_d	equilibrium dissociation constant
L (L9)	large ribosomal subunit protein
Lys	lysine
MALDI-TOF	matrix assisted laser desorption/ionization time-of-flight
MetRS	methionyl tRNA synthetase
mRNA	messenger RNA
MTF	methionyl-tRNA transformylase
MWCO	molecular weight cutoff
NA	numerical aperture
NBA	3-nitrobenzyl alcohol
NMR	Nuclear Magnetic Resonance
NOE	nuclear Overhauser effect
NOESY	nuclear Overhauser effect spectroscopy
NTA	nitrilotriacetic acid
NTD	amino-terminal domain
OB	oligomer binding
P site	peptidyl-tRNA binding site
PAGE	polyacrylamide gel electrophoresis
PDB	Protein Data Bank
PEG	polyethylene glycol
Phe	phenylalanine

P_i	inorganic phosphate
pI	isoelectric point
PMT	photomultiplier tube
PoTC	post-termination ribosomal complex
Pro	proline
PSI-BLAST	position-specific iterated basic local alignment search tool
PTC	peptidyltransferase center
r.t.	room temperature
RF	release factor
RNA	ribonucleic acid
RRF	ribosome recycling factor
rRNA	ribosomal RNA
RT	reverse transcriptase
S (30S, 50S, 70S)	Svedberg unit
S (S12)	small ribosomal subunit protein
SC	succinimidyl carbonate
SD	Shine-Dalgarno
SDGC	sucrose density gradient ultracentrifugation
smFRET	single-molecule fluorescence resonance energy transfer
T4gp32	gene product 32 from T4 bacteriophage
TCEP	tris(2-carboxyethyl)phosphine
TEMED	Tetramethylethylenediamine
TEV	tobacco etch virus
TIR	total internal reflection
TIR	translation initiation region
TIRF	total internal reflection fluorescence
TIRFM	total internal reflection fluorescence microscope
Tris	tris(hydroxymethyl)aminomethane
tRNA	transfer RNA
Tyr	tyrosine
UTR	untranslated region

Acknowledgements

This dissertation would have been impossible without the help, encouragement, and support of numerous people. Words fall short of expressing the gratitude I feel toward these mentors, colleagues, and friends.

First and foremost, I thank my research advisor, Professor Ruben L. Gonzalez, Jr. It has been a great privilege to be a member of his research group. Ruben has been endlessly patient and encouraging throughout the trials and tribulations of my graduate career, and his contagious enthusiasm has been a great source energy and inspiration. I am grateful for the excellent scientific training I have received under Ruben's guidance, and for instilling in me a love of biophysics.

I appreciate the insight and support of all my thesis committee members. It was an honor to have Prof. Joachim Frank, a giant in the field of ribosome structure and function, serve as an external examiner. Prof. Ann McDermott served on my committee throughout my graduate career and never fails to amaze me with her razor sharp observations and excellent recommendations. I have benefited immensely from Prof. Eric Greene's insightful questions and suggested experiments at my group meetings over the years. I have also admired Eric's methodical formulation of scientific questions and endless curiosity. Prof. Jack Norton has been a great support throughout my graduate career, both as chair of the graduate committee during my early years in the Ph.D. program, as a source of friendly conversation at happy hours, and as an inquisitive member of my committee since my first year. Lastly, I am especially grateful for Prof. Brian Gibney. It was my great fortune to be placed in his research group in the 2003 Columbia University Chemistry NSF-REU program, and his continued attentive mentorship and encouragement even after the program ended led to my further research training as a Fulbright fellow in Sweden and matriculation at Columbia as a graduate student. Clearly Brian's influence on my life and career has been immeasurable.

Being one of Ruben's first graduate students, I've had the pleasure of watching the lab grow and change over the course of my tenure in the lab. It's amazing to realize how much has happened in five years! I will always have many fond memories from my time in Havemeyer 614. It has been a lot of fun getting to know the other members of the Gonzalez lab, and I have benefited immensely from their support, insight, and extensive knowledge of the most delicious items on Columbia Cottage's menu, and the art and science of a good cup of coffee. These other lab members include Dan MacDougall, Dr. Jiangning Wang, Dr. Jingyi Fei, Sam Sternberg, Noam Prywes, Dr. Mike Englander, Dr. Dileep Pulukkunat, Wei Ning, Joseph Ho, Rathi Srinivas, Esther Wolff, Corey Perez, Colin Kinz-Thompson, Bridget Huang, Dr. Jon Bronson, Dr. Som Mitra, Dr. Kelvin Caban, Bo Chen, Victor Naumov, Peggy Bermel, Madeleine Jensen, and Dr. Jan-Willem van der Meent. It is my great fortune to be able to call all of these great people not just labmates, but friends.

I would never have pursued a career in science or aimed to earn a Ph.D. without the encouragement, inspiration, and painstaking training from a number of excellent educators and mentors throughout my life. The energy and charisma of Mrs. Sharon Nelson, a science teacher of mine at Waunakee High School, inspired me to take every science class I could in high school, and major in chemistry in college. At Viterbo University, Drs. Michael Collins, Vaughn Rodgers, Kyle Backstrand, Ron Amel, and Glena Temple were excellent science educators and always eager to support my dreams. They, along with many other Viterbo University professors including Earl Madary and Fr. Tom O'Neill, continue to influence my life today. Prof. Kai Griebenow (University of Puerto Rico), Prof. Brian Gibney (Columbia University), and Dr. Cecilia Tommos (Stockholm University) took the time to carefully train me in laboratory research methods and principles. I am grateful for their patience and belief in me.

I would also like to acknowledge the friends and coworkers who have been there all along:

- Kristina Westerlund – För att vara en lojal vän.
- Bruce Berry – For pointing out: "If you can do graduate school, you can do anything."

- Sam Hay – You're my scientific hero. Thanks for putting up with me when I was green.

- Sean Moran – a loyal and long-time friend. I appreciate your keen observations of the ridiculous (e.g. the studying in the cafeteria/eating in the library paradox).

- Koon (Koonie) Ching – "Ante-chamber-doors-R-us"

- Amit Reddi – "I'm not a ribosome! I'm a human being!"

- Eric McArthur – For your patience, and mongoose-like reflexes.

- Dan MacDougall – What about sea monkeys?!

- Jingyi Fei! – Really?! Okayyyyyyyy.

- Jiangning Wang – ^_^

- Noam Prywes – A kindred spirit in the sense of humor department.

- Heather Horgan – For dealing with a lot of nonsense, and having great taste in restaurants

- Jenny Baxter – For picnics, convertible rides through the Bronx, and sane conversations.

- Stephan Mackowiak – For your wisdom (e.g. "Every pot has a cover") and knowledge of German geography (especially the Weißwurstäquator). Oren's is still in business because of our coffee breaks!

- Corey ("Cody") Valdez – For your hilarious way with words ("All of them chemists!"), willingness to seek out NYC fun with me since Day 1 of STAT, and affinity for small, furry animals.

- Will Parks – We are the lucky ones!

- Sarah (Beara) Bohlman – I'm so glad we both ended up in NYC and have truly been friends forever.

- Pamela (Pammie) Rodriguez – For "punching" me in the face over a cookie. ☺

- Kathleen Kristian – My euchre partner in crime (dos puntos!), both in the sweltering heat of Rio Piedras and my knickknack-filled living room. ♣♦♥♠

- Jackie Catalano – a born event planner. Thanks for being a reliable source of information about departmental policies and events!

- Rachel Lemonik – Penne Pisticci!

- Yonah Lemonik – Who knew there was such a range of recipes for ten root vegetables?!

- Ian Halim – For hikes, bike rides, and $2 PBRs.

- Kat Allen – For all the delicious freshly baked bread, and a great living situation!

- Saráh Hale – For being a great friend and roommate, and for good tunes! ♪

Lastly, I am eternally grateful to my family, including my sister, Anne Marie Jarmuz, for being so generous toward me throughout my whole life. Thanks for all the visits and for handling that turkey… I thank my parents, Jim and Jean Elvekrog, who have always been unfailingly supportive of my love of science. From chauffeuring me to biotechnology seminars at UW Madison when I was 15, to cutting out newspaper clippings of scientific breakthroughs, I am thankful that you have always encouraged my career aspirations, even when these dreams have taken me far from home. Most of all, I am grateful for your love and guidance, throughout all of life's ups and downs.

“Satisfaction of one’s curiosity is one of the greatest sources of happiness in life.”
- Linus Pauling

“Somewhere, something incredible is waiting to be known.”
- Carl Sagan

Chapter 1
Introduction

1.1 Protein synthesis and the ribosome

Much of the work done within a cell is performed by proteins, and protein synthesis, folding, and incorporation into macromolecular assemblies represent the final steps in the gene expression pathway. Protein synthesis occurs on enzymatic ribonucleoprotein complexes called ribosomes. This process involves translation of the genetic code carried by a messenger RNA (mRNA) template by the ribosome using aminoacyl-transfer RNAs (aa-tRNAs) substrates. Although significant differences exist regarding the size of ribosomes and the complexity of the translation process across the three domains of life, the role of the ribosome in catalyzing protein synthesis is universally conserved [1].

In all species, the ribosome consists of two unequally-sized subunits. In bacteria, the small subunit has a relative sedimentation coefficient of 30S (Svedberg units), the large subunit has a sedimentation coefficient of 50S, and the intact 2.5 MDa ribosome sediments at 70S. Eukaryotic ribosomes are significantly larger, having 40S and 60S subunits, and an 80S ribosome. The differences between eukaryotic and prokaryotic ribosomes are exploited in medicine; antibiotics that interfere specifically with bacterial ribosomes have been used for many years to treat bacterial infections. The following discussion will be limited to prokaryotic ribosomes and translation.

As a ribonucleoprotein macromolecular complex, the ribosome is composed of approximately two thirds RNA and one third proteins. The 30S subunit consists of 21 proteins and a ribosomal RNA (rRNA) molecule of ~1,500 nucleotides which sediments at 16S. The 50S subunit has two rRNA molecules sedimenting at 5S and 23S (composed of 120 and 2,900 nucleotides, respectively), and 34 proteins [2]. Structural features of the ribosome are presented in Figure 1.1A.

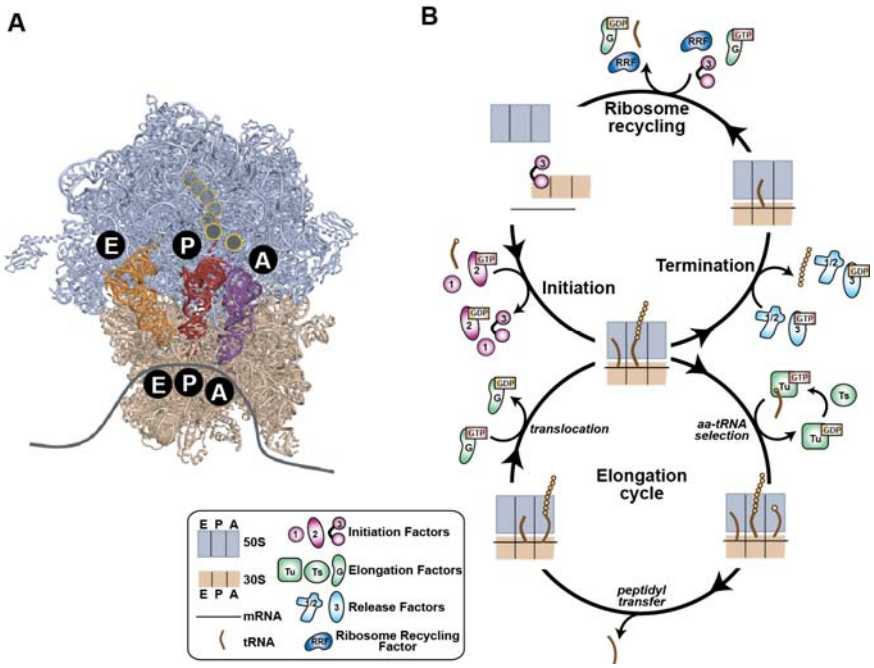

Figure 1.1 The 70S ribosome and the translation cycle. (A) The 70S ribosome with the three tRNA binding sites indicated. A = aminoacyl-tRNA (aa-tRNA) binding site, P = peptidyl-tRNA binding site, E = deacylated-tRNA binding ("exit") site. The large subunit is shown in blue and the small subunit in tan. The approximate mRNA binding site is drawn in gray. Figure from Ref. [3]. **(B)** The assembly of a translationally active ribosome occurs during initiation with the aid of three initiation factors, IF1, IF2, and IF3. The ribosome, assembled on the translation initiation region of the mRNA and containing initiator tRNA in its P site, then enters the elongation phase as the next tRNA is delivered to A site of the translational machinery in a ternary complex with elongation factor EF-Tu(GTP). Peptide bond formation occurs between the P- and A-site aminoacyl tRNAs and the growing polypeptide chain is transferred to the A-site tRNA. The tRNAs and mRNA next translocate through the ribosome with the aid of EF-G(GTP), placing the P-site and A-site tRNAs in the E and P sites, respectively. The A site is now free to accept the next aa-tRNA in the next round of elongation. The elongation cycle continues until a stop codon enters the A site. Release factor (RF) 1 or 2 recognizes the stop codon and binds to the ribosome. RF1/2 catalyzes polypeptide release, freeing the assembled polypeptide chain from the P-site tRNA, where it can now leave the ribosome through the exit tunnel. RF3(GTP) binds to the ribosome, prompting release of RF1/2, GTP hydrolysis, and its own release. The post-termination ribosomal complex (PoTC) is then recycled by ribosome recycling factor (RRF), EF-G(GTP), and IF3. Figure adapted from Ref. [3].

Translation is a polymerization process, and thus can be conceptualized as having initiation, elongation, and termination phases, followed by a ribosome recycling phase. A schematic overview of the translation cycle is found in Figure 1.1B, and a brief description of each stage can be found in the figure caption. The elongating ribosome performs three main functions: (1) a genetic function – the decoding of mRNA; (2) an enzymatic function – as a peptidyltransferase; and (3) a translocation function – as a conveyer of the mRNA chain and tRNA molecules through the ribosome during elongation (see Figure 1.1B) [4]. Inspection of base

pairing between an mRNA codon and a tRNA anticodon ('decoding') is performed on the 30S subunit, while peptide bond catalysis occurs in the peptidyltransferase center (PTC) of the 50S subunit. The two subunits cooperatively perform translocation.

The focus of my Ph.D. work has been translation initiation, and each of the molecular components of initiation are introduced and described in Section 1.2. Next, Section 1.3 describes details of the initiation factors (IFs), three essential protein translation factors that regulate initiation. My Ph.D. work has focused on IF3, and specifically the mechanism of IF3's role in regulating the fidelity of translation initiation. IF3's structure, function, and interactions with the ribosome are detailed in Section 1.3.3. The current state of our knowledge of the structural features of the 30S initiation complex (30S IC) and 70S initiation complex (70S IC) is presented in Section 1.3.4. The motivation for my Ph.D. work is described in Section 1.3.5, and an overview of the biophysical technique primarily employed for this work – single-molecule fluorescence resonance energy transfer (smFRET) – is presented in Section 1.4. Lastly, Section 1.5 summarizes this chapter and outlines the following chapters in this dissertation.

1.2 Translation Initiation – a highly regulated portal to gene expression

Initiation is the rate-limiting step in translation [5-8], taking several seconds *in vivo* [9], during which the two subunits of the ribosome and an aminoacylated and formylated initiator tRNA (*N*-formylmethionyl-tRNA) assemble at the start codon of an mRNA transcript. This assembly process sets the reading frame of the mRNA and thus is critical to the fidelity of gene expression. This 70S IC is the result of a multi-step assembly process involving the formation of several successive intermediate ribosomal complexes which differ in composition and conformation (Figure 1.2). The combined action of three protein initiation factors, IF1, IF2, and IF3, are essential for ensuring the speed and fidelity of each step in this process [10, 11].

There are three major intermediate ribosomal complexes assembled during the translation initiation pathway (reviewed in [11]). Much remains unknown about these initiation complexes, however, such as the timing and organization of IF1, IF2, and IF3 binding to the first complex, or the IFs' timing of dissociation from the last complex. The first IC is a relatively

unstable complex referred to as the 30S pre-initiation complex (30S pre-IC). Here, an mRNA is bound to a 30S subunit on its translation initiation region (TIR), along with a P (peptidyl-tRNA binding) site-bound initiator tRNA (fMet-tRNA^{fMet}) and IF1, IF2, and IF3. There is no physical interaction between the mRNA and initiator tRNA at this point. The 30S pre-IC then undergoes a rate-limiting conformational change to bring the mRNA start codon and initiator tRNA into direct contact in the 30S P site, where they can now form a codon-anticodon interaction. This new, stable complex is referred to as the 30S initiation complex (30S IC). The next step in the process is the association of the 50S subunit to the 30S IC to form a 70S IC. Stable 50S subunit joining is thought to occur concomitantly with dissociation of IF1 and IF3, GTP hydrolysis by IF2 followed by P_i release and IF2 dissociation. During this process, the initiator tRNA is also adjusted in the P site of the 50S subunit so that it is optimally positioned in the PTC for peptide bond formation with the first aminoacyl-tRNA that will be delivered to the ribosome during translation elongation. Consequently, the binding of the 50S subunit to the 30S IC, the positioning of the initiator tRNA, and the ejection of the IFs to form the 70S initiation complex (70S IC) mark the irreversible transition from the initiation to the elongation phase of translation.

Initiation is arguably the most critical stage of translation, however, as it sets the reading frame of a gene transcript, and, with *cis* and *trans* regulatory elements, determines the fate of a particular mRNA (i.e. whether or not the mRNA will enter translation elongation for rapid decoding). Indeed, regulation of gene expression at the translational level generally occurs during the initiation stage of translation and, in eukaryotes, numerous translationally regulated genes involved in both health and pathology have been identified [12, 13]. Although initiation is more complicated in eukaryotes and archaea than prokaryotes, important structural and mechanistic features are conserved between these three domains [12, 13]. The structures of eukaryotic, archael, and prokaryotic ribosomes reveal very similar mRNA binding and decoding sites, and all three prokaryotic IFs have structural or functional counterparts in eukaryotes and archaea (IF1 = e/aIF1A [14, 15]; IF2 = e/aIF5B [16]; IF3 = e/aIF1 [17, 18]). Thus, a mechanistic understanding of

the roles of the prokaryotic initiation factors is likely to be somewhat transferable to eukaryotic initiation.

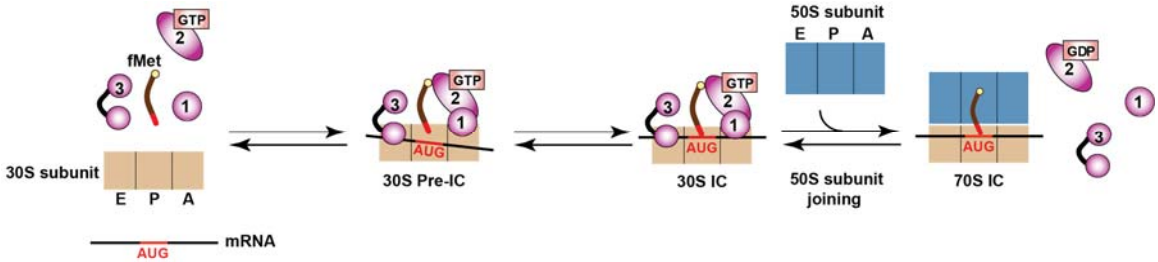

Figure 1.2 A minimal kinetic model of translation initiation. Three protein initiation factors, IFs 1, 2, and 3, assemble with an initiator tRNA, fMet-tRNA^{fMet}, and a start codon-carrying mRNA on the 30S ribosomal subunit (30S pre-IC). A rate-limiting conformational change of the initiation complex occurs during codon-anticodon interaction and 30S IC formation. The 50S ribosomal subunit rapidly joins to the correctly assembled 30S IC and is accompanied by dissociation of the IFs. The precise timing of these components' association and dissociation remains unknown. The figure cartoons are as follows: "30S subunit" indicates the small ribosomal subunit and "50S subunit" indicate the large ribosomal subunit; "E", "P", and "A" refer to the deacylated-, peptidyl-, and aminoacyl-tRNA binding sites, respectively; "1", "2", and "3" represent IF1, IF2, and IF3; "GTP" is guanosine triphosphate; "fMet" is *N*-formylmethionyl-tRNA^{fMet} (initiator tRNA); "mRNA" is a messenger RNA containing, at minimum, a translation initiation region (Shine-Dalgarno sequence, spacer, and AUG start codon); "30S IC" refers to a fully assembled 30S initiation complex containing a 30S ribosomal subunit, all three IFs, mRNA with an AUG start codon in the P site, and initiator tRNA bound to the P site; "70S IC" is a 70S ribosome containing P-site initiator tRNA bound to an AUG start codon.

1.2.1 The small (30S) ribosomal subunit

The 30S subunit is frequently described as morphologically resembling a "chick", with a head, neck, body, spur, and beak (Figure 1.3). Its 16S rRNA forms the scaffold for the subunit, with the ribosomal proteins mainly found on the periphery of the particle. The side of the 30S subunit that interfaces with the 50S subunit upon subunit association is mainly composed of rRNA, while its solvent-exposed back side contains many of the ribosomal proteins. The interface side also contains the binding sites for the mRNA, tRNAs, and translation factors [19-21]. The 30S and 50S subunits are associated into the 70S ribosome through 12 intersubunit bridges (see Table 1.1), which are formed from either rRNA-rRNA, rRNA-protein, or protein-protein contacts [22]. The standard nomenclature used to refer to ribosomal proteins and rRNA nucleotides is as follows: "S" and "L" for small subunit and large subunit proteins, respectively, and "h" and "H" for 16S rRNA and 23S rRNA helices, respectively.

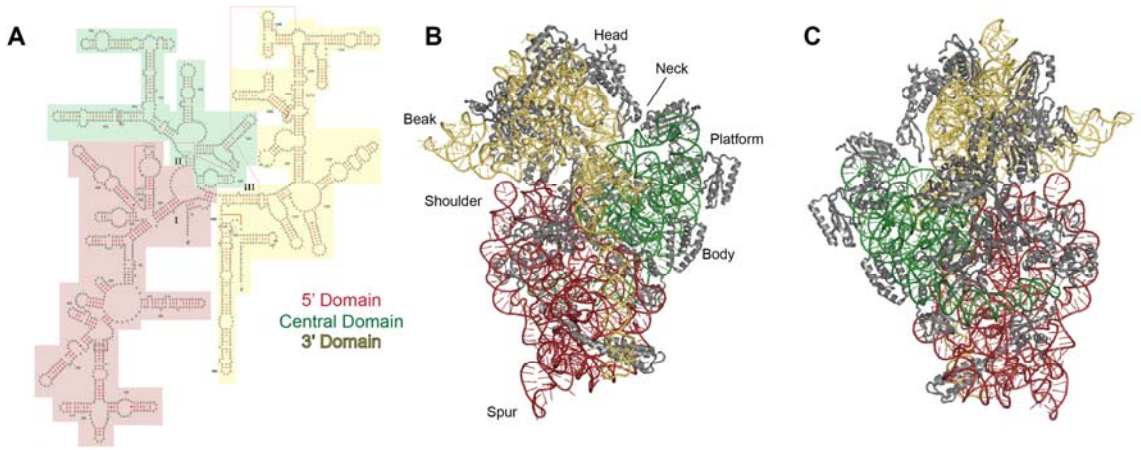

Figure 1.3 Overview of the 30S ribosomal subunit. (A) Secondary structure diagram of 16S RNA from *E. coli*. Adapted from Ref. [23, 24]. Domains are colored as indicated. **(B & C)** Front (subunit interface side) and back (solvent-exposed side) views of the 30S ribosomal subunit (PDB code 2J00). The morphological features are indicated and the domains of the ribosomal RNA are colored as in (A). Ribosomal proteins are indicated in gray. Modeled after a figure in Ref. [19]. The ribosomal subunits were rendered in PyMOL [25].

Table 1.1 Intersubunit bridges in the 70S ribosome. From Ref. [26]. Numbers refer to nucleotide (nt) positions in 16S or 23S rRNA. H, helix; S, small subunit protein; L, large subunit protein; P-P, protein-protein; R-R, RNA-RNA; R-P, RNA-protein; P-R, protein-RNA.

Interaction type	30S component (16S rRNA helix/nt or S protein)	50S component (23S rRNA helix/nt or L protein)
Bridge B1a		
P-R	S13	H38, 886-888
Bridge B1b		
P-P	S13	L5
Bridge B2a		
R-R	H44, 1408-1410, 1494-1495	H69, 1913-1914, 1918
Bridge B2b		
R-R	H24, 784-785, 794	H67/h69, 1836-1837, 1922
R-R	H45, 1516-1519	H71/h69, 1919-1920, 1932
Bridge B2c		
R-R	H24/H27, 770-771, 900-901	H67, 1832-1833
Bridge B3		
R-R	H44, 1484-1486	H71, 1947-1948, 1960-1961
Bridge B4		
R-R	H20, 763-764	H34, 717-718
P-R	S15	H34, 713, 717
Bridge B5		
R-R	H44, 1418-1419	H64, 1768-1769
R-P	H44, 1420-1422	L14
R-R	H44, 1474-1476	H62, 1689-1690
R-R	H44, 1474-1476	H64, 1989
Bridge B6		
R-R	H44, 1429-1430, 1474-1476	H62, 1689-1690, 1702-1705
R-P	H44, 1431	L19
Bridge B7a		
R-R	H23, 698, 702	H68, 1848-1849, 1896
Bridge B7b		
R-P	H23, 712-713	L2
R-P	H24, 773-776	L2
Bridge B8		
R-P	H14, 345-347	L14

1.2.2 The translation initiation region of prokaryotic mRNAs

Unlike eukaryotic mRNAs, bacterial mRNAs are normally polycistronic, containing multiple coding regions, and bind multiple ribosomes [2]. Another feature separating eukaryotes from prokaryotes is the presence of transcription and translation processes in the same cellular compartment, which allows co-transcriptional translation. The coupling of these processes helps minimize ribosomes' search for start codons.

To prevent translation initiation at internal methionine codons, all canonical mRNAs have TIRs with a few distinct elements. The TIR is found in the 5' untranslated region (UTR) upstream of the start codon and includes a purine-rich Shine-Dalgarno (SD) sequence that can be 3-9 bases in length [27]. The SD is complementary to nucleotides located at the 3' end of the 16S rRNA (the anti-SD (ASD) sequence: 5'-CACCUCCUU-3') (see Figure 1.4). The binding of the SD to the ASD anchors the mRNA to the 30S subunit and helps reduce the space to be searched in the hunt for the start codon [27, 28]. Once the SD-ASD interaction is established, the start codon is now positioned in the vicinity of the P site, where it can be recognized and ultimately locked into the P site by the initiator tRNA [29]. The length and sequence of the SD, as well as the length and sequence of the spacer between the SD and the start codon, are variable and can influence the efficiency with which a particular mRNA is translated. For example, extended SD-ASD complementarities, and consequently a short spacer between the SD and start codon, can reduce the expression level of a particular mRNA [30]. The start codon in *E. coli* is most frequently (90% of all genes) AUG, while GUG and UUG are used in 8% and 1% of all *E. coli* genes. The unusual start codons AUU and AUC are used in the remaining 1% of *E. coli* genes [31-33].

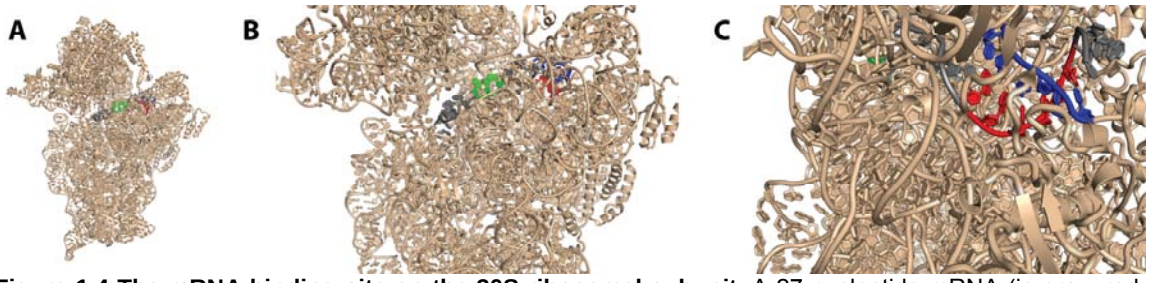

Figure 1.4 The mRNA binding site on the 30S ribosomal subunit. A 27 nucleotide mRNA (in gray, red, and green) is bound in the neck region of the 30S subunit. The AUG start codon (in green) is positioned near the P site. The Shine-Dalgarno (SD) sequence is indicated in red. The complementary anti-SD sequence at the 3' end of the 16S RNA is indicated in blue. **(A)** The mRNA-bound 30S subunit from *T. thermophilus* (PDB code 2HGR). **(B)** A closer view of the mRNA binding site, with the start codon and SD:ASD interaction depicted. **(C)** A 90° rotation about the vertical axis to more clearly view the SD:ASD interaction.

Another critical element in the TIR is the presence/absence of secondary structure within this region [34, 35]. Base-pairing within this region restricts ribosome-binding, making this a common means of translational control of gene expression. For example, riboswitches are sequences within the 5' UTRs of prokaryotic mRNAs whose secondary structures are altered in response to the binding of a metabolite, with the consequence of secluding or exposing the SD sequence and preventing or permitting translation under certain metabolic or stress conditions [36, 37].

1.2.3 The initiator tRNA

In prokaryotes, and also in the chloroplasts and mitochondria of eukaryotes, all polypeptide chains begin with *N*-formyl-methionine, though in many cases this *N*-terminal methionine is post-translationally or co-translationally removed or modified [38]. The initiator tRNA that delivers this first amino acid is $tRNA^{fMet}$, which differs from the elongator methionine tRNA in three distinct ways [39] (see Figure 1.5). $tRNA^{fMet}$ has a mismatch in its acceptor arm (C1•A72), which prevents elongation factor Tu (EF-Tu) from binding to $tRNA^{fMet}$ and delivering it to the ribosomal A (aminoacyl-tRNA binding) site during elongation [40]. This mismatch is also critical for interactions between Met-$tRNA^{fMet}$ and the methionyl-tRNA transformylase enzyme [41]. There are three highly conserved consecutive GC base pairs in the anticodon stem of $tRNA^{fMet}$ which rigidify the stem and help target the tRNA to the P site [39, 42, 43]. Two of these base pairs interact with P site 16S rRNA nucleotides, specifically: the tRNA base pair G29-C41 with the rRNA nucleotide

G1338, and the tRNA base pair G30-C40 with rRNA nucleotide A1339 [44, 45]. The third distinctive element in $tRNA^{fMet}$ is the presence of a purine-11-pyrimidine-24 in the dihydrouridine ("D") stem, which is unlike the pyrimidine-11-purine-24 base pair that is always found in elongator tRNAs. Both $tRNA^{Met}$ and $tRNA^{fMet}$ are aminoacylated by methionyl tRNA synthetase due to its recognition of their same anticodon: CAU. Met-$tRNA^{fMet}$ goes on to be formylated by methionyl-tRNA transformylase, which catalyzes the transfer of a formyl group from N^{10}-formyltetrahydrofolate to the α-amino group of the methionine [39]. fMet-$tRNA^{fMet}$ is recognized by IF2 by virtue of its formyl group [46], and possibly its acceptor arm [47] and elbow region [48], while IF3 recognizes its anticodon stem and loop elements [49]. These features can be introduced into non-initiator tRNAs, and cause them to be recognized as an initiator tRNA by these IFs [40, 46, 49, 50].

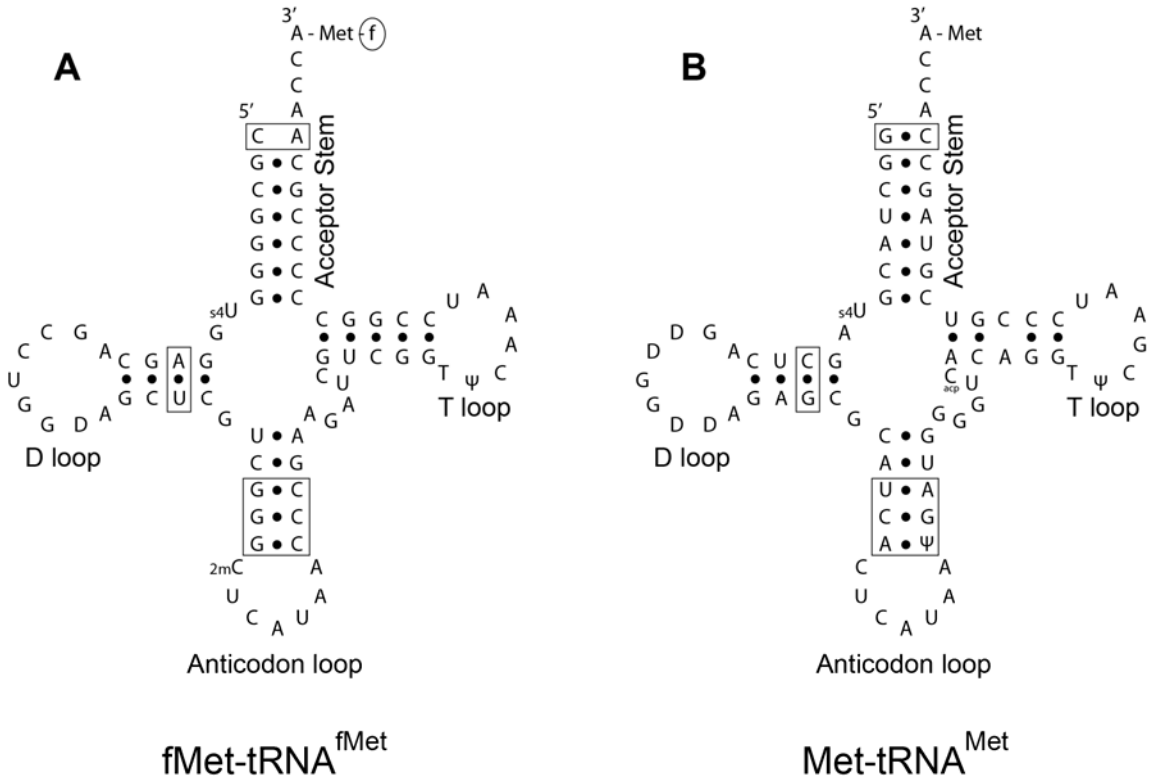

Figure 1.5 Initiator and elongator methionine tRNAs. Cloverleaf representation of the secondary structure of **(A)** the initiator fMet-$tRNA^{fMet}$, and **(B)** elongator Met-$tRNA^{Met}$. The regions critical for initiator tRNA recognition by the 30S ribosomal subunit, IF2, IF3, and the methionyl-tRNA transformylase are indicated in boxes. Additional details can be found in the text.

1.3 The Initiation Factors (IFs)

The 30S subunit shows no inherent preference for a start codon-initiator tRNA codon-anticodon interaction over a cognate elongator codon-anticodon interaction. With this in mind, it starts to become obvious that external chaperones must be used to ensure the fidelity of 30S IC and 70S IC formation. IF1, IF2, and IF3 perform this fidelity function in prokaryotes. All three of these factors are essential for cell viability [51-53] and their functions are thought to be at least partially conserved across all three domains of life [15-17].

1.3.1 Initiation Factor 1 (IF1)

IF1 is the smallest, and least understood, of the IFs. Homologs of IF1 are present in prokaryotes, archaea, and eukaryotes, clearly indicating a universal role for this initiation factor [14, 15, 54]. In *E. coli*, IF1 is encoded by the gene *infA* and is composed of 71 amino acids for a total molecular weight of 8.2 kDa. Structurally, it is composed of a rigid five-stranded β-barrel flanked by flexible and disordered extremities, with the residues connecting β3 and β4 (residues 38-44) forming a flexible 3_{10}-helix (Figure 1.6A). The N-terminal tail (residues 1-6) and C-terminal tail (residues 67-71) are highly flexible, as revealed by their intense (^{15}N,^{1}H)-nuclear Overhauser effects (NOEs), as measured by solution nuclear magnetic resonance (NMR) spectroscopy [55]. IF1 belongs to the common oligomer binding (OB) fold family of proteins, along with other nucleic acid-binding proteins including ribosomal protein S1, cold shock proteins CspA and CspB, and the N-terminal domain of aspartyl-tRNA synthetase [55].

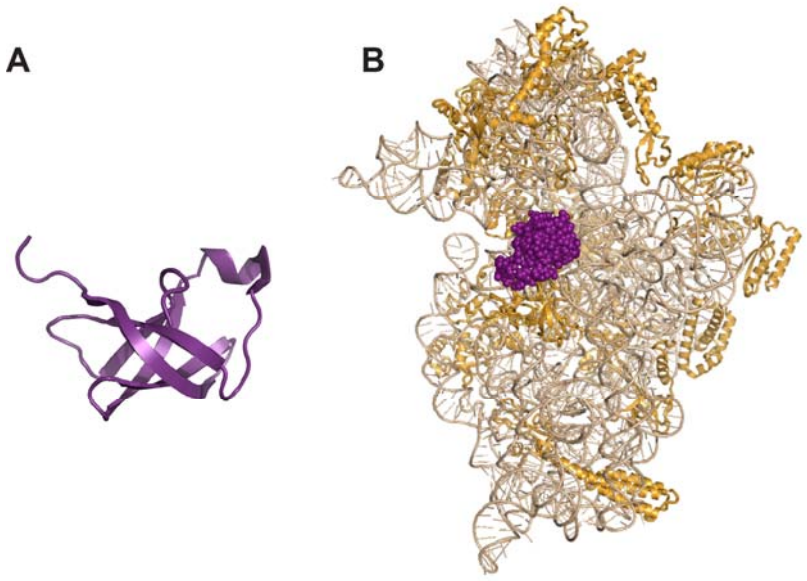

Figure 1.6 Initiation Factor 1. (A) Ribbon diagram of the structure of IF1 from *E. coli* determined with multidimensional NMR spectroscopy. Main-chain superposition of 19 refined structures. The dominant features are the five-stranded β-barrel and 3_{10} helix. PDB code 1AH9. **(B)** Crystal structure of a complex of IF1 with the *T. thermophilus* 30S ribosomal subunit. A space-filling model of IF1 is in purple, and a ribbon diagram of the 30S subunit is in wheat (RNA) and orange (ribosomal proteins). PDB code 1HR0. Structures rendered in PyMOL [25].

IF1's interactions with the 30S subunit are mainly electrostatic, as evidenced by the dependence of its affinity on buffer ionic strength, and involve the positively charged Lys and Arg residues on the protein interacting with the negatively charged 16S rRNA phosphate backbone [56-58]. Binding of IF1 to the 30S subunit and its equilibrium dissociation constant (K_d) from the 30S subunit has been measured by fluorescence polarization and is highly affected by the presence of IF2 and/or IF3. Alone, IF1 has a weak affinity for the 30S subunit (K_d = 2 μM). Adding IF3 decreases this value ~75-fold to 28 nM, while adding IF2 decreases it over 200-fold (K_d = 9 nM). Including IF2 and IF3 results in a K_d of 4 nM [59].

A crystal structure of *Thermus thermophilus* IF1 bound to a *T. thermophilus* 30S subunit was solved to 3.2 Å resolution [60] (Figure 1.6). This structure agreed with previous biochemical and mutagenesis work in localizing IF1's binding site to the A site of the 30S subunit [61, 62]. IF1 binds in a cleft between the 530 loop and h44 of 16S rRNA and ribosomal protein S12. IF1 causes several localized conformational changes upon binding the A site. These include flipping

of A1492 and A1493 such that they are no longer stacked within h44 and are then buried in pockets between IF1 and S12. This local conformational change induces longer-range conformational changes, including a rotation of the 30S head, platform and shoulder toward the A site. This IF1-bound conformational state of the 30S subunit is thought to affect the association-dissociation equilibrium of the ribosomal subunits during subunit joining [63].

IF1 is clearly an important cellular component, as it is essential for cell viability and influences the rate of translation: omission of IF1 from otherwise complete protein synthesis reaction mixtures results in a much-reduced translation rate [64]. An autonomous role of IF1 in translation has not yet been identified, but IF1 is known to function by enhancing the activities of IF2 and IF3 [65]. It enhances IF3's subunit anti-association activity (see section 1.3.3.7), and is thought to stabilize, and possibly position, IF2 for effective interaction with the initiator tRNA and optimal 50S subunit joining [66]. IF1 may also be required for dissociation of IF2 from 70S ICs [67]. In mammalian mitochondria, a 37 amino acid insert in the sequence of $IF2_{mt}$ has functionally replaced IF1 [68]. Comparable to IF1, this $IF2_{mt}$ insert region also binds in the A site of the 30S subunit [69].

1.3.2 Initiation Factor 2 (IF2)

IF2, like EF-Tu, elongation factor G (EF-G), and release factor 3 (RF3), is a ribosome-dependent guanosine triphosphatase (GTPase). It's the largest of the three IFs and is encoded by *infB* gene. IF2 has homologs in both archaea (aIF5B) and eukarya (eIF5B) [16]. There are three IF2 isoforms in *E. coli*: IF2α (97.3 kDa), IF2β (79.7 kDa), and IF2γ (78.8 kDa), which are tandemly translated from three independent start sites of the *infB* gene [70]. Their separate *in vivo* activities are indistinguishable and *E. coli* only needs one isoform to survive, however optimal growth conditions require all three isoforms [71, 72].

The structure of IF2 from *E. coli* has not yet been solved, but structures of the C-terminal core of aIF5B from *Methanobacterium thermoautotrophicum* have been solved [73] (Figure 1.7) in the nucleotide-free, GDP-bound, and GTP-bound forms. Overlays of these three structures reveal local conformational changes in the G domain that affect the C terminus, over 90 Å away [73]. IF2

contains six major domains [74]. Domains I-III are found in the N-terminal region. Both the primary structure and length are highly variable, and no specific function has been assigned to the N-terminal region. In fact, this region is not even required for cell viability in yeast (i.e. in the homologous eIF5B) [73]. The N-terminal region is connected to domain IV of the C-terminal region through a highly flexible linker [75]. Domains IV-VI make up the highly conserved C-terminal region. Domain IV is the G domain – the domain responsible for GTP binding and hydrolysis. It is an eight-stranded β-sheet adjacent to six α-helices and a 3_{10} helix. The G domain displays the conserved sequence elements characteristic of GTP binding proteins, such as Ras, Rho, and the signal recognition particle. Domain V is composed of a β-barrel, domain VI-I is an αβα sandwich connected to domain VI-2 by an extended α-helix. All of IF2's known functions are performed by the C-terminal region [76]. These functions include: (i) catalysis of 50S subunit association to the 30S IC, (ii) hydrolysis of GTP, and (iii) recognition of the initiator tRNA through its formyl group and 3' acceptor end [77, 78].

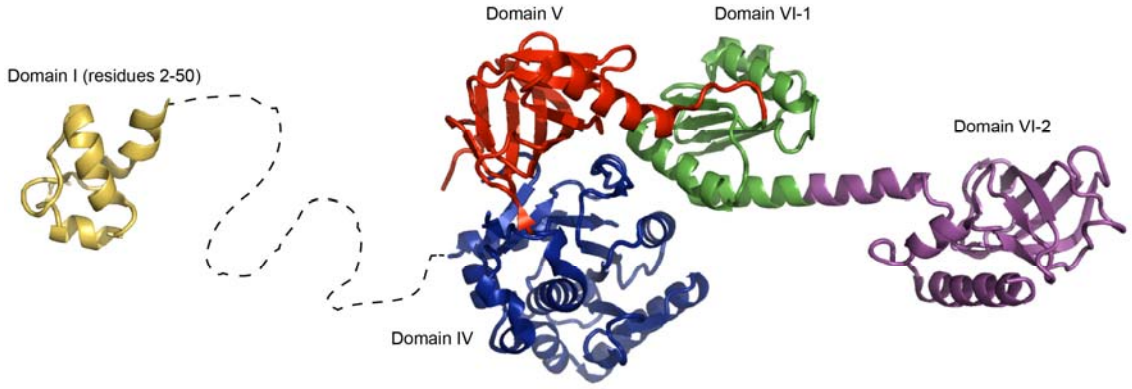

Figure 1.7 Initiation Factor 2. Ribbon diagrams of the IF2N domain (residues 2-50) from *E. coli* (left) and domains IV–VI-2 from the IF2 homolog aIF5B from *M. thermoautotrophicum* complexed with GDPNP (right). Each domain is indicated with a separate color and labeled according to *E. coli* nomenclature. Domains II and III, and part of I, (residues 51-390) are missing. The IF2N structure was solved by multidimensional NMR spectroscopy and the structure of aIF5B(GDPNP), along with aIF5B(GDP) and aIF5B, was solved with X-ray crystallography. PDB codes 1HD9 and 1G7T. Figure adapted from Ref. [11] and rendered in PyMOL [25].

IF2 is the only initiation factor that has high affinity for both ribosomal subunits [11]. Its binding to the 30S ribosomal subunit is through both ionic and hydrophobic interactions, and is highly affected by the presence of IF1, IF3, initiator tRNA, and GTP [79]. In their presence, IF2

has a K_d of 5.6 nM, while alone its K_d is 37 nM [79]. Early crosslinking studies indicated that IF2 and IF1 are in close proximity on the ribosome [62]. Cryogenic electron microscopy (Cryo-EM) studies differ regarding the presence of a direct contact between IF2 and IF1; one was absent in the 30S IC [80], but seemingly present in the 70S IC structure [47] (see section 1.3.4).

1.3.3 Initiation Factor 3 (IF3)

1.3.3.1 The gene and amino acid sequences of IF3

The gene encoding *E. coli* IF3, *infC*, has a 540-nucleotide open reading frame [31]. An interesting feature of the *infC* coding sequence is its use of the non-canonical start codon AUU, making *infC* one of only two known genes in *E. coli* to use this start codon, with the other being *pcnB* coding for poly(A) polymerase [31, 33]. The use of this AUU codon is conserved across all known prokaryotic *infC* sequences, with the exception being the *dsg* gene (a homolog of *infC*) of *Myxococcus xanthus*, which uses AUC instead [81]. In *E. coli*, this AUU start codon is located eleven bases downstream of the center of a strong Shine and Dalgarno (SD) sequence: GGAGGAA [27, 31]. IF3 from *E. coli* is largely basic in character, with an isoelectric point (pI) of 10.3 [31, 82], and is composed of 180 amino acid residues [83] for a total molecular mass of 20.5 kDa [31].

1.3.3.2 Structural features of IF3

IF3 is an elongated protein, composed of two similarly sized, globular domains connected by a hydrophilic, flexible linker peptide [84-90]. All attempts at crystallizing intact IF3 have thus far been unsuccessful [85, 91] apparently due to the interdomain linker's flexibility. The linker peptide is highly susceptible to proteolysis by numerous proteases, including trypsin, omptin, and *Staphylococcus aureus* endoproteinase Glu-C [84, 85, 92], while IF3's two globular domains are quite resistant to proteolysis. The proteolytic susceptibility of the linker has been exploited to investigate the structural features of IF3's globular domains separately [85-88].

The structures of the isolated N-terminal domain (NTD) and C-terminal domain (CTD) of IF3 from *Bacillus stearothermophilius* and *E. coli* were solved by X-ray crystallography [86] and NMR spectroscopy [87, 88], respectively. The sequence identity of IF3 from these two species is

54% [93] (see Figure 1.10). The NTD and CTD each have α/ß topology and contain an exposed ß-sheet, a motif common in many ribosomal, and other RNA-binding, proteins [86].

The crystal structure of IF3 NTD from *B. stearothermophilius* (residues 3 to 78) [86], and the structure of *E. coli* IF3 NTD (residues 7 to 83) solved by solution NMR spectroscopy [87], both reveal a globular α/ß fold containing one helix, H1, packed against a mixed four-stranded ß-sheet with central strands S1 and S4 parallel to each other and peripheral strands S2 and S3 antiparallel (Figure 1.8). 2D nuclear Overhauser effect spectroscopy (NOESY) is an NMR technique commonly employed to determine protein structure via correlation of through-space nuclear interactions within the protein. Here, 2D NOESY spectra of the NTD isolated by both proteolytic cleavage of the full-length IF3 and by overexpression of the truncated gene were indistinguishable [87], indicating that the domain was well folded and unaffected by proteolysis.

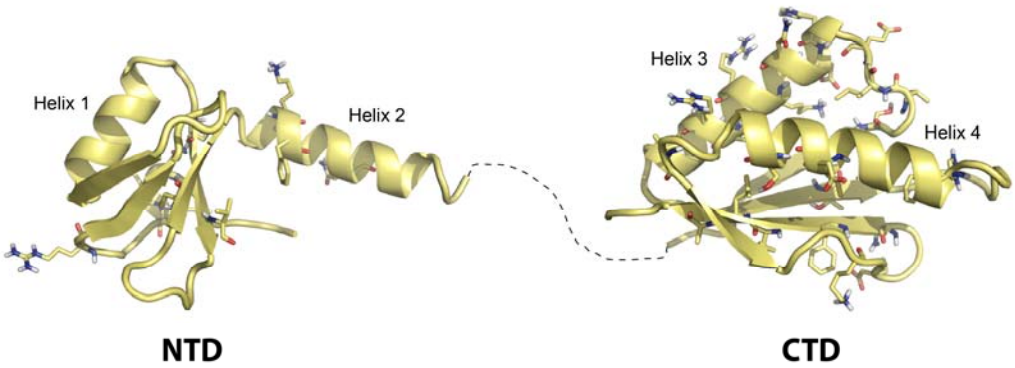

Figure 1.8 Initiation Factor 3. Ribbon diagrams of the crystal structures of IF3 N-terminal domain (NTD) (left) and C-terminal domain (CTD) (right) from *B. stearothermophilus*. The amino acids indicated by stick structures are those residues involved in 30S subunit binding, as determined by NMR spectroscopy (see text for details). The protruding helix in the NTD has been shown to be disordered and flexible in *E. coli*. Interdomain linker residues 78 to 82 (*B. stearo* numbering) are missing. PDB entries 1TIF and 1TIG. Structures rendered in PyMOL [25].

The structures of IF3's CTD from *B. stearothermophilus* and *E. coli*, solved by X-ray crystallography and NMR, respectively, reveal that it also has an α/ß fold, consisting of two parallel α-helices, H3 and H4 (Figure 1.8), which are packed against a mixed four-stranded ß-sheet. The first two strands of this sheet, S5 and S6, are parallel, while the next strand, S8, is anti-parallel to both S6 and S7 [86, 88]. The CTD very closely resembles U1A, a small protein that binds to the spliceosome [88].

1.3.3.3 Structural aspects of IF3's interdomain linker

One discrepancy between the crystal structure and NMR structure of IF3's NTD is the presence of an α-helix (H2) (residues 62 to 75 in *B. stearothermophilus* numbering, residues 70 to 83 in *E. coli*) at the C-terminus of the NTD that protrudes out from the globular α/ß domain which was seen in the crystal, but not the NMR, structure. In the NMR structure, the first six and last six amino acid residues (Ser78 to Glu83) were disordered. Eight of these H2 residues are interdomain linker residues (residues Glu76 to Lys89, *E. coli* numbering), which led to the proposal that IF3's interdomain linker is α-helical, not a flexible random coil. All other structural work on IF3 has refuted this conclusion, however. Circular dichroism and NMR spectroscopy were used to investigate the helical content of peptides containing the interdomain linker sequences from *E. coli* and *B. stearothermophilus* [90]. In solution, and under physiologically relevant temperatures for this mesophile and thermophile, the peptides were only 3% and 17% α-helical, respectively [90]. Under conditions comparable to those used to crystallize the NTD [85, 86]), the *B. stearothermophilus* peptide was 68% α-helical, suggesting that the crystallization conditions employed may have induced these linker residues to adopt a conformation that would not normally be stably sampled [90].

The conformation of intact *E. coli* IF3 was studied with heteronuclear NMR spectroscopy at 30°C and the rotational correlation times of the two globular domains were uncorrelated, indicating that their relative motions are independent and show little or no interaction, implying unrestricted motion of the interdomain linker [89]. In line with this, the amide protons within the linker region of the protein were shown to be in fast exchange with water [89]. IF3's lysine residues were selectively labeled with [α-^{15}N] lysine and heteronuclear relaxation studies of the linker residues displayed increased ^{15}N T_2 values and negative ^{1}H{^{15}N} NOEs, further indicating that the lysine-rich interdomain linker is highly flexible [89]. Additional evidence for the linker's disorder comes from its high degree of proteolytic susceptibility, both in solution and bound to

rRNA, as limited proteolysis can only occur in exposed flexible regions and not in helical chain segments of a protein [94].

The majority of all structural work on IF3, with the *B. stearothermophilus* IF3 NTD crystal structure as the sole exception, indicates that, in solution, IF3 behaves like two rigid globules connected by a flexible chain, and is thus able to adopt a large variety of interdomain conformations. These conformers were computationally modeled using the protein structure prediction program DIANA [95] by fixing the NTD's and CTD's tertiary structures but allowing unrestrained sampling of dihedral angle space for the linker [89]. The structure prediction algorithm produced a set of conformers with inter-domain distances ranging from 28 to 65 Å, with an average distance between the centers of mass of 46 Å [89], see Figure 1.9. This is consistent with a neutron scattering study of *B. stearothermophilus* IF3 in solution, where the average distance between the centers of mass of the two domains was also 46 Å [85].

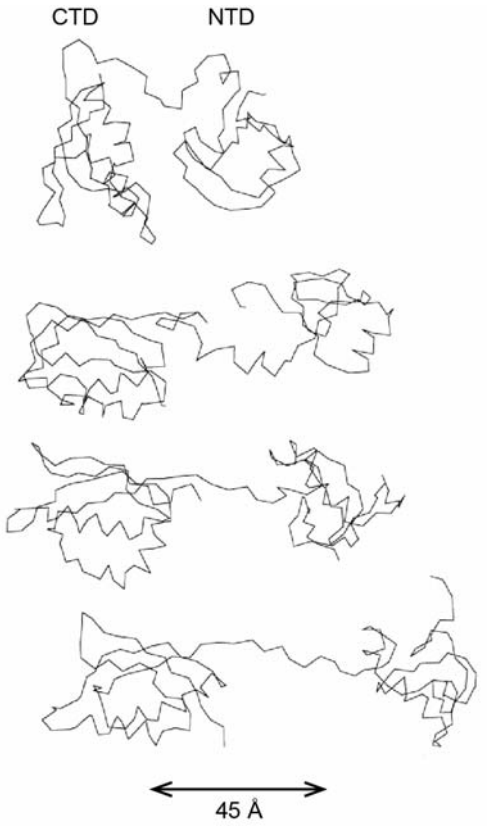

Figure 1.9 The IF3 interdomain linker is disordered in solution. Figure from Ref. [89]. Cα traces of four sample IF3 conformers obtained using DIANA, a protein structure prediction software program for NMR data [95]. The structures were obtained by leaving the linker region (residues 78 to 89) unrestrained, but restraining the backbone φ and Ψ dihedral angles of all residues within the structured part of the NTD and CTD (residues 11 to 77, and 90 to 176) to within ±1° of their respective values in the minimized mean structure obtained for each isolated domain. A set of 200 conformers was calculated and the 100 conformers having the lowest value for the target function were retained for statistical analysis. Of these, the interdomain distance varied from 28.2 to 64.5 Å, and the average distance between the centers of mass of the two domains was 46.3(±8.9) Å.

1.3.3.4 Conservation of the interdomain linker's flexibility, length, and hydrophilicity

In *E. coli*, the interdomain linker includes residues Glu76 to Lys89, and the corresponding *B. stearothermophilus* residues are Glu68 to Lys81. Deletion of the linker to produce separate CTD and NTD polypeptides results in a lethal phenotype and co-expression of the independently expressed CTD and NTD, even at high concentrations, cannot rescue the cells [94]. Clearly a physical link between the CTD and NTD is critical for IF3's function. A comparison of the length of this region, and the identity and character of its residues, across prokaryotes revealed that the specific sequence is not highly conserved, however its length, and the hydrophilic and basic

character of its residues, is very highly conserved through evolution [89, 94] (see Figure 1.10). The high conservation of its length and charge strongly suggest that either or both of these factors are important in IF3's function.

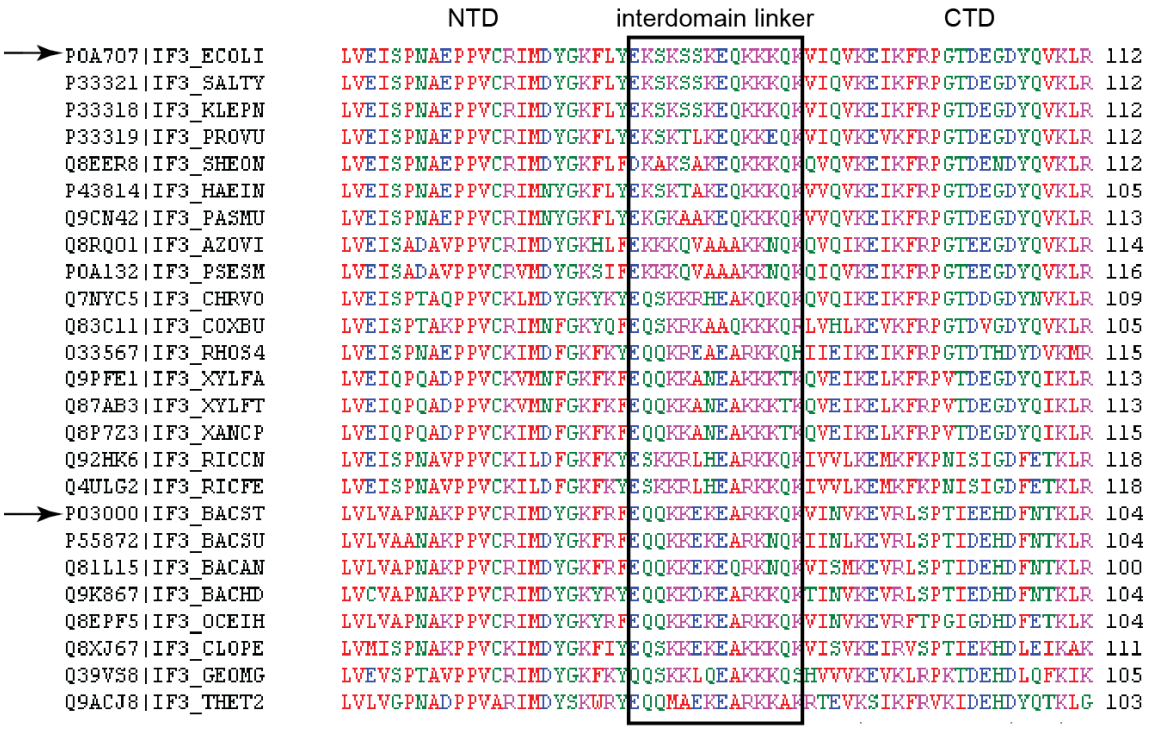

Figure 1.10 The IF3 interdomain sequence is highly conserved in length and hydrophilicity. Multiple sequence alignment of the IF3 sequence from 25 diverse bacterial species performed using position-specific iterated basic local alignment search tool (PSI-BLAST) and visualized in CLUSTALW [96, 97]. The interdomain linker region is indicated by the black box. The sequences for IF3 from *E. coli* and *B. stearothermophilus* are indicated with arrows.

Interestingly, a systematic *in vivo* mutational analysis of the length and character of the residues within the linker concluded that the linker could be shortened by up to six residues before affecting cell viability [94]. Additionally, the positively charged linker residues could be deleted or replaced with uncharged hydrophilic residues without loss of cell viability [94]. Replacing the hydrophilic linker with a poly-proline linker did result in a lethal phenotype however, indicating the need for flexibility within the linker. NMR spectroscopy confirmed that the structures of the poly-proline mutant and a $\Delta 8$ residue mutant were comparable to wild-type IF3 and binding studies showed that these IF3 mutants still bind with wild-type-like affinity to 30S subunits, thus demonstrating that loss of function was not the result of misfolding or impaired binding. The

results of this mutational study contradict the hypothesis that the linker's length and charge are critical to IF3's function, but instead imply that the linker's flexibility and hydrophilicity are most important.

1.3.3.5 Interactions between IF3 and the 30S ribosomal subunit

Throughout the 1980s, a number of studies were done to determine IF3's binding site on the ribosome using various crosslinking approaches, site-directed mutagenesis, and chemical and enzymatic protection studies. Although a few studies reported that IF3 bound to 50S subunits, today it is accepted that IF3's binding site is on the 30S subunit. Despite decades of effort, however, IF3's precise positioning on the 30S subunit remains unknown, though cryo-EM, directed hydroxyl radical probing, and X-ray crystal structures have helped narrow down its 30S subunit binding site, as will be described in the following paragraphs.

From crosslinking [62, 98-104], mutagenesis [105-107], and chemical protection [108] experiments, a number of ribosomal proteins and 16S rRNA nucleotides were identified that are either in direct contact with IF3, or whose accessibility changes upon IF3 binding. The ribosomal proteins that were repeatedly seen to be crosslinked with IF3, using a variety of different chemical, photochemical, and UV crosslinking agents, include S7, S11, S12, S13, S18, S19, and S21 (Figure 1.11A) [62, 98, 101-103]. Interestingly, these ribosomal proteins span a large surface area of the interface side of the 30S subunit, including the head, platform, and body regions of the subunit. This observation lead to suggestions that IF3 binds to both the head and platform of the 30S subunit [102]. Crosslinking between IF3 and 16S rRNA also revealed two distinct 16S rRNA regions that IF3 interacts with: the central domain (nucleotides 819-859) and the 3' end (helix 45: nucleotides 1506-1529) [62]. Later studies, using site-directed mutagenesis, identified additional nucleotides in 16S rRNA helix 24 that were critical for IF3 binding, including universally conserved nucleotides G791 [105] and A792 [106], as well as nucleotides G1530 and A1531 [107]. Chemical protection studies revealed that some helix 23 and 24 16S rRNA nucleotides, including G700, G703, and G791, are protected from kethoxal attack when IF3 was bound [108], and hydroxyl radical protection revealed an extensive footprint of nucleotides in helix 23 (685-688 and

693-703) and helix 24 (774-776, 783-793, 799-802, and 807-810) that surround these bases [109]. Thus, it seems that IF3 interacts with helices 23, 24, and 45 of the platform region of the 30S subunit (Figure 1.11B).

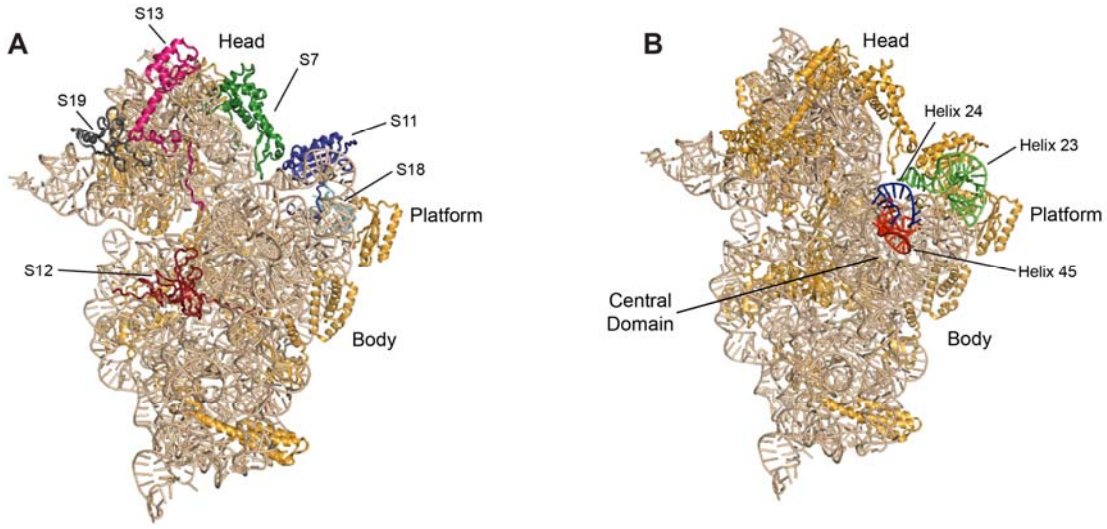

Figure 1.11 Regions of the 30S subunit involved in IF3 binding. (A) Ribosomal proteins that have been crosslinked with IF3 or protected from hydroxyl radicals or chemical modification by IF3. Proteins S1 and S21 are not shown. Details on the probing techniques and affected proteins can be found in Appendix C. **(B)** The helices in the central domain region of the 16S rRNA that IF3 has been shown to interact with. PDB code 2J00.

Many early studies were also done to identify the amino acid residues within IF3 that are critical for its interactions with the 30S subunit. Again, site-directed mutagenesis and chemical modification approaches were taken which revealed a number of residues that are essential for its 30S binding. These include NTD residue Tyr71, and CTD residues Tyr107, Tyr109, Lys110, Lys112, and Pro176 [110-113]. Later, (^{15}N,^{1}H)-HSQC NMR spectroscopy was used to more directly probe which IF3 NTD and CTD residues are involved in subunit binding during a 30S subunit titration [114]. These residues were identified by intensity changes in the cross peaks belonging to amides of individual IF3 amino acids during titrations of IF3 into 30S subunits. The affected residues are indicated in Figure 1.8, clearly demonstrating that most IF3-30S interactions are through the CTD.

A low resolution (27 Å) cryo-EM structure of an IF3-30S subunit complex confirmed that IF3's binding site lies in the cleft between the head and platform of the 30S in agreement with the

position implied in Figures 1.11A and B, however relative placement of the NTD and CTD in the positive electron density could not be unambiguously determined due to their similar sizes and morphologies, and the resolution level of the electron maps [115]. Further confidence in the specific location of IF3 on the 30S subunit came with a directed hydroxyl radical probing study using Fe(II)-1-[*p*-(bromoacetamido) benzyl]-EDTA (BABE) probes attached to 14 single Cys residues introduced into IF3 (5 on the NTD, 2 on the linker, 7 on the CTD) [109]. Fe(II)-BABE reacts to form thioether bonds with cysteine residues on a protein through the BABE bromoacetyl group, and iron redox chemistry, which generates hydroxyl radicals, is stimulated by the addition of ascorbate and hydrogen peroxide [116]. These hydroxyl radicals cleave RNA chains found within 20 Å of the Fe(II)-BABE probe. IF3's interactions with both 16S rRNA and P site- and mRNA-bound initiator tRNA were probed with this directed approach. Only the location of the CTD was determined with a relatively high confidence level, however, as only one of the five NTD probes cleaved rRNA, thus making the placement of this domain rather tenuous. The CTD was localized to the platform interface of the 30S, near the P site, while the NTD was placed near the E site, on the opposite side of the 30S- and mRNA-bound initiator tRNA. The placement of the CTD was in agreement with the cryo-EM study and the position implied in Figure 1.11, while the placement of the NTD differed (see Figure 2.1 in chapter 2 for the placement of IF3 on the 30S subunit determined in these studies). An X-ray crystal structure of *T. thermophilus* IF3-CTD bound to a *T. thermophilus* 30S subunit [117] reported that IF3-CTD binds on the upper end of the platform on the solvent side of the 30S subunit, between helices 23, 26, and the 3' end of helix 45. This location may represent a secondary binding site of IF3, and does not rule out the interface side of the platform region of the 30S subunit as the primary binding site. The platform region of the 30S subunit coincides with crystal contacts in the *T. thermophilus* 30S crystals, thus making it unlikely that IF3, soaked into these crystals, would be able to bind to this location without disrupting crystal packing [109].

Collectively, these biochemical and structural studies reveal the general vicinity of the 30S subunit that is involved in IF3 binding (Figure 1.12A). This region of the 30S subunit is also

involved in the formation of a number of intersubunit bridges, including B1a, B1b, B2a, B2b, and B7a (see Table 1.1 and Figure 1.12B), which may enable IF3 to perform its anti-50S subunit association function during ribosome recycling, as well as regulate 50S subunit joining in the later stages of translation initiation (see section 1.3.3.7). IF3's precise binding site, however, both in the absence and presence of mRNA, initiator tRNA, and the other initiation factors, awaits higher resolution cryo-EM or X-ray crystal structures.

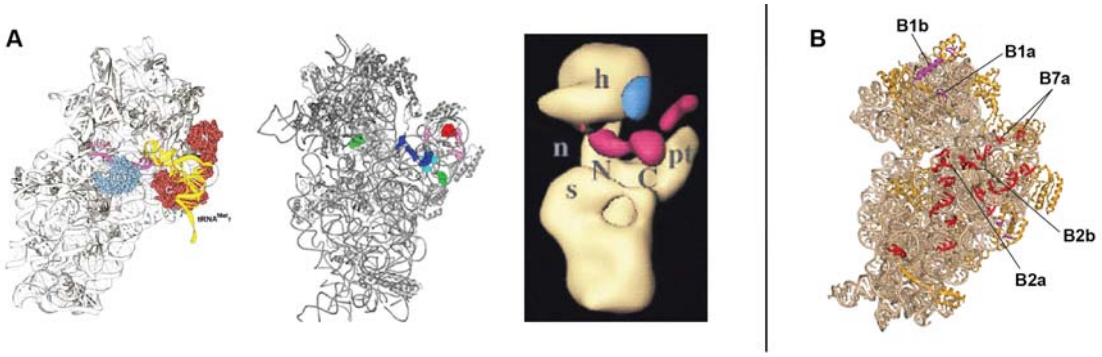

Figure 1.12 The approximate 30S subunit binding site of IF3. A. Directed hydroxyl radical probing [109], time-resolved chemical footprinting [118], and cryo-EM [115], respectively, were used to investigate the location of IF3 on the 30S subunit. Discrepancies exist regarding the relative position of IF3's NTD and CTD, but all three studies place IF3 on the platform of the 30S, near the neck. **B.** Regions of the 30S subunit involved in intersubunit bridges. 16S rRNA in wheat and ribosomal proteins in orange except for those rRNA residues (red) and r-protein residues (magenta) involved in RNA-RNA, RNA-protein, or protein-protein intersubunit bridge interactions. Figure from Mr. Wei Ning. IF3's binding site coincides with that of bridges B1a, B1b, B2a, B2b, and B7a (indicated) [26].

1.3.3.6 IF3's affinity for the 30S ribosomal subunit

The K_d for fluorescein-labeled IF3's interaction with 30S subunits was measured with fluorescence polarization to be 33 nM [119, 120]. IF3's two globular domains have different affinities for the 30S subunit, as evidenced by an NMR titration of intact IF3 with 30S subunits [114], and also through binding studies of the isolated NTD and CTD. The CTD is the 16S rRNA-binding domain of IF3 and has a relatively high affinity (K_d = ~3 μM) for the 30S subunit even in the absence of the NTD [121]. The NTD, on the other hand, has far fewer specific interactions with 16S rRNA nucleotides (Figure 1.8) and instead is thought to interact mainly with ribosomal proteins. In the absence of the CTD, the isolated NTD has a very weak, unmeasureable affinity for the 30S subunit [88, 121]. These differential affinities may explain the preferred order of IF3's

binding with the 30S subunit, which has been observed with both NMR spectroscopy and time-resolved chemical footprinting to occur first through IF3's CTD, and then through its NTD [114, 118] (see Figure 1.13). The presence/absence of IF1 and/or IF2 in the ribosomal complex does not affect IF3's binding affinity for the 30S subunit [120]. Likewise, their presence/absence does not affect the IF3-dependent hydroxyl radical footprinting pattern of the 16S rRNA, indicating that IF3's binding may be similar in the presence or absence of the other initiation factors [109]. Notably, nearly all of the detected signals from the Fe(II)-BABE probes were from IF3's CTD, so the influence of the other initiation components on IF3's NTD remains to be shown.

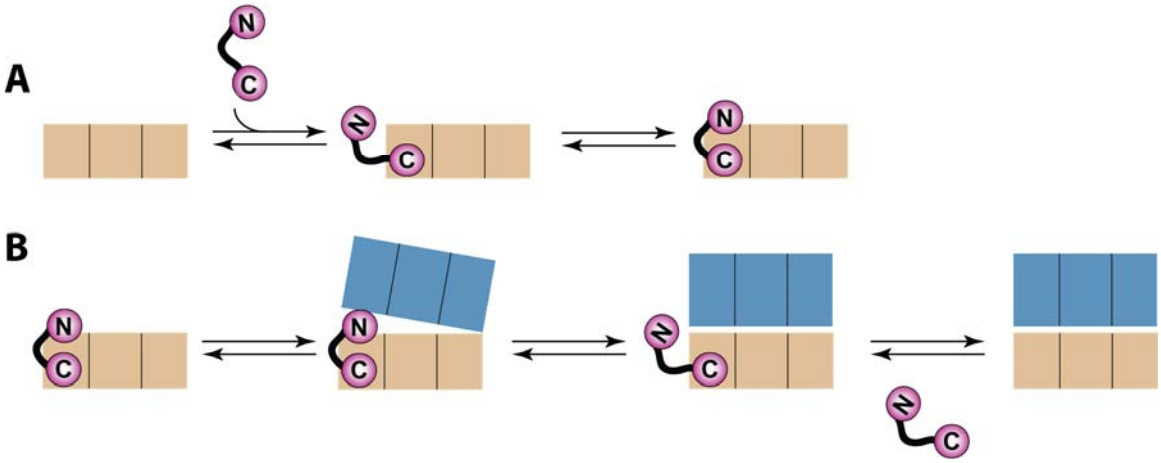

Figure 1.13 Cartoon depiction of IF3's stepwise association with and dissociation from the 30S subunit. A. Time-resolved chemical probing [118] and titrations of ^{15}N-labeled IF3 with 30S subunits monitored by two-dimensional (15N,1H)-HSQC NMR spectroscopy [114] reveal that IF3's CTD associates with the 30S subunit prior to NTD binding. **B.** Time-resolved chemical probing [118] also revealed that IF3 dissociates in the reverse order: first the NTD dissociates, then the CTD. The presence of IF3 hinders the establishment of a subset of intersubunit bridges, including B7a (see Figure 1.12B).

1.3.3.7 IF3 regulates translation initiation

One of the first functions of IF3 that was identified and thoroughly investigated is its role as a ribosome dissociation factor [122-124]. This title was eventually changed to "anti-association factor" since it was shown, through various kinetic studies, that IF3 functions by binding to vacant (i.e. lacking mRNA, tRNA, IF1, and IF2) 30S subunits and preventing premature 50S association (Scheme 1), rather than through active dissociation of 70S ribosomes (Scheme 2) [125-128].

$$70S \rightleftarrows 50S + 30S \qquad \text{Scheme 1}$$

$$IF3 + 30S \rightleftarrows 30S\bullet IF3$$

$$\text{IF3} + 70S \rightleftarrows [70S\bullet IF3] \qquad \text{Scheme 2}$$
$$[70S\bullet IF3] \rightleftarrows 50S + 30S\bullet IF3$$

In the absence of tRNA, mRNA, and the other initiation factors, and at physiologically relevant Mg^{2+} concentrations (3-6 mM), IF3 binds tightly to 30S subunits (K_d = 33 nM, see section 1.3.3.6) and inhibits stable 70S ribosome formation possibly by occluding regions of the 30S subunit that are involved in essential intersubunit bridges, including bridges B1a, B1b, B2a, B2b, and B7a (see section 1.3.3.5) [109]. This function is somehow relaxed on complete and correctly assembled 30S ICs, however, as 50S subunits will rapidly join to 30S ICs containing IF1, IF2(GTP), IF3, mRNA, and initiator tRNA [129-131]. It remains to be shown, however, how IF3 blocks subunit joining under some conditions, but relaxes this block under other conditions. Empty 70S ribosomes readily form from free 30S and 50S subunits in the absence of IF3, even if IF1, IF2, and initiator tRNA are present in the solution [129-131]. Thus, IF3 plays an important role in preventing the formation of these 'dead-end' 70S ribosomes.

IF3's role as a subunit anti-association factor is in play during both the late stages of ribosome recycling (see Figure 1.1) and the early stages of 30S IC assembly [132, 133]. Its role as an "initiation factor" extends beyond this anti-association role, however, as IF3 is critically important in substrate selection during initiation. In the absence of the other IFs, 30S•IF3 complexes select against all codon-anticodon interactions other than $tRNA^{fMet}$-AUG, $tRNA^{fMet}$-GUG, and $tRNA^{fMet}$-UUG [49, 65, 134-137]. All other P-site codon-anticodon interactions, even if they are a cognate pair (e.g. $tRNA^{Phe}$-UUU), are destabilized if IF3 is present on the 30S subunit. IF3, or the 30S•IF3 complex, shows a particular preference for the anticodon stem and loop of $tRNA^{fMet}$, and even a minimal $tRNA^{fMet}$ mimic, containing just these elements (see section 1.2.3), can be selected if an AUG, GUG, or UUG codon is bound in the P site of the 30S•IF3 complex [49]. Whether this selection occurs through an IF3-induced 30S subunit conformational change that reveals a preference by the 30S subunit for initiator codon-anticodon interactions [44], or

through direct recognition of the codon-anticodon interaction by IF3 [49], or through any number of other possible mechanisms, remains to be clearly shown.

In contrast with the aforementioned mechanism in which IF3 inspects the codon-anticodon interaction and specifically selects an initiator tRNA interacting with a start codon, another model suggests that IF3 indiscriminately destabilizes all tRNAs, including the initiator tRNA [72], and that IF2 is primarily responsible for specifically promoting the stabilization of initiator tRNA. That is, there is a kinetic tug-of-war between IF3 and IF2 to promote selection of just fMet-tRNA^{fMet} over all other aa-tRNAs, including unformylated Met-tRNA^{fMet}. Systematic investigations of the effect of each IF on the kinetics of tRNA binding revealed that IF3 indiscriminately increases the dissociation rate constant, k_d, of all three tRNAs studied (fMet-tRNA^{fMet}, Met-tRNA^{fMet}, and Phe-tRNA^{Phe}) [130, 131]. IF1 and IF2 counter this effect by varying the association rate constant, k_a, to favor the binding of fMet-tRNA^{fMet}. Thus, there is a clear discrepancy in the translation initiation field regarding the mechanism of IF3's function in substrate selection.

Other functions of IF3 include start codon selection. IF3 also prevents initiation from internal or non-canonical start codons (non-AUG, GUG, or UUG) [138], as well as leaderless mRNAs [138, 139]. These functions were revealed through *infC* allele mutations (e.g. *infC135*), which showed increases in initiation from these types of mRNAs and codons [140-142]. These *infC* mutants also displayed decreased abilities in preventing initiation from tRNAs other than tRNA^{fMet} [138] and start codons other than AUG, GUG, and UUG [81, 140, 141].

IF3's two-domain structure and ability to perform a number of seemingly disparate functions led some to speculate that each of IF3's domains were uniquely responsible for distinct functions [121]. Although IF3's CTD alone can perform the subunit anti-association role *in vitro* [121], the intact protein is required for IF3's substrate discrimination functions. This is evident from point mutations in the NTD, CTD, and linker that individually inactivate IF3's discriminatory abilities without disrupting its 30S binding affinity [113, 141]. Additionally, *in vivo* work investigating the length and amino acid character requirements of the linker peptide

demonstrated the absolute requirement for intact IF3, albeit with some allowed plasticity in the linker's length and composition, in IF3's functions [94].

1.3.3.8 Autoregulation of *infC* expression

The gene for IF3, *infC*, begins with the unusual start codon AUU in all known prokaryotic species (see section 1.3.3.1) [31]. In line with IF3's known functions in start codon and tRNA discrimination, IF3 controls expression of its own gene at the translational level [143, 144]. In a clear demonstration that IF3's role in initiation fidelity is also performed *in vivo*, decreased intracellular IF3 concentrations trigger an increase in spurious initiation [51], as do 16S rRNA mutations that decrease IF3's affinity for the 30S subunit [145, 146]. When *infC*'s AUU codon is changed to AUG, IF3 expression increases 30-fold [144, 147, 148], while conversion of other genes' canonical start codons to AUU results in repression of their expression unless they are expressed in a mutant *infC* background [141]. Similarly, IF3 expression is derepressed in mutants deficient for IF3 function (e.g. *infC135*) [140].

1.3.4 Structural insights into the 30S initiation complex and 70S initiation complex

The 30S IC is a relatively stable intermediate along the translation initiation pathway in which the reading frame of an mRNA is established (see section 1.2). Its assembly requires two main steps: 30S subunit binding to the TIR of an mRNA, and stable association of an initiator tRNA with the mRNA's start codon in the P site of the 30S subunit. These events are accomplished *in vivo* with the aid of IF1, IF2(GTP), and IF3, thus making the 30S IC a seven-component macromolecular complex. At this time, no structures have been solved of a completely assembled 30S IC. There is, however, a cryo-EM structure of a 30S IC lacking just IF3, thus composed of a 30S subunit, mRNA, fMet-tRNA^{fMet}, IF1, and IF2(GTP) [80] (Figure 1.14). This structure is valuable for its insights into the interactions of IF2 and fMet-tRNA^{fMet} on the 30S subunit; however, since this initiation complex was formed in the absence of IF3, one needs to exercise caution in drawing too many conclusions about the physiological relevance of this complex and its components' binding sites.

T. thermophilus initiation components were used to assemble these 30S ICs [80]. Five subpopulations of particles were parsed apart from the heterogeneous sample, though the major subpopulation (40% of the total particles) was the completely assembled 30S IC (though lacking IF3) (see Figure 1.14). In that major subpopulation, fMet-tRNA^{fMet} is stabilized on the 30S subunit through two major interactions: (1) its decoding stem is bound to the P site; (2) its acceptor end is in direct contact with domain IV of IF2 (see section 1.2.3). The conformation of fMet-tRNA^{fMet} is similar to that seen in an X-ray crystal structure of a 70S ribosome carrying an AUG start codon and initiator tRNA at the P site [149], though differs enough to assign the tRNA's 30S IC conformation to an intermediate state: "30S P/I." The specific differences include a slight, clockwise rotation of the decoding stem, bending it towards the initiation codon, and a re-positioning of the tRNA's elbow towards the exit site. These conformational changes may require the presence of IF2, since they were not seen in the 70S crystal structure lacking IF2. IF2's location and position on the 30S IC are also revealing. Its N-terminal and G1 domains were not visible in the cryo-EM maps, suggesting that they are disordered within the 30S IC. The location of its G domain is comparable to those of other ribosome-dependent GTPases bound to the ribosome. IF2 is anchored to the 30S subunit *via* its interaction with fMet-tRNA^{fMet}'s acceptor arm and also through interactions of domains I/II near helices h5 and h14 of 16S rRNA. Interestingly, this structure revealed the lack of a direct interaction between IF2 and IF1, contrary to long-standing hypotheses [62, 82]. The two factors are mutually stabilizing, however, and the absence of a direct interaction between the two suggests that their stabilization may occur through factor-induced conformational changes of the 30S subunit.

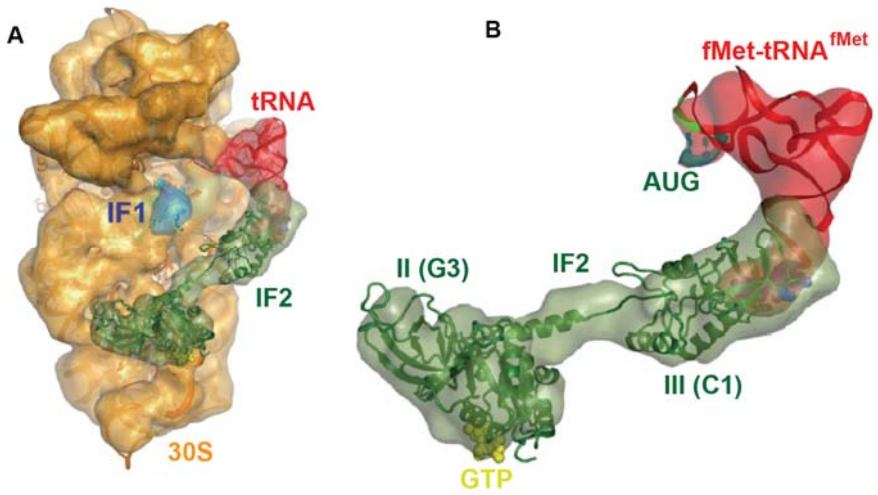

Figure 1.14 Cryo-EM structures of a 30S "IC". (A) The 30S "IC" includes a 30S ribosomal subunit, mRNA, IF1, IF2(GTP), and fMet-tRNA^{fMet} from *Thermus thermophilus*. **(B)** Rotation of the IF2-fMet-tRNA^{fMet} sub-complex relative to (A). Those domains of IF2 that were visible in the cryo-EM density are labeled. Figure from Ref. [80].

The 50S subunit-joining step is a major checkpoint for translational regulation. Efficient conversion of a 30S IC into the 70S IC is critical for entry into further steps in the translation cycle, and thus efficient translation of a particular mRNA into its corresponding polypeptide. The 70S IC is composed of an mRNA-bound 70S ribosome containing fMet-tRNA^{fMet} in the P site of both the 30S subunit and 50S subunit (that is, in the classical P/P configuration). Conversion of a 30S IC to a 70S IC involves the joining of a 50S subunit to a 30S IC, GTP hydrolysis by IF2, and the dissociation of the three IFs. The temporal organization of all of these steps remains unknown and widely debated [63, 118, 130, 131, 150, 151], though it has been established that GTP hydrolysis requires 50S subunit docking. The molecular consequences of GTP hydrolysis, however, remain poorly defined.

Three cryo-EM structures of intermediates along the 30S IC → 70S IC pathway have been solved. Two of these structures contain *T. thermophilus* 30S and 50S subunits, fMet-tRNA^{fMet}, mRNA, and IF2 with (a) the non-hydrolyzable GTP analog GMPPCP, or (b) GDP [48]. The other structure contained *E. coli* 30S and 50S subunits, IF1, IF2(GDPNP), IF3, mRNA, and fMet-tRNA^{fMet} [47]. Thus, the first two structures represent a pseudo 70S IC (a) before GTP hydrolysis and (b) after GTP hydrolysis, but before IF2 release. Both of these structures represent

the 70S IC after release of IF1 and IF3. The third cryo-EM structure represents a state before GTP hydrolysis and release of any of the IFs. With respect to the former two pseudo 70S ICs prepared in the absence of IF1 and IF3, the conformation of the 70S IC may not be the same, or even similar, to a 70S IC which has undergone IF-mediated subunit association [54].

As with the 30S IC cryo-EM structure, these 70S IC-like structures offer insight into the conformations of IF2 and fMet-tRNA^{fMet} within a 70S IC. In the *T. thermophilus* 70S, fMet-tRNA^{fMet}, mRNA, IF2(GMPPCP/GDP) structures, IF2 interacts with rRNA in both the 30S and 50S subunits. It is close to helix H89 of the 23S rRNA and the GTPase center of the 50S subunit, while it is close to helices h5, h15 and h17 of the 16S rRNA and protein S12 of the 30S subunit. IF2's G-domain and domain III interact with the 50S, while its domain II and NTD interact with the 30S subunit. In contrast to what was seen in the 30S IC cryo-EM structure, IF2 contacts the D-loop of fMet-tRNA^{fMet}, not the acceptor arm. Both IF2 and the ribosome undergo significant conformational rearrangements upon GTP hydrolysis, as evidenced by comparison of the GMPPCP and GDP structures. IF2 has fewer interactions within the ribosome in the GDP-bound state, suggesting a conformational state of IF2 just prior to its dissociation. In contrast, before GTP hydrolysis, IF2 is involved in many interactions with both ribosomal subunits.

The *E. coli* 70S IC structure [47] shows both similarities and differences with the *T. thermophilus* structure. Both show similar IF2-ribosome interactions, but differ with respect to the interactions between IF2 and fMet-tRNA^{fMet}. In the *E. coli* structure, as in the *T. thermophilus* 30S IC-like structure, but unlike the *T. thermophilus* 70S IC-like structure, IF2 is in direct contact with the single-stranded CCA end of the acceptor arm of the tRNA. Also, IF1 interacts with domain II of IF2, unlike in the 30S IC structure, where no direct contacts were seen. A region of electron difference density near the tRNA exit (E) site (see Figure 1.1A) was tentatively assigned to IF3, in line with its 30S subunit location as determined by directed hydroxyl radical probing [109]. The density may indeed be due to IF3 as a number of rapid kinetic studies have shown that IF3 dissociates following 50S subunit association [63, 150], though the relative timing of its

dissociation with respect to 50S subunit association has been a point of contention in the field [63, 118, 131, 150, 152].

1.3.5 Toward a mechanistic understanding of IF3's role in regulating the fidelity of translation initiation – the motivation for my Ph.D. work

Translation initiation is a critical stage in gene expression as it sets the reading frame of the mRNA for protein synthesis and regulates which mRNAs enter translation elongation for rapid decoding. Initiation is tightly regulated by the three IFs, all of which are essential for cell viability and are required to ensure the speed and fidelity of the multi-step initiation process. IF3 regulates substrate selection by helping discriminate the codon-anticodon interaction, favoring start codon and initiator tRNA selection. It also has roles in regulating the conversion of the 30S IC to the 70S IC by permitting rapid 50S subunit joining only to fully assembled 30S ICs. IF3 is a two domain protein connected by a lysine-rich linker that is highly conserved across prokaryotes in both length and hydrophilicity [89, 94]. The integrity of the linker is required *in vivo*, and its flexibility and hydrophilicity are critical to its function [94]. The linker's flexibility allows IF3 to adopt a range of interdomain distances in solution, spanning ~30 to 65 Å [89], and also enables the two domains' stepwise association with and dissociation from the 30S subunit [114, 118]. This ordered binding indicates that the CTD and NTD have significantly different affinities for the ribosome and behave independently. The CTD of IF3, as well as the archaeal and eukaryotic orthologs a/eIF1 [17, 153-155], binds with high affinity to the rRNA on the small subunit near the P site and occupies the site where 50S subunit 23S rRNA helix 69 binds upon subunit joining, thus the CTD of IF3 sterically blocks formation of intersubunit bridge B2b [109, 156]. The NTD fails to bind to the 16S rRNA of the 30S subunit [121] but instead is likely to interact with ribosomal proteins [114, 157], though there are discrepancies among IF3-30S structural studies regarding its binding site [109, 115, 118]. IF3 has been crosslinked with numerous ribosomal proteins spanning a broad area of the 30S subunit interface [62, 98, 101, 103, 158-160], suggesting the possibility of multiple NTD binding sites. Although an earlier model proposed that the CTD performs all IF3 functions and the NTD and linker play accessory roles in stabilizing the

CTD on the 30S subunit [121], the identification of loss-of-function point mutations in each domain clearly indicate that both domains as well as the linker are required for IF3's function [113, 142]. The two IF3 domains' ordered binding to, and dissociation from, the 30S subunit, as well as the requirement for CTD rearrangement for stable 50S subunit joining and the probable existence of multiple NTD binding sites, suggest that IF3 may be conformationally dynamic on the 30S subunit and that the relative position of the two domains of IF3 may be critical to its function.

Fluorescence resonance energy transfer (FRET) can serve as a spectroscopic ruler due to its sensitivity to the distance between two fluorophores within the biophysically relevant window of 1-10 nm. Thus, it should be possible to monitor relative changes in the position of 30S subunit-bound IF3's NTD and CTD by FRET between fluorophores attached to each domain. The identity of each FRET state and the frequency with which each state is sampled, along with the presence and behavior of any conformational dynamics, are mechanistically important. Unfortunately, these dynamics and the unique interdomain distance information from each 30S subunit-bound IF3 molecule in a heterogeneous sample are lost in an ensemble experiment as only the average FRET efficiency of the population is captured (see Figure 1.15). Single-molecule FRET (smFRET), however, enables one to capture this information for each molecule in a population and identify sub-populations, rarely sampled FRET states, and any conformational dynamics exhibited by individual molecules within the population. Monitoring the relative interdomain distance of 30S subunit-bound IF3's NTD and CTD with smFRET on a variety of 30S ICs (i.e. in the presence and absence of IF1, IF2, and tRNA) enables mechanistically important measurements on the range of interdomain distances sampled as a function of the 30S IC identity. Any interdomain dynamics on the hundreds of milliseconds time scale can also be measured as a function of 30S IC identity and further elucidate the molecular mechanism underlying IF3's function in translation initiation. Thus, smFRET was employed along with a biochemically active, dual fluorescently labeled IF3 and highly purified *in vitro* translation system to investigate the range of interdomain distances sampled by IF3 on 30S subunits, as well as any

interconversions between these states, with the aim of elucidating the molecular mechanism of IF3's relative interdomain distance with its function in regulating translation initiation.

1.4 Fluorescence Resonance Energy Transfer (FRET) – a powerful tool for measuring macromolecular conformational changes

Fluorescence (or Förster, after Theodor Förster, who proposed the theory [161]) resonance energy transfer (FRET) is a non-radiative energy transfer process that occurs between an electronically excited fluorophore (donor) and a second chromophore (acceptor), frequently another fluorophore, that are in close proximity [162]. The fluorescence emission spectrum of the donor must overlap with the absorption spectrum of the acceptor fluorophore, and the two fluorophores must be within a minimal spatial radius (usually 1-10 nm), for the donor to transfer its excitation energy to the acceptor through a long-range, induced-dipole induced-dipole interaction. Because of this dipole-dipole mechanism, the efficiency of resonance energy transfer is highly sensitive to the separation distance between the two dipoles (i.e. the fluorophores) and the ability of a donor fluorophore to transfer its excitation energy to an acceptor fluorophore decreases sharply with increasing distance between the molecules (Figure 1.15), as evidenced by the dependence of FRET efficiency (*E*) on the inverse sixth power of the separation distance (*R*) (Eqn. 1.1).

$$E = \frac{1}{1 + \left(\frac{R}{R_0}\right)^6} \qquad [1.1]$$

The distance dependence of FRET makes it an excellent tool for monitoring relative distance changes between fluorophores on the nanometer scale, with applications ranging from macromolecular conformational changes [163-169] to the oligomerization of membrane proteins in a living cell [170, 171] to protein folding [172-176], and RNA folding and catalysis [177-187]. Absolute distance measurements, however, should be interpreted with great care, for reasons described below.

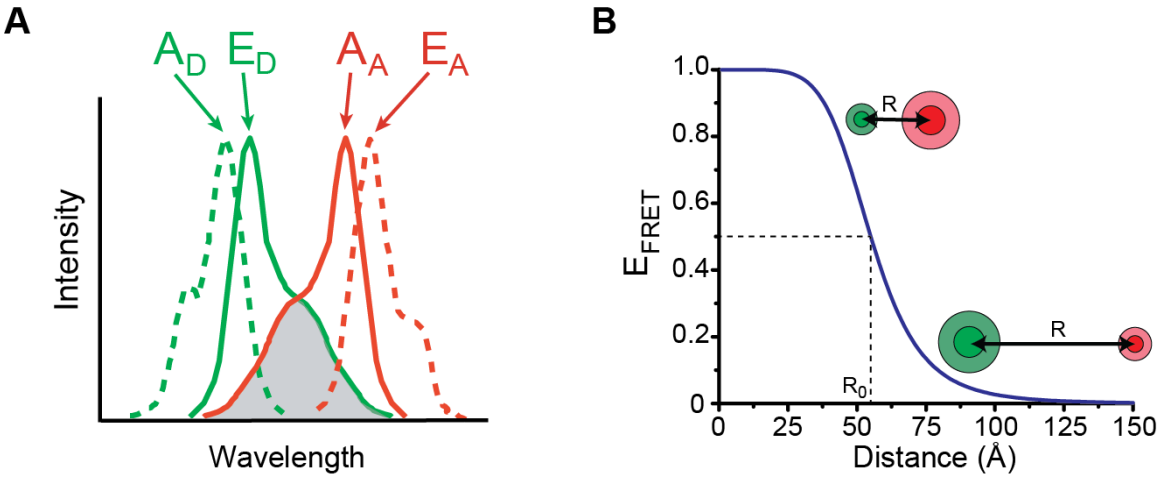

Figure 1.15 Fluorescence Resonance Energy Transfer. A. Efficient FRET requires significant spectral overlap between a donor fluorophore's emission spectrum (E_D) and an acceptor fluorophore's absorption spectrum (A_A). **B.** FRET with the fluorophore pair Cy3 and Cy5 is sensitive to distance changes between 35 and 65 Å. Figure adapted from Ref. [188].

Along with the inter-fluorophore separation, R, FRET efficiency is also dependent on R_0, the Förster distance, which depends on the particular donor-acceptor pair and is a function of a number of other parameters. It is defined as the donor-acceptor separation distance when transfer efficiency is 0.5 (Eqn. 1.2).

$$R_0 = \left(\frac{9000(\ln\ 10)\kappa^2 Q_D J(\lambda)}{128\pi^5 N n^4} \right)^{-6}, \qquad [1.2]$$

where κ^2 describes the relative orientation of the transition dipoles of the donor and acceptor, Q_D is the quantum yield of the donor in the absence of the acceptor, $J(\lambda)$ is the overlap integral, which is a function of the degree of spectral overlap between the donor emission and the acceptor absorption, and n is the refractive index of the medium, which is typically assumed to be n = 1.4 for biomolecules in aqueous solution [162]. Both Q_D and $J(\lambda)$ depend on the environment of the fluorophores, thus R_0 should be determined under the same conditions as those employed in the FRET experiments. Of all these variables, κ^2 is the most difficult to experimentally determine. κ^2 can take on a value from 0 to 4, with a value of 1 corresponding to parallel transition dipoles, and a value of 4 resulting from dipoles that are both parallel and collinear [162].

Fluorescence anisotropy measurements of the donor and acceptor fluorophores can allow limits to be determined for κ^2 variation [162]. If the two probes undergo unrestricted isotropic motion, and give fundamental fluorescence anisotropy, *r*, values for both fluorophores less than 0.2, then κ^2 is approximately equal to 2/3. Frequently, Förster distances are calculated using an assumed value of 2/3 for κ^2. This is generally considered acceptable if only relative, not absolute, changes in distance are measured [189]. For absolute distance measurements, it is critical that κ^2, and the other R_0 parameters, be precisely determined.

The efficiency of FRET can be estimated from the ratio of the acceptor's emission intensity (I_A) to the total emission intensity ($I_A + I_D$) (Eqn 1.3).

$$E_{FRET} = \frac{I_A}{I_A + I_D} \qquad [1.3]$$

One of FRET's advantages stems from it being a ratiometric method, making it largely immune to instrumental noise and drift.

1.4.1 A single-molecule approach to measuring conformational dynamics

Fluorescence- and force-based measurements of single biological molecules have become immensely popular over the past couple of decades due to numerous advantages in taking a single-molecule approach to measuring various dynamic processes [190-193]. smFRET has proven to be one of the most widely implemented single-molecule techniques. The advantages of smFRET over ensemble FRET stem from the ability to probe biological events directly without the population averaging that arises in ensemble studies, or the need to synchronize complex biochemical reactions (see Figure 1.16 and the figure caption for details). Distributions of distances between a FRET pair within a heterogeneous population, rather than an average distance, can be determined with smFRET. It is an especially powerful tool for monitoring the dynamics of individual, surface-immobilized molecules for extended timescales (milliseconds to minutes) – timescales especially relevant for translation initiation. This enables one to observe rare conformational transitions and non-accumulating reaction intermediates,

events that are hard to detect with ensemble FRET, as well as to probe the detailed kinetics of structural changes without the need for synchronization.

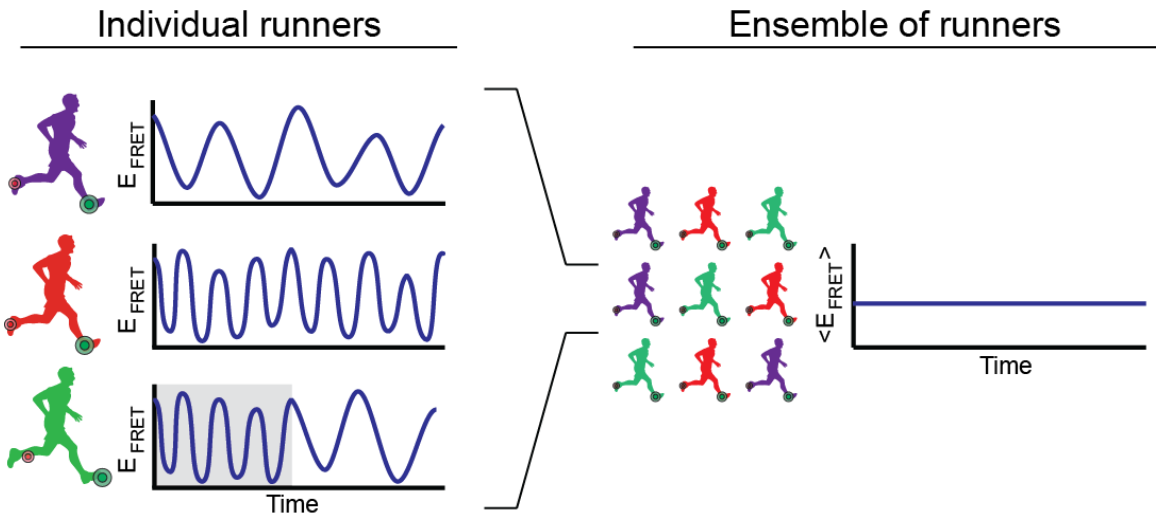

Figure 1.16 FRET at the single-molecule and ensemble levels. A macroscopic example of the power of single-molecule observations is cartooned here. FRET monitored between a donor-acceptor fluorophore pair attached to a runner's shoes can measure the runner's stride length and running rate. Individual runners within a population of marathoners vary in their running speed and step size both among each other, and within each individual's overall marathon run (e.g. bottom runner alternates between fast (gray box) and slow speeds). This heterogeneity can be captured by monitoring each runner individually (smFRET), while this information is lost by measuring the average speed of the ensemble of runners over the whole marathon. Figure prepared by Prof. Ruben L. Gonzalez, Jr. and found in Ref. [188].

1.4.2 Technical aspects of smFRET

Experimentally, smFRET can be employed using fluorescently labeled molecules that are either freely diffusing or surface-immobilized. The former are restricted to a diffusion-limited observation time window, making them appropriate only for very fast (<10 ms) phenomena or equilibrium/thermodynamic measurements. Surface immobilization of fluorescently labeled biomolecules, employed in my studies, extends the observation window to a point where it is now only limited by the rate of fluorophore photobleaching.

Molecules can be specifically immobilized on quartz microscope slide surfaces which have been coated with either biotinylated bovine serum albumin (BSA) or biotinylated polyethylene glycol (PEG), and then streptavidin [189, 194, 195]. The choice of BSA or PEG can be determined by the overall charge of the macromolecule of interest [195]. DNA- and RNA-only studies can be done with BSA-biotin surfaces since quartz, BSA, and streptavidin are all

negatively charged, preventing binding to the surface except at biotin attachment points. PEG surfaces are better for protein studies, as BSA surfaces can be quite adhesive to positively charged regions of proteins [195]. Biotin-PEG surfaces were used for all the work described in this thesis, and biotin-PEG was mixed with non-biotin-PEG solution at 0.1% final concentration to ensure spatially well-separated biotin anchors on the quartz surface for ribosome immobilization.

smFRET is typically performed using two main classes of fluorescence microscopes: (1) A confocal or near-field scanning optical microscope with photomultiplier tube (PMT) or avalanche photodiode (APD) single element point detection, or (2) a wide-field microscope with a two-dimensional detector such as an electron-multiplying charge coupled device (EMCCD) camera [194]. The latter is advantageous due to its high-throughput sampling, allowing hundreds of single molecules to be detected simultaneously, though its time resolution (>1 ms) and sensitivity are not as good as the point detection route. Wide-field microscopy can be performed via epi-illumination or evanescent field excitation. Again, the latter is advantageous since the evanescent field of excitation light extends only ~100-200 nm from the interface where the sample is bound, thus greatly reducing background fluorescence. An evanescent field is generated by total internal reflection (TIR) of the excitation light at the quartz-water interface. TIR requires an incident angle of the laser excitation source larger than the critical angle for the interface. This angle can be achieved using either a prism or the edge of a high numerical aperture (NA) objective (NA $\geq$ 1.4). In prism-based TIR fluorescence microscopy, employed here, an inverted microscope is used to hold the sample flowcell with a fused silica prism on top (see Figure 1.17). The incident laser beam is focused through a long focal length lens, enters the prism, passes through refractive index-matching oil and is internally reflected at the quartz-water interface. The emitted fluorescence signal from the surface-immobilized molecules is collected through the objective (1.2 NA, 60x magnification), below the sample. Donor and acceptor fluorescence emission is wavelength separated onto two halves of an EMCCD camera using a dichroic mirror and bandpass filters, allowing the two images to be obtained simultaneously [194].

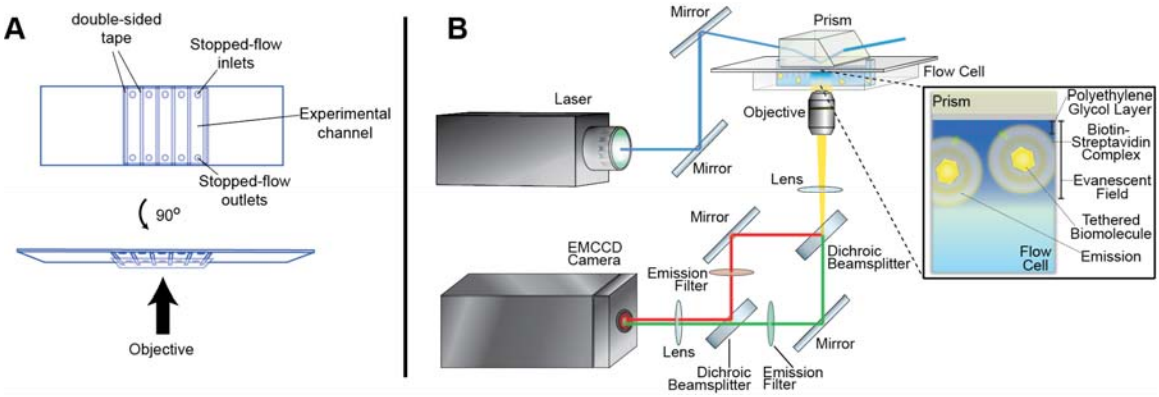

Figure 1.17 Experimental set-up for TIRFM-based smFRET. A. A microfluidic flowcell. Flowcells are prepared with a quartz microscope slide that has been passivated with biotinylated PEG, a glass coverslip, and thin strips of double-sided tape. **B.** Schematic of a prism-based TIRF microscope for single-molecule fluorescence imaging. The inset shows an enlarged view of a surface-tethered, fluorescently labeled biomolecule illuminated by the evanescent field produced at the interface between the quartz surface and the aqueous solution. Figure prepared by Mr. Colin Kinz-Thompson and found in Ref. [196].

1.4.3 An overview of smFRET data analysis

Donor (Cy3) and acceptor (Cy5) fluorophore emission intensities are collected simultaneously on two halves of a capacitor array within an EMCCD camera. Fluorescence intensity versus time trajectories for each Cy3 and Cy5 molecule are obtained by first identifying the fluorescent regions in an image that correspond to single molecules. This initial step in the data analysis procedure is performed using the MetaMorph software package (Molecular Dynamics, Inc.). Regions corresponding to single Cy3 and Cy5 molecules are identified by setting a minimum intensity threshold significantly higher than the background signal and then selecting the regions that lie above this threshold intensity. Cy3 and Cy5 emission intensities are co-localized by first identifying the regions corresponding to Cy5 fluorophores, followed by the transfer of the coordinates of these regions to the other half of the image where they overlay with Cy3 emission intensities. The Cy3 and Cy5 intensity versus frame number trajectories are plotted for each of the 200-400 regions within the image, where the time between frames is determined by the time resolution of the experiment (typically 100 msec). Each of these fluorescence versus time trajectories are visually inspected for the presence of characteristic smFRET and single-fluorophore behavior: (i) the Cy3 and Cy5 intensities are within the range expected for single

fluorophores, (ii) fluctuations of the Cy3 and Cy5 intensities exhibit anticorrelation in time, (iii) the Cy3 and Cy5 trajectories exhibit single-step photobleaching during the observation period (usually 1-2 minutes). Following this selection process, the Cy3 and Cy5 trajectories are baseline corrected using the average intensity of the last 25 frames of each trajectory. The Cy5 trajectories are also corrected for the bleed through from Cy3 emission into the Cy5 channel that results from imperfect emission filters. This value was experimentally determined to be ~7% for our TIRFM [197], therefore 7% of the baseline-corrected Cy3 intensity is subtracted from the raw Cy5 intensity at every time point in the trace before the Cy5 trajectory is baseline corrected. FRET efficiency is then calculated as the ratio of Cy5 fluorescence emission intensity over the sum of the total emission intensities from Cy3 and Cy5 (equation 1.3).

Figure 1.18 depicts representative plots and histograms obtained in the course of smFRET data analysis. After plotting the fluorescence intensity versus time trajectory (Figure 1.18A), and the corresponding smFRET versus time trajectory (Figure 1.18B) for each identified single molecule, all of these trajectories are combined to plot a one-dimensional FRET efficiency histogram (Figure 1.18C). These histograms are useful in obtaining an initial estimate of the number of FRET states present in a data set, as well as the relative occupancy of each FRET state. The histogram is fit with multiple Gaussian distributions to determine the mean and standard deviation of the Gaussian representing each FRET state. The standard deviation is determined as the full width at half maximum (FWHM) of each Gaussian distribution. An effective means to observe how the relative occupancy of a FRET state changes over time is through a two-dimensional surface contour plot of the time evolution of population FRET (Figure 1.18D).

The raw, noisy smFRET trajectories are then idealized using a hidden Markov model in order to more rigorously identify the number of conformational states underlying the observed FRET signals and their FRET efficiency values (Figure 1.18E) [198-201]. A Markov model describes a discrete stochastic process in which the condition of the present state is independent of any past or future state of the system. A hidden Markov model describes a process in which an observed time series is conditionally dependent on a hidden discrete state variable (i.e.

conformational state) [199]. Trajectory idealization identifies the number of states and transitions between these states, enabling the measurement of the dwell time spent in each FRET state prior to transitioning to another FRET state. Histograms of these dwell times can be fit to single- or multi-exponential functions to determine the mean lifetime (τ) spent in a state between transitions (Figure 1.18F). The transition rate is calculated as the reciprocal of the mean lifetime (1/τ). Taken together, these data analysis approaches are a powerful way to extract information about the number of conformational states sampled by single molecules within a population, as well as the transition rates between these states.

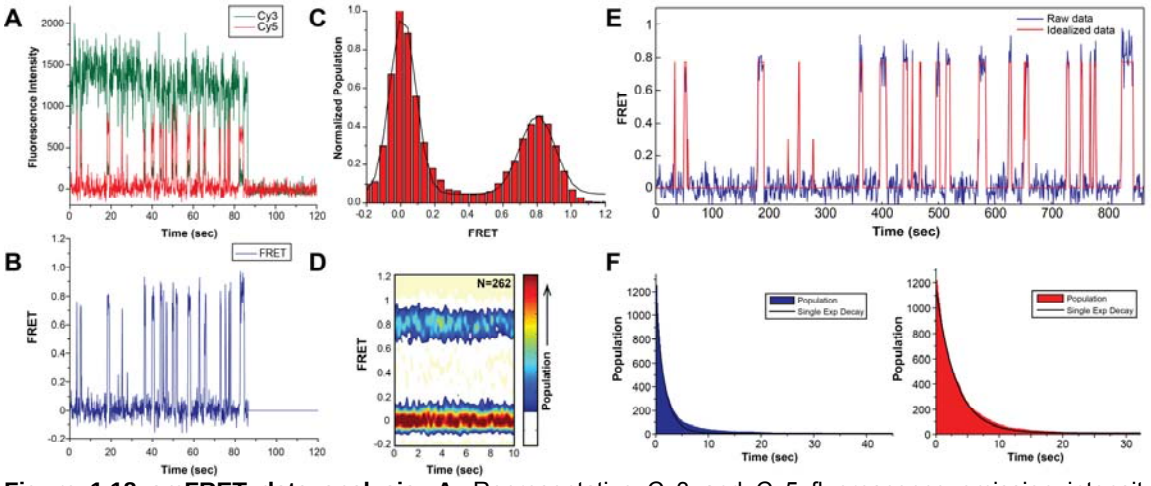

Figure 1.18 smFRET data analysis. A. Representative Cy3 and Cy5 fluorescence emission intensity versus time trajectories are shown in green and red, respectively. **B.** The corresponding smFRET versus time trajectory (FRET = $I_{Cy5}/(I_{Cy5}+I_{Cy3})$). **C.** One dimensional FRET efficiency histogram. There are 30 bins in the range of -0.2 to 1.2 FRET efficiencies (bin size = 0.047 FRET). The histogram is normalized to the most populated bin. **D.** Two dimensional surface contour plot of the time evolution of population FRET containing all (N=262) individual smFRET trajectories in the data set. The contours are plotted from tan (lowest population) to red (highest population). **E.** Hidden Markov modeling of smFRET trajectories. The raw data was idealized with the vbFRET software package [199]. **F.** Lifetime analysis. Dwell times spent at one FRET state before transitioning to another FRET state are plotted as one-dimensional histograms and fitted with an exponential decay function. Dwell times spent in the non-zero FRET state before transitioning to the zero FRET state (left) and dwell times spent in the zero FRET state before transitioning to the non-zero FRET state (right) The data are normally well described by a single exponential decay of the form $A_1(\exp(-t/\tau_1))$, where τ_1 is the lifetime of the corresponding FRET state. The transition rate is calculated as $1/\tau_1$. The representative lifetime analysis is exemplified with IF3(Cy3)-IF1(Cy5) data.

1.5 Summary and thesis overview

Translation initiation is a complicated, multi-step assembly pathway that is a key regulatory checkpoint in gene expression. Initiation factors 1, 2, and 3 play critical roles in controlling the rate and fidelity of initiation, yet much remains unknown about the mechanistic

details of their regulatory roles. The conformational dynamics of each IF, as well as of the ribosomal initiation complexes (30S pre-IC, 30S IC, 70S IC), likely play important roles in their functions, and may be critical to a better understanding of the regulation of translation initiation. The conformational and compositional heterogeneities inherently present in ribosomal complexes during translation initiation make smFRET an ideal technique for revealing mechanistically important subpopulations obscured in ensemble studies. smFRET has been used extensively in the past few years to investigate the aa-tRNA selection and translocation steps of translation elongation (reviewed in [202]). Far fewer single-molecule investigations into translation initiation have been performed. In our research group, smFRET has been used investigate the role of IF2 dynamics in regulating translation initiation [203], and in other labs it has been used to investigate inter-ribosomal subunit dynamics and the role of GTP hydrolysis by IF2 during late events in translation initiation [204].

Chapter Two of this thesis describes investigations on the conformational dynamics of 30S subunit-bound IF3 and how these dynamics are affected by the presence of IF1 and/or IF2, as well as the presence and identity of the P-site start codon and aa-tRNA. These results provide novel insight into the ribosome-bound conformational state and dynamics of IF3, especially as a function of the identity of the 30S P site substrates. They also reveal a role for IF3's two-domain nature in signaling 30S IC assembly. In Chapter Three, the role of the conformational dynamics of 30S-subunit bound IF3 in regulating 50S subunit joining following 30S IC assembly is investigated. Some preliminary results are presented, along with potential future directions for the project. Chapter Four presents measurements of IF1's 30S subunit association and dissociation kinetics investigated through the use of an IF1-IF3 smFRET signal. These results provide insight into IF1's role in the regulation of translation initiation by demonstrating its stabilizing effect on IF2. Finally, detailed materials and methods for all of the experiments described herein are found in Chapter Five.

1.6 References

1. Rodnina, M.V. and W. Wintermeyer, *Recent mechanistic insights into eukaryotic ribosomes.* Curr Opin Cell Biol, 2009. **21**(3): p. 435-43.

2. Alberts, B., A. Johnson, J. Lewis, M. Raff, K. Roberts, and P. Walter, *Molecular Biology of the Cell*. 5th ed2008, New York: Garland Science.

3. Fei, J., J. Wang, S.H. Sternberg, D.D. MacDougall, M.M. Elvekrog, D.K. Pulukkunat, M.T. Englander, and R.L. Gonzalez, Jr., *A highly purified, fluorescently labeled in vitro translation system for single-molecule studies of protein synthesis.* Methods Enzymol, 2010. **472**: p. 221-59.

4. Spirin, A.S., *Ribosome as a molecular machine.* FEBS Lett, 2002. **514**(1): p. 2-10.

5. Palmiter, R.D., *Differential Rates of Initiation on Conalbumin and Ovalbumin Messenger-Ribonucleic-Acid in Reticulocyte Lysates.* Journal of Biological Chemistry, 1974. **249**(21): p. 6779-6787.

6. Lodish, H.F. and M. Jacobsen, *Regulation of Hemoglobin Synthesis - Equal Rates of Translation and Termination of Alpha-Globin, and Beta-Globin Chains.* Journal of Biological Chemistry, 1972. **247**(11): p. 3622-&.

7. Walden, W.E., T. Godefroycolburn, and R.E. Thach, *The Role of Messenger-Rna Competition in Regulating Translation .1. Demonstration of Competition Invivo.* Journal of Biological Chemistry, 1981. **256**(22): p. 1739-1746.

8. Mathews, M.B., N. Sonenberg, and J.W. Hershey, *Origins and Principles of Translational Control*, in *Translational Control in Biology and Medicine*, M.B. Mathews, N. Sonenberg, and J.W. Hershey, Editors. 2007, Cold Spring Harbor Laboratory Press. p. 1-40.

9. Simonetti, A., S. Marzi, L. Jenner, A. Myasnikov, P. Romby, G. Yusupova, B.P. Klaholz, and M. Yusupov, *A structural view of translation initiation in bacteria.* Cell Mol Life Sci, 2009. **66**(3): p. 423-36.

10. Boelens, R. and C.O. Gualerzi, *Structure and function of bacterial initiation factors.* Curr Protein Pept Sci, 2002. **3**(1): p. 107-19.

11. Laursen, B.S., H.P. Sorensen, K.K. Mortensen, and H.U. Sperling-Petersen, *Initiation of protein synthesis in bacteria.* Microbiol Mol Biol Rev, 2005. **69**(1): p. 101-23.

12. Sonenberg, N. and A.G. Hinnebusch, *New modes of translational control in development, behavior, and disease.* Mol Cell, 2007. **28**(5): p. 721-9.

13. Sonenberg, N. and A.G. Hinnebusch, *Regulation of translation initiation in eukaryotes: mechanisms and biological targets.* Cell, 2009. **136**(4): p. 731-45.

14. Kyrpides, N.C. and C.R. Woese, *Universally conserved translation initiation factors.* Proc Natl Acad Sci U S A, 1998. **95**(1): p. 224-8.

15. Roll-Mecak, A., B.S. Shin, T.E. Dever, and S.K. Burley, *Engaging the ribosome: universal IFs of translation.* Trends Biochem Sci, 2001. **26**(12): p. 705-9.

16. Lee, J.H., S.K. Choi, A. Roll-Mecak, S.K. Burley, and T.E. Dever, *Universal conservation in translation initiation revealed by human and archaeal homologs of bacterial translation initiation factor IF2.* Proc Natl Acad Sci U S A, 1999. **96**(8): p. 4342-7.

17. Lomakin, I.B., N.E. Shirokikh, M.M. Yusupov, C.U. Hellen, and T.V. Pestova, *The fidelity of translation initiation: reciprocal activities of eIF1, IF3 and YciH.* EMBO J, 2006. **25**(1): p. 196-210.

18. Pestova, T.V. and V.G. Kolupaeva, *The roles of individual eukaryotic translation initiation factors in ribosomal scanning and initiation codon selection.* Genes Dev, 2002. **16**(22): p. 2906-22.

19. Wimberly, B.T., D.E. Brodersen, W.M. Clemons, Jr., R.J. Morgan-Warren, A.P. Carter, C. Vonrhein, T. Hartsch, and V. Ramakrishnan, *Structure of the 30S ribosomal subunit.* Nature, 2000. **407**(6802): p. 327-39.

20. Schluenzen, F., A. Tocilj, R. Zarivach, J. Harms, M. Gluehmann, D. Janell, A. Bashan, H. Bartels, I. Agmon, F. Franceschi, and A. Yonath, *Structure of functionally activated small ribosomal subunit at 3.3 angstroms resolution.* Cell, 2000. **102**(5): p. 615-23.

21. Yusupov, M.M., G.Z. Yusupova, A. Baucom, K. Lieberman, T.N. Earnest, J.H. Cate, and H.F. Noller, *Crystal structure of the ribosome at 5.5 A resolution.* Science, 2001. **292**(5518): p. 883-96.

22. Gao, N. and J. Frank, *A library of RNA bridges.* Nat Chem Biol, 2006. **2**(5): p. 231-2.

23. Cannone, J.J., S. Subramanian, M.N. Schnare, J.R. Collett, L.M. D'Souza, Y. Du, B. Feng, N. Lin, L.V. Madabusi, K.M. MÜller, N. Pande, Z. Shang, N. Yu, and R.R. Gutell, *The Comparative RNA Web (CRW) Site: An Online Database of Comparative Sequence and Structure Information for Ribosomal, Intron, and Other RNAs.* BioMed Central Bioinformatics, 2002. **3**(15).

24. Sykes, M.T. and J.R. Williamson, *A Complex Assembly Landscape for the 30S Ribosomal Subunit.* Annual Review of Biophysics, 2009. **38**: p. 197-215.

25. DeLano, W., *The PyMOL Molecular Graphics System, Version 1.3, Schrödinger, LLC.* 2008.

26. Liiv, A. and M. O'Connor, *Mutations in the intersubunit bridge regions of 23 S rRNA.* J Biol Chem, 2006. **281**(40): p. 29850-62.

27. 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. 1342-6.

28. Steitz, J.A. and K. Jakes, *How ribosomes select initiator regions in mRNA: base pair formation between the 3' terminus of 16S rRNA and the mRNA during initiation of protein synthesis in Escherichia coli.* Proc Natl Acad Sci U S A, 1975. **72**(12): p. 4734-8.

29. Gold, L. and G.D. Stormo, *High-level translation initiation.* Methods Enzymol, 1990. **185**: p. 89-93.

30. Ma, J., A. Campbell, and S. Karlin, *Correlations between Shine-Dalgarno sequences and gene features such as predicted expression levels and operon structures.* J Bacteriol, 2002. **184**(20): p. 5733-45.

31. Sacerdot, C., G. Fayat, P. Dessen, M. Springer, J.A. Plumbridge, M. Grunberg-Manago, and S. Blanquet, *Sequence of a 1.26-kb DNA fragment containing the structural gene for E.coli initiation factor IF3: presence of an AUU initiator codon.* EMBO J, 1982. **1**(3): p. 311-5.

32. Chalut, C. and J.M. Egly, *AUC is used as a start codon in Escherichia coli.* Gene, 1995. **156**(1): p. 43-5.

33. Binns, N. and M. Masters, *Expression of the Escherichia coli pcnB gene is translationally limited using an inefficient start codon: a second chromosomal example of translation initiated at AUU.* Mol Microbiol, 2002. **44**(5): p. 1287-98.

34. Vellanoweth, R.L. and J.C. Rabinowitz, *The influence of ribosome-binding-site elements on translational efficiency in Bacillus subtilis and Escherichia coli in vivo.* Mol Microbiol, 1992. **6**(9): p. 1105-14.

35. Studer, S.M. and S. Joseph, *Unfolding of mRNA secondary structure by the bacterial translation initiation complex.* Mol Cell, 2006. **22**(1): p. 105-15.

36. Schlax, P.J. and D.J. Worhunsky, *Translational repression mechanisms in prokaryotes.* Mol Microbiol, 2003. **48**(5): p. 1157-69.

37. Mandal, M. and R.R. Breaker, *Gene regulation by riboswitches.* Nat Rev Mol Cell Biol, 2004. **5**(6): p. 451-63.

38. Kozak, M., *Comparison of initiation of protein synthesis in procaryotes, eucaryotes, and organelles.* Microbiol Rev, 1983. **47**(1): p. 1-45.

39. RajBhandary, U.L., *Initiator transfer RNAs.* J Bacteriol, 1994. **176**(3): p. 547-52.

40. Seong, B.L. and U.L. RajBhandary, *Mutants of Escherichia coli formylmethionine tRNA: a single base change enables initiator tRNA to act as an elongator in vitro.* Proc Natl Acad Sci U S A, 1987. **84**(24): p. 8859-63.

41. Mangroo, D., X.Q. Wu, and U.L. RajBhandary, *Escherichia coli initiator tRNA: structure-function relationships and interactions with the translational machinery.* Biochem Cell Biol, 1995. **73**(11-12): p. 1023-31.

42. Sprinzl, M., T. Hartmann, F. Meissner, J. Moll, and T. Vorderwulbecke, *Compilation of tRNA sequences and sequences of tRNA genes.* Nucleic Acids Res, 1987. **15 Suppl**: p. r53-188.

43. Mandal, N., D. Mangroo, J.J. Dalluge, J.A. McCloskey, and U.L. Rajbhandary, *Role of the three consecutive G:C base pairs conserved in the anticodon stem of initiator tRNAs in initiation of protein synthesis in Escherichia coli.* RNA, 1996. **2**(5): p. 473-82.

44. Lancaster, L. and H.F. Noller, *Involvement of 16S rRNA Nucleotides G1338 and A1339 in discrimination of initiator tRNA.* Molecular Cell, 2005. **20**(4): p. 623-632.

45. Selmer, M., C.M. Dunham, F.V.t. Murphy, A. Weixlbaumer, S. Petry, A.C. Kelley, J.R. Weir, and V. Ramakrishnan, *Structure of the 70S ribosome complexed with mRNA and tRNA.* Science, 2006. **313**(5795): p. 1935-42.

46. Wu, X.Q. and U.L. RajBhandary, *Effect of the amino acid attached to Escherichia coli initiator tRNA on its affinity for the initiation factor IF2 and on the IF2 dependence of its binding to the ribosome.* J Biol Chem, 1997. **272**(3): p. 1891-5.

47. Allen, G.S., A. Zavialov, R. Gursky, M. Ehrenberg, and J. Frank, *The cryo-EM structure of a translation initiation complex from Escherichia coli.* Cell, 2005. **121**(5): p. 703-12.

48. Myasnikov, A.G., S. Marzi, A. Simonetti, A.M. Giuliodori, C.O. Gualerzi, G. Yusupova, M. Yusupov, and B.P. Klaholz, *Conformational transition of initiation factor 2 from the GTP- to GDP-bound state visualized on the ribosome.* Nat Struct Mol Biol, 2005. **12**(12): p. 1145-9.

49. Hartz, D., J. Binkley, T. Hollingsworth, and L. Gold, *Domains of initiator tRNA and initiation codon crucial for initiator tRNA selection by Escherichia coli IF3.* Genes Dev, 1990. **4**(10): p. 1790-800.

50. Seong, B.L. and U.L. RajBhandary, *Escherichia coli formylmethionine tRNA: mutations in GGGCCC sequence conserved in anticodon stem of initiator tRNAs affect initiation of protein synthesis and conformation of anticodon loop.* Proc Natl Acad Sci U S A, 1987. **84**(2): p. 334-8.

51. Olsson, C.L., M. Graffe, M. Springer, and J.W. Hershey, *Physiological effects of translation initiation factor IF3 and ribosomal protein L20 limitation in Escherichia coli.* Mol Gen Genet, 1996. **250**(6): p. 705-14.

52. Cummings, H.S. and J.W. Hershey, *Translation initiation factor IF1 is essential for cell viability in Escherichia coli.* J Bacteriol, 1994. **176**(1): p. 198-205.

53. Laalami, S., H. Putzer, J.A. Plumbridge, and M. Grunberg-Manago, *A severely truncated form of translational initiation factor 2 supports growth of Escherichia coli.* J Mol Biol, 1991. **220**(2): p. 335-49.

54. Allen, G.S. and J. Frank, *Structural insights on the translation initiation complex: ghosts of a universal initiation complex.* Mol Microbiol, 2007. **63**(4): p. 941-50.

55. Sette, M., P. van Tilborg, R. Spurio, R. Kaptein, M. Paci, C.O. Gualerzi, and R. Boelens, *The structure of the translational initiation factor IF1 from E.coli contains an oligomer-binding motif.* EMBO J, 1997. **16**(6): p. 1436-43.

56. Gualerzi, C.O., R. Spurio, A. La Teana, R. Calogero, B. Celano, and C.L. Pon, *Site-directed mutagenesis of Escherichia coli translation initiation factor IF1. Identification of the amino acid involved in its ribosomal binding and recycling.* Protein Eng, 1989. **3**(2): p. 133-8.

57. Baan, R.A., J.J. Duijfjes, E. van Leerdam, P.H. van Knippenberg, and L. Bosch, *Specific in situ cleavage of 16S ribosomal RNA of Escherichia coli interferes with the function of initiation factor IF-1.* Proc Natl Acad Sci U S A, 1976. **73**(3): p. 702-6.

58. Celano, B., R.T. Pawlik, and C.O. Gualerzi, *Interaction of Escherichia coli translation-initiation factor IF-1 with ribosomes.* Eur J Biochem, 1988. **178**(2): p. 351-5.

59. Zucker, F.H. and J.W. Hershey, *Binding of Escherichia coli protein synthesis initiation factor IF1 to 30S ribosomal subunits measured by fluorescence polarization.* Biochemistry, 1986. **25**(12): p. 3682-90.

60. Carter, A.P., W.M. Clemons, Jr., D.E. Brodersen, R.J. Morgan-Warren, T. Hartsch, B.T. Wimberly, and V. Ramakrishnan, *Crystal structure of an initiation factor bound to the 30S ribosomal subunit.* Science, 2001. **291**(5503): p. 498-501.

61. Dahlquist, K.D. and J.D. Puglisi, *Interaction of translation initiation factor IF1 with the E. coli ribosomal A site.* J Mol Biol, 2000. **299**(1): p. 1-15.

62. Boileau, G., P. Butler, J.W. Hershey, and R.R. Traut, *Direct cross-links between initiation factors 1, 2, and 3 and ribosomal proteins promoted by 2-iminothiolane.* Biochemistry, 1983. **22**(13): p. 3162-70.

63. Milon, P., A.L. Konevega, C.O. Gualerzi, and M.V. Rodnina, *Kinetic checkpoint at a late step in translation initiation.* Mol Cell, 2008. **30**(6): p. 712-20.

64. Kung, H.F., C. Spears, T. Schulz, and H. Weissbach, *Studies on the in vitro synthesis of beta-galactosidase: necessary components in the ribosomal wash.* Arch Biochem Biophys, 1974. **162**(2): p. 578-84.

65. Hartz, D., D.S. McPheeters, and L. Gold, *Selection of the initiator tRNA by Escherichia coli initiation factors.* Genes Dev, 1989. **3**(12A): p. 1899-912.

66. Choi, S.K., D.S. Olsen, A. Roll-Mecak, A. Martung, K.L. Remo, S.K. Burley, A.G. Hinnebusch, and T.E. Dever, *Physical and functional interaction between the eukaryotic orthologs of prokaryotic translation initiation factors IF1 and IF2.* Mol Cell Biol, 2000. **20**(19): p. 7183-91.

67. Stringer, E.A., P. Sarkar, and U. Maitra, *Function of initiation factor 1 in the binding and release of initiation factor 2 from ribosomal initiation complexes in Escherichia coli.* J Biol Chem, 1977. **252**(5): p. 1739-44.

68. Gaur, R., D. Grasso, P.P. Datta, P.D. Krishna, G. Das, A. Spencer, R.K. Agrawal, L. Spremulli, and U. Varshney, *A single mammalian mitochondrial translation initiation factor functionally replaces two bacterial factors.* Mol Cell, 2008. **29**(2): p. 180-90.

69. Yassin, A.S., M.E. Haque, P.P. Datta, K. Elmore, N.K. Banavali, L.L. Spremulli, and R.K. Agrawal, *Insertion domain within mammalian mitochondrial translation initiation factor 2 serves the role of eubacterial initiation factor 1.* Proc Natl Acad Sci U S A, 2011. **108**(10): p. 3918-23.

70. Mortensen, K.K., E. Hajnsdorf, P. Regnier, and H.U. Sperling-Petersen, *Improved recombinant tandem expression of translation initiation factor IF2 in RNASE E deficient E. coli cells.* Biochem Biophys Res Commun, 1995. **214**(3): p. 1254-9.

71. Sacerdot, C., G. Vachon, S. Laalami, F. Morel-Deville, Y. Cenatiempo, and M. Grunberg-Manago, *Both forms of translational initiation factor IF2 (alpha and beta) are required for*

maximal growth of Escherichia coli. Evidence for two translational initiation codons for IF2 beta. J Mol Biol, 1992. **225**(1): p. 67-80.

72. Howe, J.G. and J.W. Hershey, *Initiation factor and ribosome levels are coordinately controlled in Escherichia coli growing at different rates.* J Biol Chem, 1983. **258**(3): p. 1954-9.

73. Roll-Mecak, A., C. Cao, T.E. Dever, and S.K. Burley, *X-Ray structures of the universal translation initiation factor IF2/eIF5B: conformational changes on GDP and GTP binding.* Cell, 2000. **103**(5): p. 781-92.

74. Mortensen, K.K., J. Kildsgaard, J.M. Moreno, S.A. Steffensen, J. Egebjerg, and H.U. Sperling-Petersen, *A six-domain structural model for Escherichia coli translation initiation factor IF2. Characterisation of twelve surface epitopes.* Biochem Mol Biol Int, 1998. **46**(5): p. 1027-41.

75. Laursen, B.S., A.C. Kjaergaard, K.K. Mortensen, D.W. Hoffman, and H.U. Sperling-Petersen, *The N-terminal domain (IF2N) of bacterial translation initiation factor IF2 is connected to the conserved C-terminal domains by a flexible linker.* Protein Sci, 2004. **13**(1): p. 230-9.

76. Caserta, E., J. Tomsic, R. Spurio, A. La Teana, C.L. Pon, and C.O. Gualerzi, *Translation initiation factor IF2 interacts with the 30 S ribosomal subunit via two separate binding sites.* J Mol Biol, 2006. **362**(4): p. 787-99.

77. Guenneugues, M., E. Caserta, L. Brandi, R. Spurio, S. Meunier, C.L. Pon, R. Boelens, and C.O. Gualerzi, *Mapping the fMet-tRNA(f)(Met) binding site of initiation factor IF2.* EMBO J, 2000. **19**(19): p. 5233-40.

78. Petersen, H.U., T. Roll, M. Grunberg-Manago, and B.F. Clark, *Specific interaction of initiation factor IF2 of E. coli with formylmethionyl-tRNA f Met.* Biochem Biophys Res Commun, 1979. **91**(3): p. 1068-74.

79. Weiel, J. and J.W. Hershey, *The binding of fluorescein-labeled protein synthesis initiation factor 2 to Escherichia coli 30 S ribosomal subunits determined by fluorescence polarization.* J Biol Chem, 1982. **257**(3): p. 1215-20.

80. Simonetti, A., S. Marzi, A.G. Myasnikov, A. Fabbretti, M. Yusupov, C.O. Gualerzi, and B.P. Klaholz, *Structure of the 30S translation initiation complex.* Nature, 2008. **455**(7211): p. 416-20.

81. Sacerdot, C., C. Chiaruttini, K. Engst, M. Graffe, M. Milet, N. Mathy, J. Dondon, and M. Springer, *The role of the AUU initiation codon in the negative feedback regulation of the gene for translation initiation factor IF3 in Escherichia coli.* Mol Microbiol, 1996. **21**(2): p. 331-46.

82. Gualerzi, C.O. and C.L. Pon, *Initiation of mRNA translation in prokaryotes.* Biochemistry, 1990. **29**(25): p. 5881-9.

83. Brauer, D. and B. Wittmann-Liebold, *The primary structure of the initiation factor IF-3 from Escherichia coli.* FEBS Lett, 1977. **79**(2): p. 269-75.

84. Fortier, P.L., J.M. Schmitter, C. Garcia, and F. Dardel, *The N-terminal half of initiation factor IF3 is folded as a stable independent domain.* Biochimie, 1994. **76**(5): p. 376-83.

85. Kycia, J.H., V. Biou, F. Shu, S.E. Gerchman, V. Graziano, and V. Ramakrishnan, *Prokaryotic translation initiation factor IF3 is an elongated protein consisting of two crystallizable domains.* Biochemistry, 1995. **34**(18): p. 6183-7.

86. Biou, V., F. Shu, and V. Ramakrishnan, *X-ray crystallography shows that translational initiation factor IF3 consists of two compact alpha/beta domains linked by an alpha-helix.* EMBO J, 1995. **14**(16): p. 4056-64.

87. Garcia, C., P.L. Fortier, S. Blanquet, J.Y. Lallemand, and F. Dardel, *1H and 15N resonance assignments and structure of the N-terminal domain of Escherichia coli initiation factor 3.* Eur J Biochem, 1995. **228**(2): p. 395-402.

88. Garcia, C., P.L. Fortier, S. Blanquet, J.Y. Lallemand, and F. Dardel, *Solution structure of the ribosome-binding domain of E. coli translation initiation factor IF3. Homology with the U1A protein of the eukaryotic spliceosome.* J Mol Biol, 1995. **254**(2): p. 247-59.

89. Moreau, M., E. de Cock, P.L. Fortier, C. Garcia, C. Albaret, S. Blanquet, J.Y. Lallemand, and F. Dardel, *Heteronuclear NMR studies of E. coli translation initiation factor IF3. Evidence that the inter-domain region is disordered in solution.* J Mol Biol, 1997. **266**(1): p. 15-22.

90. Hua, Y. and D.P. Raleigh, *On the global architecture of initiation factor IF3: a comparative study of the linker regions from the Escherichia coli protein and the Bacillus stearothermophilus protein.* J Mol Biol, 1998. **278**(4): p. 871-8.

91. Gualerzi, C.O., A. Fabbretti, L. Brandi, P. Milon, and C.L. Pon, *Role of the Initiation Factors in mRNA Start Site Selection and fMet-tRNA Recruitment by Bacterial Ribosomes.* Israel Journal of Chemistry, 2010. **50**(1): p. 80-94.

92. Lammi, M., C.L. Pon, and C.O. Gualerzi, *The NH2-terminal cleavage of Escherichia coli translational initiation factor IF3. A mechanism to control the intracellular level of the factor?* FEBS Lett, 1987. **215**(1): p. 115-21.

93. Altschul, S.F., W. Gish, W. Miller, E.W. Myers, and D.J. Lipman, *Basic local alignment search tool.* J Mol Biol, 1990. **215**(3): p. 403-10.

94. de Cock, E., M. Springer, and F. Dardel, *The interdomain linker of Escherichia coli initiation factor IF3: a possible trigger of translation initiation specificity.* Mol Microbiol, 1999. **32**(1): p. 193-202.

95. Guntert, P., W. Braun, and K. Wuthrich, *Efficient computation of three-dimensional protein structures in solution from nuclear magnetic resonance data using the program DIANA and the supporting programs CALIBA, HABAS and GLOMSA.* J Mol Biol, 1991. **217**(3): p. 517-30.

96. Larkin, M.A., G. Blackshields, N.P. Brown, R. Chenna, P.A. McGettigan, H. McWilliam, F. Valentin, I.M. Wallace, A. Wilm, R. Lopez, J.D. Thompson, T.J. Gibson, and D.G. Higgins, *Clustal W and Clustal X version 2.0.* Bioinformatics, 2007. **23**(21): p. 2947-8.

97. Goujon, M., H. McWilliam, W. Li, F. Valentin, S. Squizzato, J. Paern, and R. Lopez, *A new bioinformatics analysis tools framework at EMBL-EBI.* Nucleic Acids Res, 2010. **38**(Web Server issue): p. W695-9.

98. Heimark, R.L., L. Kahan, K. Johnston, J.W. Hershey, and R.R. Traut, *Cross-linking of initiation factor IF3 to proteins of the Escherichia coli 30 S ribosomal subunit.* J Mol Biol, 1976. **105**(2): p. 219-30.

99. Hawley, D.A., L.I. Slobin, and A.J. Wahba, *The mechanism of action of initiation factor 3 in protein synthesis. II. Association of the 30S ribosomal protein S12 with IF-3.* Biochem Biophys Res Commun, 1974. **61**(2): p. 544-50.

100. van Duin, J., C.G. Kurland, J. Dondon, and M. Grunberg-Manago, *Near neighbors of IF3 bound to 30S ribosomal subunits.* FEBS Lett, 1975. **59**(2): p. 287-90.

101. Cooperman, B.S., A. Expert-Bezancon, L. Kahan, J. Dondon, and M. Grunberg-Manago, *IF-3 crosslinking to Escherichia coli ribosomal 30 S subunits by three different light-dependent procedures: identification of 30 S proteins crosslinked to IF-3--utilization of a new two-stage crosslinking reagent, p-nitrobenzylmaleimide.* Arch Biochem Biophys, 1981. **208**(2): p. 554-62.

102. Pon, C.L., R.T. Pawlik, and C. Gualerzi, *The topographical localization of IF3 on Escherichia coli 30 S ribosomal subunits as a clue to its way of functioning.* FEBS Lett, 1982. **137**(2): p. 163-7.

103. MacKeen, L.A., L. Kahan, A.J. Wahba, and I. Schwartz, *Photochemical cross-linking of initiation factor-3 to Escherichia coli 30 S ribosomal subunits.* J Biol Chem, 1980. **255**(21): p. 10526-31.

104. Ehresmann, C., H. Moine, M. Mougel, J. Dondon, M. Grunberg-Manago, J.P. Ebel, and B. Ehresmann, *Cross-linking of initiation factor IF3 to Escherichia coli 30S ribosomal subunit by trans-diamminedichloroplatinum(II): characterization of two cross-linking sites in 16S rRNA; a possible way of functioning for IF3.* Nucleic Acids Res, 1986. **14**(12): p. 4803-21.

105. Tapprich, W.E., D.J. Goss, and A.E. Dahlberg, *Mutation at position 791 in Escherichia coli 16S ribosomal RNA affects processes involved in the initiation of protein synthesis.* Proc Natl Acad Sci U S A, 1989. **86**(13): p. 4927-31.

106. Santer, M., E. Bennett-Guerrero, S. Byahatti, S. Czarnecki, D. O'Connell, M. Meyer, J. Khoury, X. Cheng, I. Schwartz, and J. McLaughlin, *Base changes at position 792 of Escherichia coli 16S rRNA affect assembly of 70S ribosomes.* Proc Natl Acad Sci U S A, 1990. **87**(10): p. 3700-4.

107. Firpo, M.A., M.B. Connelly, D.J. Goss, and A.E. Dahlberg, *Mutations at two invariant nucleotides in the 3'-minor domain of Escherichia coli 16 S rRNA affecting translational initiation and initiation factor 3 function.* J Biol Chem, 1996. **271**(9): p. 4693-8.

108. Moazed, D., R.R. Samaha, C. Gualerzi, and H.F. Noller, *Specific protection of 16 S rRNA by translational initiation factors.* J Mol Biol, 1995. **248**(2): p. 207-10.

109. Dallas, A. and H.F. Noller, *Interaction of translation initiation factor 3 with the 30S ribosomal subunit.* Mol Cell, 2001. **8**(4): p. 855-64.

110. Bruhns, J. and C. Gualerzi, *Structure--function relationship in Escherichia coli initiation factors: role of tyrosine residues in ribosomal binding and functional activity of IF-3.* Biochemistry, 1980. **19**(8): p. 1670-6.

111. Ohsawa, H. and C. Gualerzi, *Structure-function relationship in Escherichia coli initiation factors. Identification of a lysine residue in the ribosomal binding site of initiation factor by site-specific chemical modification with pyridoxal phosphate.* J Biol Chem, 1981. **256**(10): p. 4905-12.

112. De Bellis, D., D. Liveris, D. Goss, S. Ringquist, and I. Schwartz, *Structure-function analysis of Escherichia coli translation initiation factor IF3: tyrosine 107 and lysine 110 are required for ribosome binding.* Biochemistry, 1992. **31**(48): p. 11984-90.

113. Sacerdot, C., E. de Cock, K. Engst, M. Graffe, F. Dardel, and M. Springer, *Mutations that alter initiation codon discrimination by Escherichia coli initiation factor IF3.* J Mol Biol, 1999. **288**(5): p. 803-10.

114. Sette, M., R. Spurio, P. van Tilborg, C.O. Gualerzi, and R. Boelens, *Identification of the ribosome binding sites of translation initiation factor IF3 by multidimensional heteronuclear NMR spectroscopy.* RNA, 1999. **5**(1): p. 82-92.

115. McCutcheon, J.P., R.K. Agrawal, S.M. Philips, R.A. Grassucci, S.E. Gerchman, W.M. Clemons, Jr., V. Ramakrishnan, and J. Frank, *Location of translational initiation factor IF3 on the small ribosomal subunit.* Proc Natl Acad Sci U S A, 1999. **96**(8): p. 4301-6.

116. Hermanson, G.T., *Bioconjugate Techniques*2008: Academic Press.

117. Pioletti, M., F. Schlunzen, J. Harms, R. Zarivach, M. Gluhmann, H. Avila, A. Bashan, H. Bartels, T. Auerbach, C. Jacobi, T. Hartsch, A. Yonath, and F. Franceschi, *Crystal structures of complexes of the small ribosomal subunit with tetracycline, edeine and IF3.* EMBO J, 2001. **20**(8): p. 1829-39.

118. Fabbretti, A., C.L. Pon, S.P. Hennelly, W.E. Hill, J.S. Lodmell, and C.O. Gualerzi, *The real-time path of translation factor IF3 onto and off the ribosome.* Mol Cell, 2007. **25**(2): p. 285-96.

119. Weiel, J. and J.W. Hershey, *Fluorescence polarization studies of the interaction of Escherichia coli protein synthesis initiation factor 3 with 30S ribosomal subunits.* Biochemistry, 1981. **20**(20): p. 5859-65.

120. Weiel, J. and J.W. Hershey, *Fluorescence polarization studies of the binding of fluorescein-labeled initiation factor IF3 to 30 S ribosomal subunits from Escherichia coli.* FEBS Lett, 1978. **87**(1): p. 103-6.

121. Petrelli, D., A. LaTeana, C. Garofalo, R. Spurio, C.L. Pon, and C.O. Gualerzi, *Translation initiation factor IF3: two domains, five functions, one mechanism?* EMBO J, 2001. **20**(16): p. 4560-9.

122. Subramanian, A.R. and B.D. Davis, *Activity of initiation factor F3 in dissociating Escherichia coli ribosomes.* Nature, 1970. **228**(5278): p. 1273-5.

123. Subramanian, A.R., E.Z. Ron, and B.D. Davis, *A factor required for ribosome dissociation in Escherichia coli.* Proc Natl Acad Sci U S A, 1968. **61**(2): p. 761-7.

124. Miall, S.H. and T. Tamaoki, *Dissociation of Escherichia coli ribosomes. Role of initiation factors.* Biochemistry, 1972. **11**(25): p. 4826-30.

125. Gottleib, M., B.D. Davis, and R.C. Thompson, *Mechanism of dissociation of ribosomes of Escherichia coli by initiation factor IF-3.* Proc Natl Acad Sci U S A, 1975. **72**(11): p. 4238-42.

126. Chaires, J.B., C. Pande, and A. Wishnia, *The effect of initiation factor IF-3 on Escherichia coli ribosomal subunit association kinetics.* J Biol Chem, 1981. **256**(13): p. 6600-7.

127. Grunberg-Manago, M., P. Dessen, D. Pantaloni, T. Godefroy-Colburn, A.D. Wolfe, and J. Dondon, *Light-scattering studies showing the effect of initiation factors on the reversible dissociation of Escherichia coli ribosomes.* J Mol Biol, 1975. **94**(3): p. 461-78.

128. Naaktgeboren, N., K. Roobol, and H.O. Voorma, *The effect of initiation factor IF-1 on the dissociation of 70-S ribosomes of Escherichia coli.* Eur J Biochem, 1977. **72**(1): p. 49-56.

129. Antoun, A., M.Y. Pavlov, T. Tenson, and M. Ehrenberg, *Ribosome formation from subunits studied by stopped-flow and Rayleigh light scattering.* Biol. Proced. Online, 2004. **6**(1): p. 35-45.

130. Antoun, A., M.Y. Pavlov, M. Lovmar, and M. Ehrenberg, *How initiation factors tune the rate of initiation of protein synthesis in bacteria.* EMBO J, 2006. **25**(11): p. 2539-50.

131. Antoun, A., M.Y. Pavlov, M. Lovmar, and M. Ehrenberg, *How initiation factors maximize the accuracy of tRNA selection in initiation of bacterial protein synthesis.* Mol Cell, 2006. **23**(2): p. 183-93.

132. Pavlov, M.Y., A. Antoun, M. Lovmar, and M. Ehrenberg, *Complementary roles of initiation factor 1 and ribosome recycling factor in 70S ribosome splitting.* EMBO J, 2008. **27**(12): p. 1706-17.

133. Karimi, R., M.Y. Pavlov, R.H. Buckingham, and M. Ehrenberg, *Novel roles for classical factors at the interface between translation termination and initiation.* Mol Cell, 1999. **3**(5): p. 601-9.

134. Risuleo, G., C. Gualerzi, and C. Pon, *Specificity and properties of the destabilization, induced by initiation factor IF-3, of ternary complexes of the 30-S ribosomal subunit, aminoacyl-tRNA and polynucleotides.* Eur J Biochem, 1976. **67**(2): p. 603-13.

135. Pon, C.L. and C. Gualerzi, *Effect of initiation factor 3 binding on the 30S ribosomal subunits of Escherichia coli.* Proc Natl Acad Sci U S A, 1974. **71**(12): p. 4950-4.

136. Berkhout, B., C.J. van der Laken, and P.H. van Knippenberg, *Formylmethionyl-tRNA binding to 30 S ribosomes programmed with homopolynucleotides and the effect of translational initiation factor 3.* Biochim Biophys Acta, 1986. **866**(2-3): p. 144-53.

137. Gualerzi, C., C.L. Pon, and A. Kaji, *Initiation factor dependent release of aminoacyl-tRNAs from complexes of 30S ribosomal subunits, synthetic polynucleotide and aminoacyl tRNA.* Biochem Biophys Res Commun, 1971. **45**(5): p. 1312-9.

138. O'Connor, M., S.T. Gregory, U.L. Rajbhandary, and A.E. Dahlberg, *Altered discrimination of start codons and initiator tRNAs by mutant initiation factor 3.* RNA, 2001. **7**(7): p. 969-78.

139. Tedin, K., I. Moll, S. Grill, A. Resch, A. Graschopf, C.O. Gualerzi, and U. Blasi, *Translation initiation factor 3 antagonizes authentic start codon selection on leaderless mRNAs.* Mol Microbiol, 1999. **31**(1): p. 67-77.

140. Haggerty, T.J. and S.T. Lovett, *IF3-mediated suppression of a GUA initiation codon mutation in the recJ gene of Escherichia coli.* J Bacteriol, 1997. **179**(21): p. 6705-13.

141. Sussman, J.K., E.L. Simons, and R.W. Simons, *Escherichia coli translation initiation factor 3 discriminates the initiation codon in vivo.* Mol Microbiol, 1996. **21**(2): p. 347-60.

142. Maar, D., D. Liveris, J.K. Sussman, S. Ringquist, I. Moll, N. Heredia, A. Kil, U. Blasi, I. Schwartz, and R.W. Simons, *A single mutation in the IF3 N-terminal domain perturbs the fidelity of translation initiation at three levels.* J Mol Biol, 2008. **383**(5): p. 937-44.

143. Butler, J.S., M. Springer, J. Dondon, M. Graffe, and M. Grunberg-Manago, *Escherichia coli protein synthesis initiation factor IF3 controls its own gene expression at the translational level in vivo.* J Mol Biol, 1986. **192**(4): p. 767-80.

144. Brombach, M. and C.L. Pon, *The unusual translational initiation codon AUU limits the expression of the infC (initiation factor IF3) gene of Escherichia coli.* Mol Gen Genet, 1987. **208**(1-2): p. 94-100.

145. Qin, D., N.M. Abdi, and K. Fredrick, *Characterization of 16S rRNA mutations that decrease the fidelity of translation initiation.* RNA, 2007. **13**(12): p. 2348-55.

146. Qin, D. and K. Fredrick, *Control of translation initiation involves a factor-induced rearrangement of helix 44 of 16S ribosomal RNA.* Mol Microbiol, 2009. **71**(5): p. 1239-49.

147. Butler, J.S., M. Springer, and M. Grunbergmanago, *Auu-to-Aug Mutation in the Initiator Codon of the Translation Initiation-Factor If3 Abolishes Translational Autocontrol of Its Own Gene (Infc) Invivo.* Proceedings of the National Academy of Sciences of the United States of America, 1987. **84**(12): p. 4022-4025.

148. Falconi, M., M. Brombach, C.O. Gualerzi, and C.L. Pon, *In vivo transcriptional pattern in the infC operon of Bacillus stearothermophilus.* Mol Gen Genet, 1991. **227**(1): p. 60-4.

149. Yusupova, G., L. Jenner, B. Rees, D. Moras, and M. Yusupov, *Structural basis for messenger RNA movement on the ribosome.* Nature, 2006. **444**(7117): p. 391-4.

150. Grigoriadou, C., S. Marzi, D. Pan, C.O. Gualerzi, and B.S. Cooperman, *The translational fidelity function of IF3 during transition from the 30 S initiation complex to the 70 S initiation complex.* J Mol Biol, 2007. **373**(3): p. 551-61.

151. Grigoriadou, C., S. Marzi, S. Kirillov, C.O. Gualerzi, and B.S. Cooperman, *A quantitative kinetic scheme for 70 S translation initiation complex formation.* J Mol Biol, 2007. **373**(3): p. 562-72.

152. Pon, C.L. and C.O. Gualerzi, *Mechanism of translational initiation in prokaryotes. IF3 is released from ribosomes during and not before 70 S initiation complex formation.* FEBS Lett, 1986. **195**(1-2): p. 215-9.

153. Lomakin, I.B., V.G. Kolupaeva, A. Marintchev, G. Wagner, and T.V. Pestova, *Position of eukaryotic initiation factor eIF1 on the 40S ribosomal subunit determined by directed hydroxyl radical probing.* Genes Dev, 2003. **17**(22): p. 2786-97.

154. Rabl, J., M. Leibundgut, S.F. Ataide, A. Haag, and N. Ban, *Crystal structure of the eukaryotic 40S ribosomal subunit in complex with initiation factor 1.* Science, 2011. **331**(6018): p. 730-6.

155. Passmore, L.A., T.M. Schmeing, D. Maag, D.J. Applefield, M.G. Acker, M.A. Algire, J.R. Lorsch, and V. Ramakrishnan, *The eukaryotic translation initiation factors eIF1 and eIF1A induce an open conformation of the 40S ribosome.* Mol Cell, 2007. **26**(1): p. 41-50.

156. Julian, P., P. Milon, X. Agirrezabala, G. Lasso, D. Gil, M.V. Rodnina, and M. Valle, *The Cryo-EM Structure of a Complete 30S Translation Initiation Complex from Escherichia coli.* PLoS Biol, 2011. **9**(7): p. e1001095.

157. Pon, C., S. Cannistraro, A. Giovane, and C. Gualerzi, *Structure-function relationship in Escherichia coli initiation factors. Environment of the Cys residue and evidence for a hydrophobic region in initiation factor IF3 by fluorescence and ESR spectroscopy.* Arch Biochem Biophys, 1982. **217**(1): p. 47-57.

158. Vanduin, J., C.G. Kurland, J. Dondon, and M. Grunbergmanago, *Near Neighbors of If3-Bound to 30s-Ribosomal Subunits.* Febs Letters, 1975. **59**(2): p. 287-290.

159. Gualerzi, C. and C.L. Pon, *Nature of the ribosomal binding site for initiation factor 3 (IF-3).* Biochem Biophys Res Commun, 1973. **52**(3): p. 792-9.

160. Pon, C.L., R.T. Pawlik, and C. Gualerzi, *The topographical localization of IF3 on Escherichia coli 30 S ribosomal subunits as a clue to its way of functioning.* Febs Letters, 1982. **137**(2): p. 163-7.

161. Forster, T., *Zwischenmolekulare Energiewanderung Und Fluoreszenz.* Annalen Der Physik, 1948. **2**(1-2): p. 55-75.

162. Lakowicz, J.R., *Principles of Fluorescence Spectroscopy*, 2006, Springer: New York.

163. Brasselet, S., E.J.G. Peterman, A. Miyawaki, and W.E. Moerner, *Single-molecule fluorescence resonant energy transfer in calcium concentration dependent cameleon.* Journal of Physical Chemistry B, 2000. **104**(15): p. 3676-3682.

164. Ha, T.J., A.Y. Ting, J. Liang, W.B. Caldwell, A.A. Deniz, D.S. Chemla, P.G. Schultz, and S. Weiss, *Single-molecule fluorescence spectroscopy of enzyme conformational dynamics and cleavage mechanism.* Proceedings of the National Academy of Sciences of the United States of America, 1999. **96**(3): p. 893-898.

165. Yasuda, R. and K. Kinosita, *Rapid rotation of F1-ATPase detected with micro-beads attached to the rotor.* Molecular Biology of the Cell, 1999. **10**: p. 10A-10A.

166. Chen, Y., D.H. Hu, E.R. Vorpagel, and H.P. Lu, *Probing single-molecule T4 lysozyme conformational dynamics by intramolecular fluorescence energy transfer.* Journal of Physical Chemistry B, 2003. **107**(31): p. 7947-7956.

167. Borsch, M., M. Diez, B. Zimmermann, R. Reuter, and P. Graber, *Stepwise rotation of the gamma-subunit of EF0F1-ATP synthase observed by intramolecular single-molecule fluorescence resonance energy transfer.* Febs Letters, 2002. **527**(1-3): p. 147-152.

168. Schuette, C.G., K. Hatsuzawa, M. Margittai, A. Stein, D. Riedel, P. Kuster, M. Konig, C. Seidel, and R. Jahn, *Determinants of liposome fusion mediated by synaptic SNARE proteins.* Proceedings of the National Academy of Sciences of the United States of America, 2004. **101**(9): p. 2858-2863.

169. Diez, M., B. Zimmermann, M. Borsch, M. Konig, E. Schweinberger, S. Steigmiller, R. Reuter, S. Felekyan, V. Kudryavtsev, C.A.M. Seidel, and P. Graber, *Proton-powered subunit rotation in single membrane-bound F0F1-ATP synthase.* Nature Structural & Molecular Biology, 2004. **11**(2): p. 135-141.

170. Sako, Y., S. Minoghchi, and T. Yanagida, *Single-molecule imaging of EGFR signalling on the surface of living cells.* Nat Cell Biol, 2000. **2**(3): p. 168-72.

171. Nguyen, V.T., Y. Kamio, and H. Higuchi, *Single-molecule imaging of cooperative assembly of gamma-hemolysin on erythrocyte membranes.* EMBO J, 2003. **22**(19): p. 4968-79.

172. Deniz, A.A., T.A. Laurence, G.S. Beligere, M. Dahan, A.B. Martin, D.S. Chemla, P.E. Dawson, P.G. Schultz, and S. Weiss, *Single-molecule protein folding: diffusion fluorescence resonance energy transfer studies of the denaturation of chymotrypsin inhibitor 2.* Proc Natl Acad Sci U S A, 2000. **97**(10): p. 5179-84.

173. Jia, Y.W., D.S. Talaga, W.L. Lau, H.S.M. Lu, W.F. DeGrado, and R.M. Hochstrasser, *Folding dynamics of single GCN4 peptides by fluorescence resonant energy transfer confocal microscopy.* Chemical Physics, 1999. **247**(1): p. 69-83.

174. Talaga, D.S., W.L. Lau, H. Roder, J.Y. Tang, Y.W. Jia, W.F. DeGrado, and R.M. Hochstrasser, *Dynamics and folding of single two-stranded coiled-coil peptides studied by fluorescent energy transfer confocal microscopy.* Proceedings of the National Academy of Sciences of the United States of America, 2000. **97**(24): p. 13021-13026.

175. Schuler, B., E.A. Lipman, and W.A. Eaton, *Probing the free-energy surface for protein folding with single-molecule fluorescence spectroscopy.* Nature, 2002. **419**(6908): p. 743-747.

176. Rhoades, E., E. Gussakovsky, and G. Haran, *Watching proteins fold one molecule at a time.* Proceedings of the National Academy of Sciences of the United States of America, 2003. **100**(6): p. 3197-3202.

177. Ha, T., X.W. Zhuang, H.D. Kim, J.W. Orr, J.R. Williamson, and S. Chu, *Ligand-induced conformational changes observed in single RNA molecules.* Proceedings of the National Academy of Sciences of the United States of America, 1999. **96**(16): p. 9077-9082.

178. Kim, H.D., G.U. Nienhaus, T. Ha, J.W. Orr, J.R. Williamson, and S. Chu, *Mg2+-dependent conformational change of RNA studied by fluorescence correlation and FRET on immobilized single molecules.* Proceedings of the National Academy of Sciences of the United States of America, 2002. **99**(7): p. 4284-4289.

179. Bartley, L.E., X.W. Zhuang, R. Das, S. Chu, and D. Herschlag, *Exploration of the transition state for tertiary structure formation between an RNA helix and a large structured RNA.* Journal of Molecular Biology, 2003. **328**(5): p. 1011-1026.

180. Bokinsky, G., D. Rueda, V.K. Misra, M.M. Rhodes, A. Gordus, H.P. Babcock, N.G. Walter, and X.W. Zhuang, *Single-molecule transition-state analysis of RNA folding.* Proceedings of the National Academy of Sciences of the United States of America, 2003. **100**(16): p. 9302-9307.

181. Russell, R., X.W. Zhuang, H.P. Babcock, I.S. Millett, S. Doniach, S. Chu, and D. Herschlag, *Exploring the folding landscape of a structured RNA.* Proceedings of the National Academy of Sciences of the United States of America, 2002. **99**(1): p. 155-160.

182. Zhuang, X., L.E. Bartley, H.P. Babcock, R. Russell, T. Ha, D. Herschlag, and S. Chu, *A single-molecule study of RNA catalysis and folding.* Science, 2000. **288**(5473): p. 2048-51.

183. Zhuang, X.W., H. Kim, M.J.B. Pereira, H.P. Babcock, N.G. Walter, and S. Chu, *Correlating structural dynamics and function in single ribozyme molecules.* Science, 2002. **296**(5572): p. 1473-1476.

184. Nahas, M.K., T.J. Wilson, S.C. Hohng, K. Jarvie, D.M.J. Lilley, and T. Ha, *Observation of internal cleavage and ligation reactions of a ribozyme.* Nature Structural & Molecular Biology, 2004. **11**(11): p. 1107-1113.

185. Tan, E., T.J. Wilson, M.K. Nahas, R.M. Clegg, D.M.J. Lilley, and T. Ha, *A four-way junction accelerates hairpin ribozyme folding via a discrete intermediate.* Proceedings of the National Academy of Sciences of the United States of America, 2003. **100**(16): p. 9308-9313.

186. Hohng, S., T.J. Wilson, E. Tan, R.M. Clegg, D.M.J. Lilley, and T.J. Ha, *Conformational flexibility of four-way junctions in RNA.* Journal of Molecular Biology, 2004. **336**(1): p. 69-79.

187. Solomatin, S.V., M. Greenfeld, S. Chu, and D. Herschlag, *Multiple native states reveal persistent ruggedness of an RNA folding landscape.* Nature, 2010. **463**(7281): p. 681-4.

188. MacDougall, D.D. and R.L. Gonzalez, Jr., *Exploring the structural dynamics of the translational machinery using single-molecule fluorescence resonance energy transfer*, in *Ribosomes: Structure, Function, and Dynamics*, M.V. Rodnina, R. Green, and W. Wintermeyer, Editors. 2011, Springer-Verlag: New York.

189. Ha, T., *Single-molecule fluorescence resonance energy transfer.* Methods, 2001. **25**(1): p. 78-86.

190. Cornish, P.V. and T. Ha, *A survey of single-molecule techniques in chemical biology.* ACS Chem Biol, 2007. **2**(1): p. 53-61.

191. Weiss, S., *Fluorescence spectroscopy of single biomolecules.* Science, 1999. **283**(5408): p. 1676-83.

192. Weiss, S., *Measuring conformational dynamics of biomolecules by single molecule fluorescence spectroscopy.* Nat Struct Biol, 2000. **7**(9): p. 724-9.

193. Ishijima, A. and T. Yanagida, *Single molecule nanobioscience.* Trends Biochem Sci, 2001. **26**(7): p. 438-44.

194. Roy, R., S. Hohng, and T. Ha, *A practical guide to single-molecule FRET.* Nat Methods, 2008. **5**(6): p. 507-16.

195. Rasnik, I., S.A. McKinney, and T. Ha, *Surfaces and orientations: much to FRET about?* Acc Chem Res, 2005. **38**(7): p. 542-8.

196. Tinoco, I., Jr. and R.L. Gonzalez, Jr., *Biological mechanisms, one molecule at a time.* Genes Dev, 2011. **25**(12): p. 1205-31.

197. Fei, J., J.E. Bronson, J.M. Hofman, R.L. Srinivas, C.H. Wiggins, and R.L. Gonzalez, Jr., *Allosteric collaboration between elongation factor G and the ribosomal L1 stalk directs tRNA movements during translation.* Proc Natl Acad Sci U S A, 2009. **106**(37): p. 15702-7.

198. Eddy, S.R., *What is a hidden Markov model?* Nat Biotechnol, 2004. **22**(10): p. 1315-6.

199. Bronson, J.E., J. Fei, J.M. Hofman, R.L. Gonzalez, Jr., and C.H. Wiggins, *Learning rates and states from biophysical time series: a Bayesian approach to model selection and single-molecule FRET data.* Biophys J, 2009. **97**(12): p. 3196-205.

200. Qin, F., A. Auerbach, and F. Sachs, *Estimating single-channel kinetic parameters from idealized patch-clamp data containing missed events.* Biophys J, 1996. **70**(1): p. 264-80.

201. McKinney, S.A., C. Joo, and T. Ha, *Analysis of single-molecule FRET trajectories using hidden Markov modeling.* Biophys J, 2006. **91**(5): p. 1941-51.

202. Frank, J. and R.L. Gonzalez, Jr., *Structure and dynamics of a processive Brownian motor: the translating ribosome.* Annu Rev Biochem, 2010. **79**: p. 381-412.

203. Wang, J., *Regulation of IF2 Binding Kinetics and 30S IC Conformational Dynamics during Translation Initiation*, 2010, Columbia University.

204. Marshall, R.A., C.E. Aitken, and J.D. Puglisi, *GTP hydrolysis by IF2 guides progression of the ribosome into elongation.* Mol Cell, 2009. **35**(1): p. 37-47.

Chapter 2

An interdomain reconfiguration of IF3 signals proper initiator tRNA and mRNA start codon selection during translation initiation*

2.1 Introduction

Translation initiation, the rate-limiting step of protein synthesis and a significant regulatory checkpoint in gene expression [1, 2], requires a multi-step, initiation factor-regulated assembly of a 30S IC followed by a 70S IC containing an accurately selected initiator tRNA and properly positioned start codon within its P site (see Figure 1.2 and section 1.2). This assembly process sets the mRNA reading frame, thus incorrect start codon selection can trigger a frameshifting error resulting in production of truncated and consequently misfolded protein [1, 2], and selection of a non-initiator tRNA results in an incorrect N-terminal protein sequence. In prokaryotes, IF1, IF2, and IF3 play essential roles in ensuring proper substrate selection [3]. Of these, IF3 is particularly critical for maintaining the fidelity of translation initiation, having known functions in discriminating the start codon [4-7], tRNA [8-11], and codon-anticodon interaction [8, 9, 12]. In the absence of IF1 and IF2, 30S subunit-bound IF3 can, either directly or through an IF3-induced conformational change of the 30S subunit, select against all codon-anticodon interactions other than AUG-fMet, GUG-fMet, and UUG-fMet [8, 9], while kinetic studies have shown that IF3 indiscriminately destabilizes all P-site tRNAs [13, 14]. Beyond its ability to assist in proper tRNA selection, IF3 further regulates fidelity during 70S IC assembly by preventing efficient 50S subunit association with incompletely or incorrectly assembled 30S ICs [13, 15-17]. Despite its numerous known roles in regulating the fidelity of initiation complex assembly, mechanistic explanations for how IF3 performs these roles have remained elusive.

---

* *Elvekrog, MM and Gonzalez Jr, RL (2011) An interdomain reconfiguration of IF3 signals proper initiator tRNA and mRNA start codon selection during translation initiation (Submitted)*

IF3 is composed of two globular and similarly sized, N- and C-terminal domains (NTD and CTD, respectively) that are separated by a highly conserved and flexible linker which is essential for IF3's *in vivo* functions [18-21]. When free in solution, the flexible linker allows IF3 to adopt a range of conformations in which the interdomain distance varies from 28 to 64 Å [20] (see Figure 1.9). IF3 also exhibits dynamic behavior with stepwise association with, and dissociation from, the 30S subunit due to the two domains' independent behavior and different affinities for the 30S subunit, with the CTD having a much higher affinity for the 30S subunit relative to the NTD (see Figure 1.13) [22-24]. Biochemical and structural studies of IF3's interactions with an empty 30S subunit demonstrate that IF3's binding site includes the G700 region (helix 23) of the 16S rRNA on the upper part of the platform of the 30S subunit and the G790 region (helix 24) and 3' region (helix 45) of the 16S rRNA at the interface between the platform and body domains of the 30S subunit, near the P and E sites [22, 25-29] (see discussion in section 1.3.3.5 and Figures 1.11 and 1.12). Although most studies agree that the CTD of IF3, as well as the orthologous a/eIF1 in archaea and eukaryotes, binds at the small subunit platform region near the P site [22, 27, 28, 30-34], discrepancies exist regarding the placement of the NTD [22, 27, 28] as well as the possible existence of alternative IF3 binding sites on the 30S subunit [35, 36]. IF3's interactions with a fully assembled 30S IC, when IF1, IF2, mRNA, and tRNA are also bound to the 30S subunit, may differ signficantly from its interactions with an empty 30S subunit. Regardless, multiple studies implicate 16S rRNA residues and ribosomal proteins on the 30S subunit for IF3 binding that are also involved in the formation of some subset of intersubunit bridges [37, 38] (see Table 1.1 and Figure 1.12). This shared binding site implies a need for conformational rearrangement and/or dissociation of IF3 in order for intersubunit bridge formation and stable 70S IC assembly. Hence, although the conformational dynamics of IF3 may present an obstacle to structural studies of free and ribosome-bound IF3, they are critical to understanding its functional roles in regulating the fidelity of translation initiation.

The aim of this work was to test the hypothesis that 30S subunit-bound IF3 is conformationally dynamic during the 30S IC assembly pathway and that its dynamics are

essential to IF3's roles in substrate selection. Employing an IF3 variant labeled with a Cy3 FRET donor fluorophore and a Cy5 FRET acceptor fluorophore at its NTD and CTD, I developed an intramolecular NTD-CTD IF3 single-molecule fluoresence resonance energy transfer (smFRET) signal that reports on the interdomain dynamics of 30S IC-bound IF3. Using smFRET, I show that the 30S IC-bound IF3 exists in a dynamic equilibrium between at least three stable interdomain configurations. Moreover, I show that recognition of a cognate start codon-initiator tRNA interaction during translation initiation uniquely shifts this interdomain configuration equilibrium toward a single state. These findings reveal a novel, structural rationale for IF3's role in regulating the fidelity of substrate selection in translation initiation.

The materials and methods used to obtain the results described herein are presented in section 2.2. Sections 2.3.1 to 2.3.6 describe the design, preparation, and biochemical activity testing of a dual fluorescently labeled IF3 for smFRET studies of its 30S subunit-bound interdomain configurations. The results of these smFRET investigations are found in the remainder of section 2.3, from 2.3.7 to 2.3.13, and the conclusions and possible future directions are found in section 2.4.

2.2 Experimental Methods

2.2.1 Buffers

A low-salt version of the standard Tris-polymix buffer developed and optimized for *in vitro* translation work was used for all smFRET and biochemistry work [39-41]. The buffer conditions, as optimized for translation initiation work by Dr. Jiangning Wang and described in her Ph.D. dissertation [42], are: 10 mM Tris-acetate (Tris-HOAc) ($pH_{25°C}$ = 7.5), 20 mM KCl, 5-15 mM $Mg(OAc)_2$ (with the exact concentration depending on the experiment), 1 mM NH_4OAc, 0.1 mM $Ca(OAc)_2$, 0.1 mM EDTA, 6 mM β-mercaptoethanol (BME), 5 mM putrescine-HCl, and 1 mM spermidine-free base. See Chapter 5 for a complete list of all buffers used and their compositions. All reagents were of molecular biology grade or higher. Those buffers used for protein purification are described in detail in Ref. [43] and Chapter 5.

2.2.2 30S ribosomal subunits

Highly active, tightly-coupled 70S ribosomes were purified by Dr. Jiangning Wang from *E. coli* strain MRE600 using sucrose density gradient ultracentrifugation (SDGC) as previously described [43]. These intact ribosomes were then dissociated into their component 30S and 50S subunits by dialyzing against buffer containing 1 mM Mg^{2+}, followed by purification of the 30S and 50S subunits using SDGC in 1 mM Mg^{2+} buffer. The 30S subunits were pelleted, resuspended, and stored in small aliquots at -80°C in a buffer containing 7.5 mM Mg^{2+} until further use.

2.2.3 Messenger RNAs

5'-biotinylated mRNAs with sequences derived from the mRNA encoding gene product 32 from T4 bacteriophage (T4gp32) were chemically synthesized by Dharmacon, Inc. (Thermo Scientific). The mRNAs were purified by polyacrylamide gel electrophoresis (PAGE) and 2'-deprotected/desalted by Dharmacon. The sequence of the short (64 nucleotide) T4gp32 mRNA variant (#1) is found in Table 2.1. Two other mRNA variants were used for work in this chapter in which the start codon was modified to AUU (mRNA #2), and in which the first two codons were swapped, placing UUC in the start codon position and AUG in position two (mRNA #3). The sequence of the mRNA used for the primer extension inhibition ("toeprinting") biochemical activity assay discussed in section 2.3.5 is also listed in Table 2.1 (mRNA #4). This mRNA was generated by *in vitro* transcription.

Table 2.1 mRNA sequences. The nucleotides in bold are the Shine-Dalgarno sequence with the core AGGA sequence, the spacer nucleotides are italicized, and the P-site (start) codon is underlined. Those mRNAs modified to contain a 5' biotin moiety are indicated with "Bi."

#	mRNA nucleotide sequence
1	5'-Bi-CAACCUAAAACUUACACAAAUUAAA**AAGGA***AAUAGAC*AUGUUCAAAGUCGAAAAAUCUACUGCU-3'
2	5'-Bi-CAACCUAAAACUUACACAAAUUAAA**AAGGA***AAUAGAC*AUUUUCAAAGUCGAAAAAUCUACUGCU-3'
3	5'-Bi-CAACCUAAAACUUACACAAAUUAAA**AAGGA***AAUAGAC*UUCAUGAAAGUCGAAAAAUCUACUGCU-3'
4	5'GGCAACCUAAAACUUACACAGGGCCC**UAAGGA***AAUAAAA*AUGUUUAAAGAAGUAUACACUGCUGAACUCGCU GCACAAAUGGCUAAACUGAAUGGCAAUAAAGGUUUUUCUUCUGAAGAUAAAGGCGAGUGGAAACUGAAACUCG AUAAUGCGGGUAACGGUCAAGCAGUAAUUCGUUUUCUUCCGUCUAAAAAUGAUGAACAAGCACCAUUCGCAAU UCUUGUAAAUCACGGUUUCAAGAAAAAUGGUAAAUGGUAUAUUGAAACAUGUUCAUCUACCCAUGGUGAUUAC GAUUCUUGCCCAGUAUGUCAAUACAUCAGUAAAAAUGAUCUAUACAACACUGACAAUAAAGAGUACAGUCUUG UUAAACGUAAAACUUCUUACUGGGCUAACAUUCUUGUAGUAAAAGACCCAGCUGCUCCAGAAAACGAAGGUAA

AGUAUUUAAAUACCGUUUCGGUAAGAAAAUCUGGGAUAAAAUCAAUGCAAUGAUUGCGGUUGAUGUUGAAAUG
GGUGAAACUCCAGUUGAUGUAACUUGUCCGUGGGAAGGUGCUAACUUUGUACUGAAAGUUAAACAAGUUUCU
GGAUUUAGUAACUACGAUGAAUCUAAAUUCCUGAAUCAAUCUGCGAUUCCAAACAUUGACGAUGAAUCUUUCC
AGAAAGAACUGUUCGAACAAAUGGUCGACCUUUCUGAAAUGACUUCUAAAGAUAAAUAAGG-3'

2.2.4 Transfer RNAs

N-formylmethionine-specific tRNA from *E. coli* was purchased from MP Biomedicals, and phenylalanine- and lysine-specific tRNAs from *E. coli* were purchased from Sigma-Aldrich. tRNAs were aminoacylated and, in the case of $tRNA^{fMet}$, formylated as described previously [43] and detailed in section 5.1.1. Aminoacylation and formylation yields were determined by hydrophobic interaction chromatography (HIC) on a TSKgel Phenyl-5PW column (8.0 mm (ID) x 7.5 cm (L)) (Tosoh Bioscience) and were consistently $\geq$90% for $tRNA^{fMet}$ and $tRNA^{Phe}$, and $\geq$60% for $tRNA^{Lys}$ [43]. fMet-(Cy3)$tRNA^{fMet}$, labeled at the 4-thiouridine at nucleotide position 8 (s^4U8), was prepared by Dr. Jiangning Wang as previously described [42, 43].

2.2.5 Translation factors

The genes encoding IF1, IF2γ, and IF3 from *E. coli* were cloned, and the proteins overexpressed and purified as previously described [43] and detailed in section 5.2. Briefly, all genes were cloned into a pProEX-HTb plasmid (Invitrogen) under control of an isopropyl β-D-1-thiogalactopyranoside (IPTG)-inducible pTrc promoter. This plasmid introduces a six-histidine (6xHis) affinity tag followed by a tobacco etch virus (TEV) protease cleavage site at the amino terminus of each translation factor. The cloning strategy and plasmid employed necessitated that the *N*-terminal ends of the IFs differ slightly from their wild-type *E. coli* gene sequences. Specifically, the *N*-terminal ends are: G-A-M1 (IF1), G-A-Q-D-D-M1 (IF2γ), and G-A-M-A-K2 (IF3), where the underlined amino acid and sequence position denote the beginning of the wild-type gene sequence.

The three initiation factors were purified using standard Ni^{2+}-nitrilotriacetic acid (NTA) affinity purification procedures (resin from Qiagen) [43]. The 6xHis affinity tag was removed from each initiation factor by cleavage with TEV protease followed by a second Ni^{2+}-NTA column. Additional purification steps vary by initiation factor. IF1 is further purified by size exclusion

column chromatography (HiLoad 16/60 Superdex 75 prep grade–GE Biosciences), while IF2 and IF3 require additional purification with a cation exchange column (HiTrap SP HP–GE Biosciences; see Figure 2.2 for a representative IF3 cation exchange column chromatogram). The detailed protocol for purification of IF1, IF2, and IF3 is found in Ref [43] and section 5.2. The design, mutagenesis, purification, and Cy3-/Cy5-fluorophore labeling of the IF3 variant used in this study is described in the Results section (2.3.1).

2.2.6 Preparation of 30S Initiation Complexes (30S ICs) for smFRET investigations

The nomenclature used throughout this dissertation to describe the contents of the 30S IC under investigation is as follows: $30SIC^{(-)tRNA,XYZ}_{-A/B}$, where the superscript indicates the identity of the aa-tRNA present in the complex (e.g., fMet) or lack of a tRNA (–tRNA) in the complex, and XYZ refers to the identity of the start codon if it differs from the canonical AUG. The subscript –A/B indicates which non-IF3 initiation factors are absent from the complex. For example, $30SIC^{fMet,AUU}$ is an initiation complex assembled with a 30S subunit, fMet-$tRNA^{fMet}$, an mRNA with an AUU start codon, and all three initiation factors. As an additional example, $30SIC^{-tRNA}_{-2}$ refers to an initiation complex formed from a 30S subunit, mRNA with an AUG start codon, IF1, and IF3, but lacking aa-tRNA and IF2.

30S ICs were always assembled in low-salt Tris-polymix buffer containing 5 mM Mg^{2+} by mixing 1.8 μM 5'-biotinylated mRNA, 0.9 μM IF1, 0.9 μM IF2, 1 mM GTP, 0.9 μM aa-tRNA, and 0.6 μM fluorescently-labeled IF3. Lastly, 0.6 μM 30S subunits were added to this mixture to initiate the assembly reaction. The 30S ICs were incubated at 37°C for 10 minutes, and subsequently aliquoted, flash frozen with liquid nitrogen, and stored at -80°C until further use.

2.2.7 Surface-Immobilization of 30S ICs within microfluidic flowcells

Quartz microscope slides (G. Finkenbeiner) were extensively cleaned and then chemically passivated with PEG-Succinimidyl Carbonate (SC) (m.w. 5,000; Laysan Bio, Inc.) containing 0.1% biotin-modified PEG-SC according to published protocols [44]. Flowcells were assembled on these passivated slides as depicted in Figure 1.17. Immediately prior to 30S IC

sample introduction, a solution of 10 μM "Block" containing 10 μM UltraPure BSA (Invitrogen) and 10 μM hybridized DNA duplex was loaded into the flowcell to limit non-specific binding of the fluorescently labeled molecules to the flowcell surface, followed by a mixture of 1 μM streptavidin (Invitrogen) and 10 μM "Block." 30S ICs, diluted to 100-300 pM in low-salt Tris-polymix buffer, were then added to the passivated flowcell and immobilized via the 5'-biotin-mRNA on the surface through a PEG-biotin-streptavidin-biotin interaction. Saturating concentrations of IF1 (0.9 μM), IF2 (0.9 μM), aa-tRNA (0.9 μM), and GTP (1 mM) were kept in all dilution and wash buffers (when present in the 30S IC). Single-molecule fluorescence imaging was then performed with a lab-built TIRF microscope system (see sections 1.4.2 and 2.2.8). The concentration of 30S IC was optimized in order to achieve a surface concentration of ~200-400 surface-immobilized 30S ICs per field of view (FOV) (~60 x 120 μm^2), but was always in the range of 100-300 pM. After delivery to the flowcell, 30S ICs were incubated for 5 min at room temperature (r.t.). Unbound 30S ICs were then flushed out with low-salt Tris-polymix buffer supplemented with an oxygen scavenging system (optimized for these experiments as: 165 U/mL glucose oxidase, 2170 U/mL catalase, 1% β-D-glucose) and triplet state quenchers (1 mM 1,3,5,7-cyclooctatetraene, 1 mM *p*-Nitrobenzyl alcohol) to minimize singlet oxygen-induced fluorophore damage and transitions into dark triplet states, respectively (see section 5.6). In total, 8-10 min passed between sample delivery to the flowcell and data acquisition since time was needed to mount the slide on the microscope stage and focus the illuminated FOV.

2.2.8 TIRFM imaging of surface-immobilized 30S ICs

A lab-built, wide-field prism-based TIRFM was designed and constructed as previously described [44] (see Figure 1.17). A diode-pumped 532 nm laser (CrystaLaser) was used as an excitation source and a back-thinned, 512 x 512 pixel electron-multiplying (EM)CCD camera (Cascade II: Princeton Instruments) was used as a detector. A DualView emission splitter (Photometrics) was used to split the capacitor array within the EMCCD detector into separate donor and acceptor channels, creating Cy3 and Cy5 observation areas of 60 x 120 μm^2 each that

enabled visualization of 200-400 spatially well-separated initiation complexes per movie (2 x 2 pixel binning). All data were collected under 7 mW 532 nm laser power excitation (as measured at the prism) and all movies were collected at 10 frames per second time resolution.

2.2.9 smFRET data analysis

Single Cy3 and Cy5 fluorophores were selected and their corresponding Cy3 and Cy5 fluorescence emission versus time trajectories were generated as described in section 1.4.3 and in Refs. [45-47]. Briefly, each fluorescence emission trajectory was visually inspected and those trajectories that exhibited characteristic single-fluorophore fluorescence intensities and single-step fluorophore photobleaching were kept for further analysis. Those trajectories that showed Cy3-Cy5 anti-correlated behavior were retained (see Figure 2.7 for representative trajectories) and, of these, trajectories exhibiting photobleaching of Cy3 or Cy5 within the first second (10 frames) were discarded. Following this initial selection process, the Cy3 and Cy5 signals were baseline-corrected by subtracting the average intensity of the last ten data points after photobleaching. Additionally, the Cy5 signal was adjusted to account for Cy3 signal bleed-through due to imperfect emission filters which allow a small amount of Cy3 emission to bleed through into the Cy5 emission channel. This value was experimentally determined to be 7% for our system [46]. FRET efficiency versus time trajectories were then plotted for each trajectory using the equation $FRET=I_{Cy5}/(I_{Cy3}+I_{Cy5})$, where I_{Cy3} and I_{Cy5} are the intensities of the baseline corrected Cy3 and Cy5 data points, respectively.

For each dataset, the data from just the first 0.5 seconds were used to generate a normalized population FRET histogram in order to minimize the zero FRET peak which arises from fluorophore photobleaching (see Figure 2.9 for representative example). Origin 8.0 (OriginLab) was used to fit each histogram with Gaussian distributions and the lower and upper limits for each FRET efficiency value were set by using the full width at half-maximum (FWHM) of the Gaussian distributions. Subpopulation analysis was done manually following smFRET trajectory idealization using the vbFRET software package to idealize each smFRET trajectory as

a hidden Markov model [48]. The smFRET data is also plotted as a time-evolution of population FRET histogram (see Figure 2.9 for representative example) for the first five seconds.

2.3 Results and Discussion

2.3.1 Design and preparation of an IF3 variant for dual labeling with Cy3 and Cy5

In order to probe relative distance changes between IF3's two domains, a two-Cys-containing IF3 variant was designed which could be dual fluorescently-labeled at Cys residues with maleimide-conjugated versions of the Cy3/Cy5 FRET pair (see Figure 2.1). Single residues in IF3's NTD and CTD were chosen for QuickChange site-directed mutagenesis (Stratagene) that are solvent-exposed and have low sequence conservation based on BLAST multiple sequence alignment [49]. Based on modeling of IF3 linker conformers using NMR constraints, the centers-of-mass of the NTD and CTD are able to adopt a range of separation distances from 28 to 64 Å in solution (Figure 1.9) [20]. Fortuitously, the extremes of these distances lie on both sides of the Förster distance, R_0, for the Cy3/Cy5 FRET pair (R_0 = 55 Å) [50, 51], enabling sensitive measurements of relative interdomain distance changes. Initially, IF3's wild-type single Cys65 was the target NTD residue, and two sites within the CTD (K97C and M143C) were chosen for a single step mutagenesis route. Cys65 is a highly conserved (>90%) [5] and solvent inaccessible (<30% as determined by MOLMOL [52]) residue, however, and initial attempts at labeling wild-type IF3 under non-denaturing conditions gave labeling efficiencies of only 15-20%. Additionally, although other research groups have claimed that modifying this Cys residue does not perturb IF3's ability to bind 30S subunits [28, 53, 54], I have found that the biochemical activity of IF3(Cy3 or Cy5) is significantly lower compared with unlabeled IF3 in multiple biochemical activity assays (see section 2.3.5). Thus, a modified strategy was taken to obtain a double Cys-containing IF3 variant. Cys65 was first mutated to Ala, following the route of Dallas and Noller [28]; however this mutation proved to be disruptive to IF3's biochemical activity. A C65S construct was then generated which demonstrated wild-type-like activity. Other NTD residues were then chosen for the introduction of a Cys residue. These point mutants included: R40C, E44C, E61C, Q23C, L29C, S57C, and S38C (see Table 2.2). A step-wise mutagenesis approach was taken to

introduce the C65S, CTD, and NTD mutations. The gene sequences of all mutants were verified by DNA sequencing (Genewiz, Plainfield, NJ) and the proteins were overexpressed, purified, fluorescently-labeled and screened for biochemical activity. Of all the constructs, however, the only double Cys-containing IF3 variant exhibiting high biochemical activity is IF3(C65S-S38C-K97C) (see Table 2.2 and section 2.3.5). Thus, all subsequent smFRET work was performed with a Cy3- and Cy5-labeled version of IF3(C65S-S38C-K97C) (see the following section for labeling details and section 5.2.2.1 for the exact sequence of this protein).

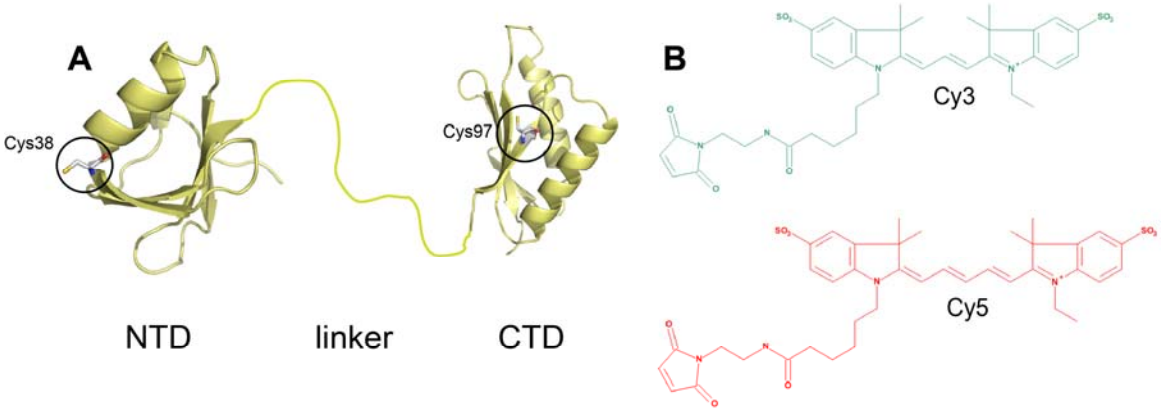

Figure 2.1 An intramolecular IF3 smFRET signal to monitor interdomain dynamics. A. Ribbon diagram of IF3 NTD and CTD from *B. stearothermophilus*. Residues 61 to 78 (*B. stearo* numbering) were removed and the flexible, interdomain linker was drawn in to accurately represent its flexibility in *E. coli* IF3 [20, 21, 55]. *E. coli* IF3 was mutagenized and then labeled with Cy3 and Cy5 at residues Cys38 and Cys97. The corresponding *B. stearothermophilus* residues are rendered as sticks. PDB codes 1TIF and 1TIG. **B.** The chemical structures of maleimide-conjugated Cy3 and Cy5 fluorophores.

Table 2.2 IF3 Cys mutants The plasmids for the following mutants were generated *via* site-directed mutagenesis, and their corresponding protein was overexpressed, purified, and tested for biochemical activity.

Mutant	Activity in fMet-(Cy3)tRNA^{fMet} dissociation assay*
IF3(K97C)	+
IF3(M142C)	–
IF3(C65A-K97C)	–
IF3(C65A-M142C)	–
IF3(C65S)	+
IF3(E61C)	–
IF3(E44C)	–
IF3(R40C)	–
IF3(C65S-M142C)	–
IF3(C65S-Q23C)	– (poor overexpression)
IF3(C65S-L29C)	– (poor overexpression)
IF3(C65S-S57C)	– (poor overexpression)
IF3(C65S-K97C)	+
IF3(C65S-K97C)Cy5	+
IF3(C65S-S38C)	+
IF3(C65S-S38C)Cy3	+
IF3(C65S-S38C-K97C)	+
IF3(C65S-S38C-K97C)Cy3-Cy5	+
IF3(C65S-S38C-K97C-Y75N)Cy3-Cy5	– (**)

*****See section 2.3.5 for assay details. "Inactive" (–) protein is defined as % fMet-(Cy3)tRNA^{fMet} falling outside the error limits of wild-type IF3 in the tRNA dissociation assay. "Active" (+) protein is within error of the wild-type IF3's activity in this assay.

******IF3-Y75N is a previously described loss of function point mutant [4, 5] (see section 2.3.4).

2.3.2 Fluorophore-labeling of IF3(C65S-S38C-K97C)

IF3(C65S-S38C-K97C) was buffer-exchanged into IF3 Labeling Buffer (100 mM Tris-OAc, pH_{rt} = 7.0; and 50 mM KCl) and incubated with a 10-fold molar excess of tris(2-carboxyethyl)phosphine (TCEP) hydrochloride for one hour at room temperature (r.t.). An equimolar mixture of Cy3- and Cy5-maleimide fluorophores (GE Healthcare) (20-fold molar excess over IF3 of each) was then added to the reaction mixture. The reaction was allowed to proceed in the dark for 1 hour at r.t. followed by overnight at 4°C. The reaction was quenched with a final concentration of 100 mM BME and the reaction volume was brought up to 1 mL with IF3 Labeling Buffer and transferred to a 10,000 MWCO slide-a-lyzer cassette (Thermo Scientific) and dialyzed at 4°C against 1 L HIC Buffer A (1 M $(NH_4)_2SO_4$ and 100 mM Na_2HPO_4 pH 7.0) to remove unreacted dye.

2.3.3 Purification and characterization of IF3(Cy3-Cy5)

The dual labeling strategy used here (i.e. labeling two Cys residues simultaneously with a mixture of two maleimide-conjugated dyes) results in numerous products, namely: unlabeled IF3, IF3 singly labeled at either its NTD or CTD with Cy3, IF3 singly labeled at either its NTD or CTD with Cy5, IF3 doubly labeled at its NTD and CTD with Cy3, IF3 doubly labeled at its NTD and CTD with Cy5, and the Cy3/Cy5 doubly labeled variants IF3-NTD(Cy3)-CTD(Cy5) and IF3-NTD(Cy5)-CTD(Cy3) (see Figure 2.3A). These products were separated using a TSKgel Phenyl-5PW HIC column (Tosoh Bioscience), which is a strategy that has been shown previously to enable comparable separations [56]. The HIC column was pre-equilibrated with HIC Buffer A, and a 0-100% HIC Buffer B (100 mM Na_2HPO_4 pH 7.0) gradient over 20 column volumes (CVs) (1 CV = 4 mL) enabled separation of the labeled IF3 species (see Figure 2.2). The two IF3(Cy3-Cy5) fractions were collected, concentrated, and stored in a total volume of 50% glycerol at -20°C. Imperfect separation of each labeled species resulted in some heterogeneity, however only those proteins labeled with both Cy3 and Cy5 (hereafter: IF3(Cy3-Cy5)) are able to undergo FRET. The

contaminating singly-labeled IF3 species will not undergo FRET, thus further "purification" occurs at the level of smFRET data acquisition and processing.

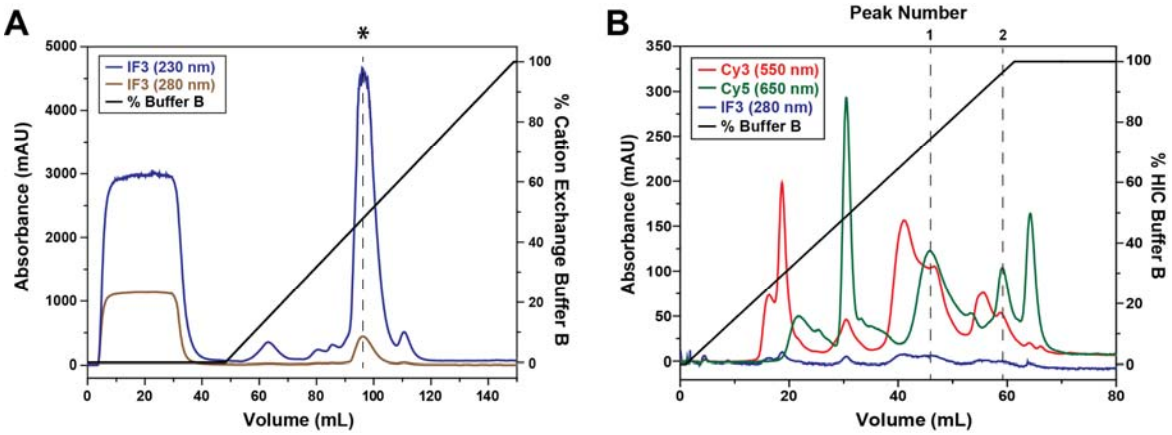

Figure 2.2 Purification and fluorescent labeling of IF3. IF3, containing an N-terminal His_6-tag, was overexpressed and purified by Ni^{2+}-NTA affinity chromatography. TEV protease was added to cleave the His_6-tag from the protein's N-terminus. A second Ni^{2+}-NTA column was performed to separate IF3 from the His_6-tag. The protein was purified to homogeneity by cation exchange column chromatography. **A.** Cation exchange column purification of IF3(C65S/S38C/K97C). The major peak (*) represents pure IF3. **B.** Purification of Cy3 and Cy5 labeled IF3 with hydrophobic interaction column (HIC) chromatography. Peaks 1 and 2 are both IF3(Cy3-Cy5), as confirmed by MALDI-TOF mass spectrometry (Appendix A) and smFRET (Figure 2.4).

The identity of the protein from peaks 1 and 2 (see Figure 2.2B) was verified through trypsin digestion of the interdomain linker followed by matrix-assisted laser desorption/ionization time-of-flight (MALDI-TOF) mass spectrometry analysis of the labeled domains (for details, see section 5.2.2.6; for mass spectra, see Appendix A). The identity of IF3 from both HIC peaks was similar: both contained a mixture of IF3-NTD(C65S-S38C)Cy3, IF3-NTD(C65S-S38C)Cy5, IF3-CTD(K97C)Cy3, and IF3-CTD(K97C)Cy5 (see Appendix A). Additionally, both of these IF3(Cy3-Cy5) samples were shown to behave biophysically similar, as shown with smFRET (see section 2.3.9) (Figure 2.4).

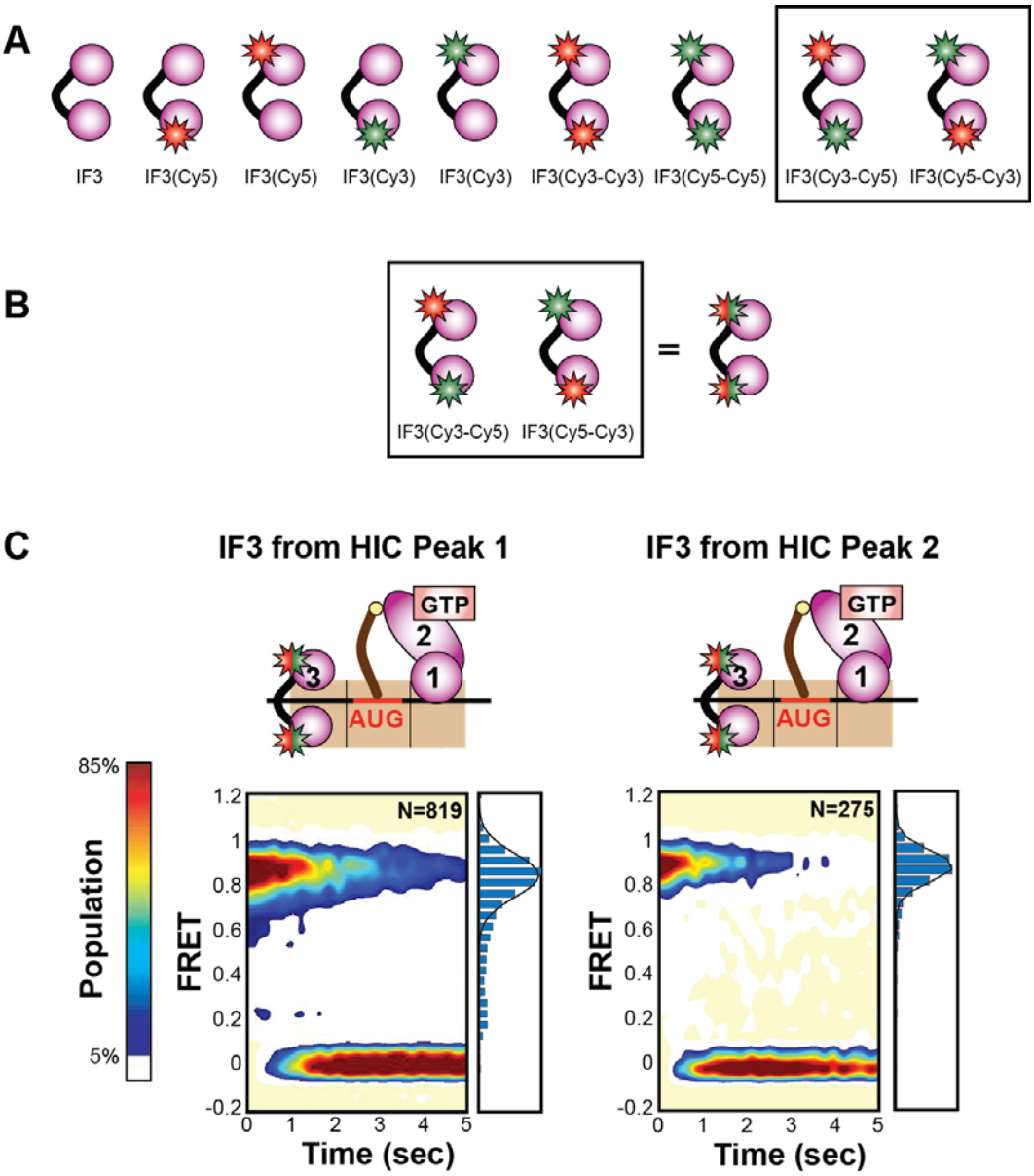

Figure 2.3 Both IF3(Cy3-Cy5) samples (peaks 1 and 2 from HIC purification) display similar smFRET behavior within 30SIC^{fMet}. A. Cartoon representation of the possible IF3 Cy3-Cy5 labeling products (unlabeled IF3, singly- and doubly-labeled IF3). The red star represents Cy5 and the green star represents Cy3. **B.** MALDI-TOF mass spectrometry confirmed that the identities of HIC peaks 1 and 2 (Figure 2.2) are both a mixture of IF3(CTD)-Cy3, IF3(CTD)-Cy5, IF3(NTD)-Cy5, and IF3(NTD)-Cy5 (see Appendix A). A cartoon representation of these heterogeneous samples is depicted on the right with a split red and green star. **C.** IF3 from both HIC peaks behaves biophysically similarly as evidenced by the use of each sample in an smFRET experiment. See section 2.3.9 for a description of the experiment and Figure 1.2 for details on the 30S IC cartoons. Time-evolution of population FRET histogram and normalized population FRET histogram of 30S IC-bound IF3(Cy3-Cy5) on complexes containing 5'-biotinylated-mRNA, IF1, IF2(GTP), and fMet-tRNA^{fMet}. **(left)** IF3(Cy3-Cy5) from peak 1 (see Figure 2.3B), **(right)** IF3(Cy3-Cy5) from peak 2. With this work and the MALDI-TOF mass spectrometry analysis as exceptions, all other experiments were performed with the IF3(Cy3-Cy5) sample from peak #1.

2.3.4 Preparation of a fluorescently-labeled loss-of-function point mutant: IF3(C65S-S38C-K97C-Y75N)-Cy3-Cy5

In addition to IF3(Cy3-Cy5), another IF3 variant was generated containing the substitution mutation Y75N. This point mutation, and its corresponding allele *infC362*, has been well described, and used previously, as a loss-of-function mutation that does not interfere with IF3's ability to bind to 30S subunits [4, 5, 57, 58]. Compared with wild-type IF3, IF3-Y75N is defective in its ability to (i) discriminate start codons, (ii) perform initiator tRNA selection, and (iii) inhibit initiation on leaderless mRNA. Considering the location of this residue on the edge of the NTD near the linker region (see Figure 2.13), I hypothesized that the loss-of-function phenotypes triggered solely by this single Tyr-to-Asn substitution may be the result of, or result in, an alteration in the conformational flexibility of the linker region, and thus the interdomain dynamics of IF3. To test this hypothesis, the Y75N point mutation was introduced into the double Cys IF3 variant and the resulting construct was fluorescently labeled with Cy3 and Cy5 as described in section 2.3.3.

2.3.5 Verification of IF3 mutants' biochemical activity

It was important to determine whether attachment of the Cy3/Cy5 fluorophores to IF3's Cys residues perturbed the biochemical activity of the protein compared with unlabeled wild-type IF3. Traditionally, ensemble biochemical approaches are taken to verify the activity of reaction components for both ensemble and single-molecule fluorescence biochemical investigations. Considering that the environment of a surface-immobilized 5'-biotinylated-mRNA-bound 30S subunit may be different than that of freely mobile components in a test tube, I sought to develop an assay that probed a biochemical activity of the fluorescently-labeled IF3 on immobilized ribosomes using TIRFM. One of IF3's functions is to indiscriminately destabilize P-site tRNAs from 30S ICs [13]. This effect is most pronounced on 30S ICs lacking IF2 but containing IF1, IF3, mRNA, and aa-tRNA [13]. Thus, a TIRFM-based assay was developed that rapidly, and visually, probes IF3's ability to destabilize tRNAs, specifically fMet-(Cy3)tRNA^{fMet} (see Figure 2.4), based on this property. Details of the experimental set-up can be found in section 5.5 and the caption for

Figure 2.4. Briefly, 30S ICs were prepared side-by-side containing 5'-biotinylated mRNA, 30S subunits, IF1, and fMet-(Cy3)tRNA^{fMet}, either with or without IF3. Equal concentrations of the +/- IF3 complexes were delivered to channels in a microfluidic flowcell of a passivated quartz microscope slide (see section 2.2.7 and Figure 1.17 for a description of the flowcell). The number of immobilized fMet-(Cy3)tRNA^{fMet}-containing 30S subunits were quantified in each of these channels by determining a fluorescence intensity threshold and quantifying the number of fluorescent regions of this minimum intensity. Raw counts were converted to "%fMet-(Cy3)tRNA^{fMet} bound" by normalizing the (+)IF3 data to the (-)IF3 data (see Figure 2.4). The activity of unlabeled IF3(C65S-S38C-K97C) and IF3(C65S-S38C-K97C)-Cy5-Cy5 (here dual labeled with Cy5 due to the use of fMet-(Cy3)tRNA^{fMet}) is within error of wild-type IF3. Notably, Cy5-labeled wild-type IF3 has much lower activity than unlabeled IF3 in this assay (see section 2.3.1) and the Y75N point mutant effectively serves as a negative control in this assay, showing activity comparable to the (-)IF3 30S IC.

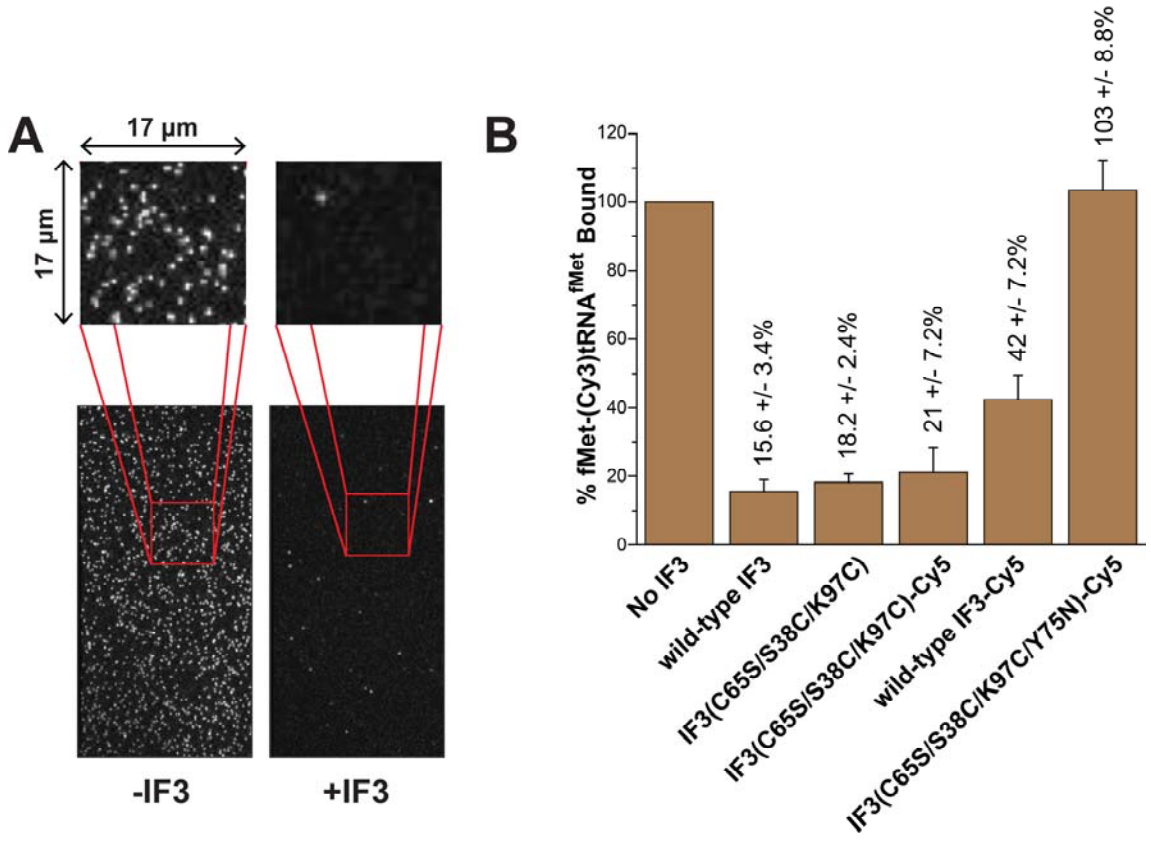

Figure 2.4 TIRFM-based IF3 activity assay. A. Representative Cy3 FOV for data collected in the absence (left) and presence (right) of wild-type IF3. **B.** Wild-type IF3, IF3(C65S/S38C/K97C) and IF3(Cy5-Cy5) display high biochemical activity in tRNA dissociation. Complexes containing biotin-mRNA, 30S subunits, IF1, IF3, and fMet-(Cy3)tRNA^{fMet} display high levels of tRNA dissociation. Error bars represent the standard deviation of the mean obtained from three independent data sets. Each of the five complexes was imaged in a separate channel in the same flow-cell. fMet-(Cy3)tRNA^{fMet} counts were normalized to those of the (-)IF3 complex imaged in the same flow-cell. Wild-type IF3 was Cy5 labeled at its native Cys65 residue. See text (section 2.3.5) for further details.

In addition to this TIRFM-based activity assay, the biochemical activity of all IF3 variants used in the work described in this chapter was also probed using a standard primer extension inhibition assay referred to as “toeprinting” which tests the activity of the ribosome and its components [8, 9, 43, 59]. This assay was employed to test the ability of the fluorescently labeled IF3 in promoting the selection of tRNA^{fMet} over tRNA^{Phe} on 30S ICs lacking IF1 and IF2 (see section 1.3.3.7) [8, 9]. Reactions were performed as developed and described by Hartz and Gold [8, 9], and detailed in Fei et al. [43] and in section 5.4. Briefly, initiation complexes were assembled with 30S subunits, a 1:10 mixture of tRNA^{fMet} to tRNA^{Phe}, and a ^{32}P-primer annealed-mRNA with a sequence derived from T4gp32 (Table 2.1, mRNA #4), containing an AUG start

codon followed by UUU, a codon cognate for $tRNA^{Phe}$. The identity of the P site-bound tRNA ($tRNA^{fMet}$ versus $tRNA^{Phe}$) can be determined based on the location of the complex on the mRNA (AUG versus UUU in the P site) (see Figure 2.5). This location is determined by reverse transcription of the ribosome-bound mRNA to generate a cDNA product. Reverse transcriptase (RT)'s cDNA synthesis is terminated when RT encounters a tRNA-bound 30S subunit on the mRNA. This signature, short cDNA appears as a "toeprint" band on a sequencing gel. RT is halted at the edge of the 30S subunit, 15 nucleotides downstream of the 5' nucleotide positioned in the P site (the "A" of the start codon if $tRNA^{fMet}$ is bound or the 5' "U" of the second codon if $tRNA^{Phe}$ is bound in the complex). The length of the cDNA product varies by three nucleotides (one codon) depending on whether $tRNA^{fMet}$ or $tRNA^{Phe}$ is bound in the complex, and the ratio of the intensity of the ^{32}P signal from these two bands is an effective read-out of the efficiency of complex formation containing either $tRNA^{fMet}$ or $tRNA^{Phe}$. In the absence of IF3, the 30S subunit does not discriminate between these two tRNAs and the resulting ratio of the toeprint band intensities is comparable to the tRNA concentrations (1:10 $tRNA^{fMet}$:$tRNA^{Phe}$). Addition of IF3 shifts tRNA selection toward $tRNA^{fMet}$, overcoming even the 10-fold excess of $tRNA^{Phe}$ (see Figure 2.5). As also seen in Figure 2.5, IF3(Cy3-Cy5), and each of the singly-labeled IF3 variants, achieves wild-type-like activity under increased IF3 concentrations. The need for slightly elevated IF3(Cy3-Cy5) concentrations to achieve wild-type-like activity suggests that introduction of the Cy3/Cy5 fluorophores may result in a binding defect of IF3. Any defect in IF3 binding did not hinder subsequent smFRET experiments, however, as the fluorescence/FRET signal indicated a 30S subunit-IF3(Cy3-Cy5) interaction, and IF3(Cy3-Cy5) remained stably bound to the immobilized 30S subunits over the course of data collection (20-30 minutes). Comparable to the results from the TIRFM-based assay, and as reported previously [5], the Y75N point mutant showed no ability to promote selection of $tRNA^{fMet}$ over $tRNA^{Phe}$ even at high concentrations in the toeprinting assay. Thus, despite this IF3 mutant's ability to bind to the 30S subunit with wild-type-like affinity [5], it is totally impaired in its ability to inspect the codon-anticodon interaction and reject $tRNA^{Phe}$.

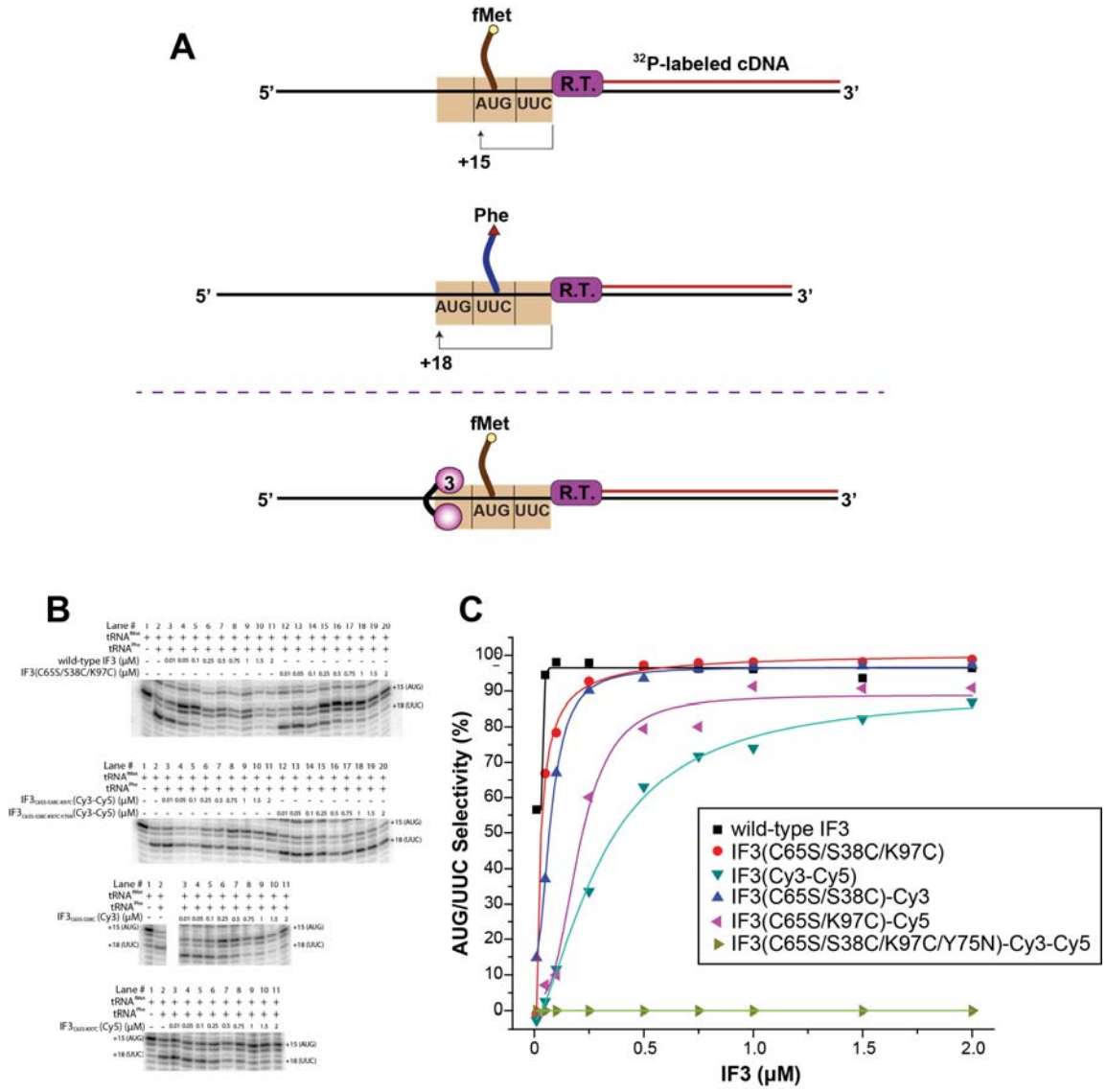

Figure 2.5 Primer extension inhibition ("toeprinting") IF3 activity assay. Toeprinting tests the ability of wild-type IF3, IF3(C65S/S38C/K97C), IF3(Cy3-Cy5), and the singly-labeled IF3 mutants IF3(C65S/S38C)-Cy3 and IF3(C65S/K97C)-Cy5 to promote selection of $tRNA^{fMet}$ over $tRNA^{Phe}$ on 30S subunits bound to ^{32}P-primer-annealed T4gp32 mRNA. **A.** Cartoon depiction of the toeprinting assay. "R.T." is reverse transcriptase. The 30S subunit shows a preference for $tRNA^{fMet}$ over $tRNA^{Phe}$ only in the presence of IF3. **B. top row.** Lane 1: (-)IF3, (-)$tRNA^{Phe}$. The strong stop at +15 indicates that $tRNA^{fMet}$ is bound to the AUG codon in the P site of the 30S subunit. Lane 2: (-)IF3, (+)$tRNA^{fMet}$, and (+)10-fold excess of $tRNA^{Phe}$ over $tRNA^{fMet}$. The strong stop at +18 indicates that $tRNA^{Phe}$ is bound to the UUU codon in the P site of the 30S subunit. Lanes 3-11: Increasing concentrations of wild-type IF3 shift the preference of the 30S subunit from $tRNA^{Phe}$ (+18) to $tRNA^{fMet}$ (+15). Lanes 12-20: Similar behavior is seen with increasing concentrations of unlabeled IF3(C65S/S38C/K97C). **B. second row.** IF3(C65S/S38C/K97C)-Cy3-Cy5 and the negative control IF3(C65S/S38C/K97C/Y75N)-Cy3-Cy5 (see section 2.3.5 of text). Lane numbering as above. **B. third row.** IF3(C65S/S38C)-Cy3. **B. bottom row.** IF3(C65S/K97C)-Cy5. **C.** Selectivity of $tRNA^{fMet}$ over $tRNA^{Phe}$ for each construct tested. Selectivity is defined as described in section 5.4.

2.3.6 30S ICs containing IF3(Cy3-Cy5) can be specifically surface-immobilized

Control experiments were performed to ensure that smFRET data collected with samples containing IF3(Cy3-Cy5) reflected the behavior of IF3 molecules that were 30S IC-bound and immobilized on the flowcell surface via a biotinylated mRNA. A channel-to-channel comparison within the same slide was performed to investigate the affinity, specificity, and interdomain configuration of 30S IC-bound IF3(Cy3-Cy5) (both $30SIC^{-tRNA}_{-1/2}$ and $30SIC^{-tRNA}_{-1}$) and free IF3(Cy3-Cy5) for the slide surface in the presence and absence of mRNA containing a biotin moiety. Complexes were prepared in parallel and delivered to each flowcell at identical concentrations. IF3(Cy3-Cy5) displays very limited affinity for the surface in the absence of ribosomal subunits (only 7% of delivered molecules were surface immobilized) (see Figure 2.6). $30SIC^{-tRNA}_{-1/2}$ complexes are >80% specifically immobilized, and the introduction of other initiation components, such as IF2 in $30SIC^{-tRNA}_{-1}$, brings that specificity to >90% (see below).

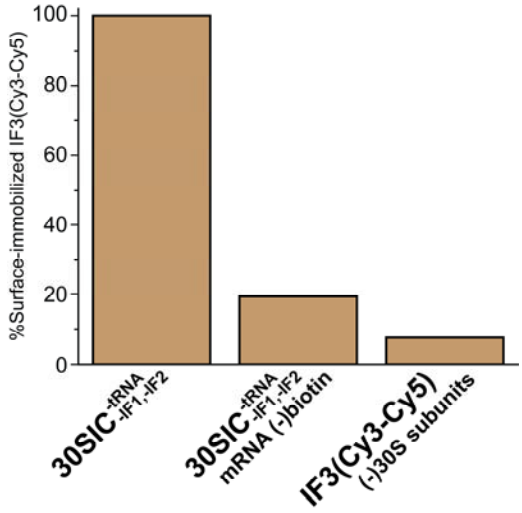

Figure 2.6 IF3(Cy3-Cy5) exhibits high 30S subunit affinity and minimal non-specific binding to biotin- and streptavidin-passivated quartz slide surfaces. A channel-to-channel comparison within the same flowcell was performed to investigate the affinity and specificity of $30SIC^{-tRNA}_{-IF1,IF2}$ and free IF3(Cy3-Cy5) for the slide surface in the presence and absence of mRNA containing a 5' biotin moiety. Complexes were prepared in parallel and injected into each flowcell at identical concentrations. The data were normalized to the $30SIC^{-tRNA}_{-IF1,IF2}$ results.

Formation of 30S complexes in the presence of IF2 further limits non-specific binding of IF3(Cy3-Cy5). A side-by-side comparison was done of $30SIC^{-tRNA}_{-IF1}$ complexes prepared identically

except with mRNA containing or lacking the 5' biotin moiety. Equal concentrations of complexes were delivered into adjacent channels of the same flowcell and for $30SIC^{-tRNA}_{-IF1}$ complexes formed on a 5'-biotinylated mRNA, the average number of immobilized molecules was 184 ± 28. An equal concentration of complexes formed with mRNA that was not biotinylated, and imaged in a separate channel of the same flowcell, showed an average of 17 ± 6 immobilized complexes. The error represents the standard deviation of the mean of ten movies collected on the same flowcell channel. Thus, $30SIC^{-tRNA}_{-IF1}$ complexes prepared with biotin-mRNA show greater than 90% biotin-streptavidin specific binding.

2.3.7 IF3 can adopt at least three distinct interdomain configurations on 30S ICs

In order to test the hypothesis that IF3 is conformationally dynamic on 30S ICs and that these dynamics may be linked to its roles in substrate selection, a series of partially, completely, and incorrectly (containing an non-initiator tRNA or non-start codon) assembled 30S ICs were prepared to represent the physiologically relevant 30S ICs in which IF3 may be found. First, the interdomain configuration of IF3 bound to an mRNA-immobilized 30S subunit in the absence of the other initiation components was investigated. Due to IF3's roles in both ribosome recycling and initiation (see section 1.3.3.6 and Figure 1.1), IF3 is expected to be the first initiation component to bind to the 30S subunit and may play roles in recruiting and stabilizing the other IFs [3]. smFRET versus time trajectories were recorded with 30S ICs lacking IF1, IF2, and tRNA ($30SIC^{-tRNA}_{-IF1,IF2}$) (see Figure 2.7). As is clear from both the individual trajectories and composite histogram (where the total number of trajectories, N, is 796), IF3(Cy3-Cy5) can adopt three distinct interdomain configurations, with the FRET efficiency values of these states centered at 0.24 ± 0.04, 0.43 ± 0.12, and 0.84 ± 0.06 (where the standard deviation is determined by ½(FWHM) for each Gaussian fitted peak) hereafter referred to as $IF3_{0.24}$, $IF3_{0.43}$, and $IF3_{0.84}$. Recalling that smFRET is an effective tool for measuring relative distance changes between fluorophores, with the FRET efficiency proportional to the inverse sixth power of the distance between the fluorophores, these three FRET efficiencies can be conceptualized as IF3

interdomain configurations. Assuming a Förster radius, R_0, of 55 Å for the Cy3/Cy5 FRET donor-acceptor pair, and isotropic and rapid rotation of Cy3 and Cy5 (i.e. $\kappa^2 = 2/3$) [50, 51], these FRET efficiencies correspond to fluorophore separation (R) values of ~67 Å, 58 Å, and 42 Å, respectively (see equation 1.1). Considering the location of the fluorophores on each of IF3's two globular domains, these values become a proxy for interdomain distances of 30S IC-bound IF3. Here these states likely reflect extended (0.24, 67 Å), intermediate (0.43, 42 Å), and compact (0.84, 42 Å) forms of IF3. Notably, the R values for 30S IC-bound IF3 are well within the range of interdomain distances that IF3 was observed to sample in an NMR study of IF3 free in solution (28 to 64 Å) [20] (see Figure 1.9). It is important to note that the broad FRET distribution of the intermediate state, $IF3_{0.43}$ (0.43 ± 0.12), arises due to its proximity to R_0, where small changes in distance (~1 Å) result in substantial changes in FRET efficiencies (see Figure 1.11B). The other states, $IF3_{0.24}$ and $IF3_{0.84}$ are further outside this sensitivity window, and consequently the observed FRET efficiency is less sensitive to changes in fluorophore distance and the corresponding distributions are therefore narrower.

While ~70% of the individual smFRET trajectories, from three independent data sets (Figure 2.9B, Table 2.3, and Appendix B), were observed to statically occupy either $IF3_{0.24}$, $IF3_{0.43}$, or $IF3_{0.84}$ prior to photobleaching, yielding fractional occupancies of 21 ± 5% ($IF3_{0.24}$), 35 ± 8% ($IF3_{0.43}$) and 13 ± 4% ($IF3_{0.84}$), ~30% of the smFRET trajectories exhibited at least one $IF3_{0.24}$ → $IF3_{0.43}$, $IF3_{0.43}$ → $IF3_{0.24}$, $IF3_{0.43}$ → $IF3_{0.84}$, or $IF3_{0.84}$ → $IF3_{0.43}$ transition. This analysis suggests that $30SIC_{-1/2}^{-tRNA}$-bound IF3 can interconvert between $IF3_{0.24}$, $IF3_{0.43}$, and $IF3_{0.84}$, but does so on a timescale much slower than our observation time (limited by the ~1-5 sec lifetime of the Cy5 fluorophore prior to photobleaching, with the exact fluorophore lifetime dependent on the FRET efficiency. See section 5.6).

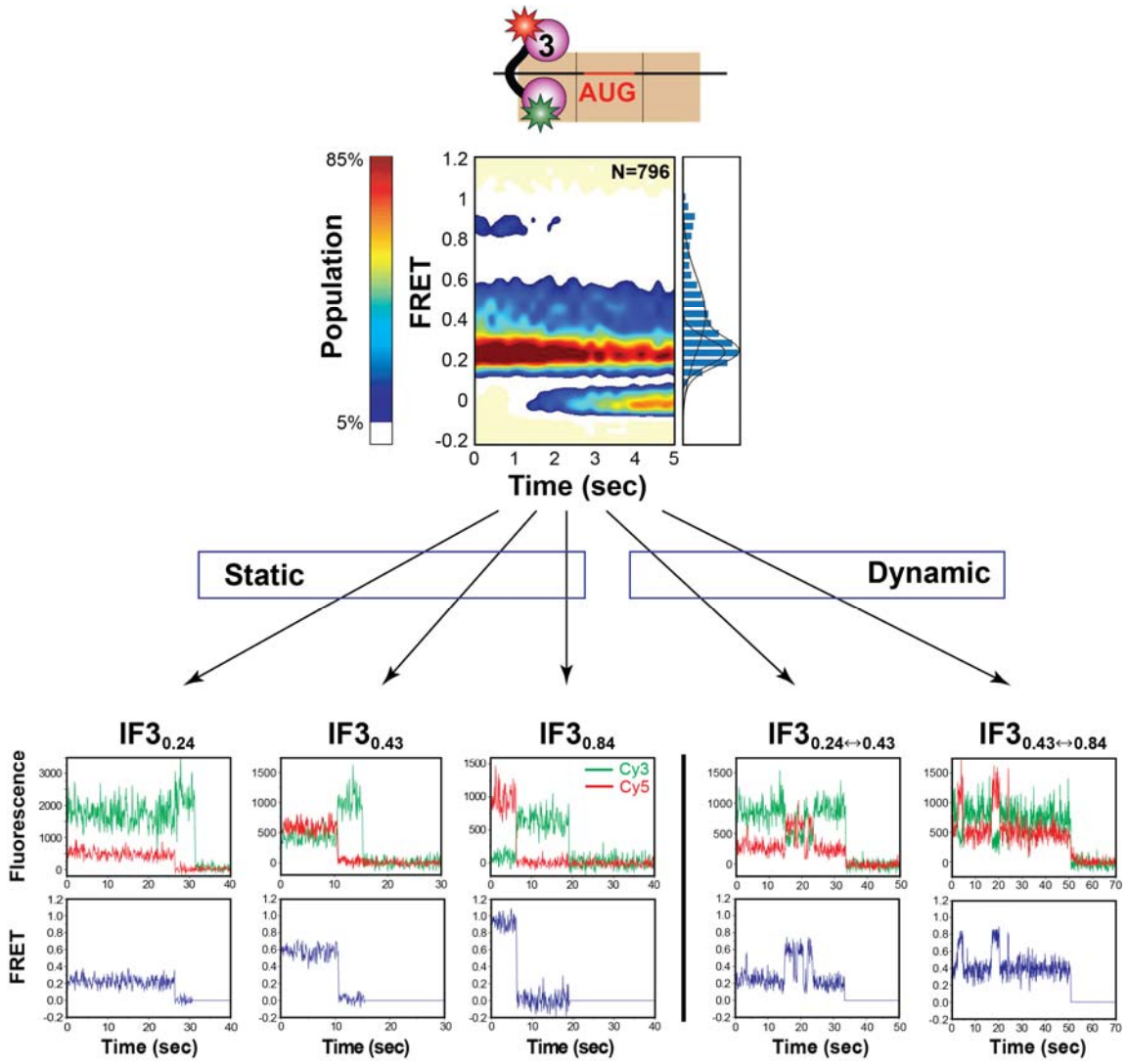

Figure 2.7 IF3 can adopt at least three distinct interdomain configurations on 30S ICs. (Top) cartoon depiction of the contents of the 30S IC. (Middle) Time-evolution of population FRET histogram. "N" represents the total number of FRET trajectories in the data set. The intensity at zero FRET efficiency arises due to Cy3 and Cy5 photobleaching.To the right of the contour plot is a normalized population FRET histogram of the first 0.5 seconds of data; bin size = 0.047 FRET. The population FRET histogram was fit with Gaussian distributions (black line). (Bottom) Representative examples of single-molecule fluorescence intensity and FRET versus time trajectories. Cy3 (green), Cy5 (red), and FRET (blue) trajectories for each of the five distinguishable subpopulations present in 30S IC-bound IF3(Cy3-Cy5) data sets. FRET efficiency is calculated as $I_{Cy5}/(I_{Cy3}+I_{Cy5})$. Representative static trajectories from $IF3_{0.24}$, $IF3_{0.43}$, and $IF3_{0.84}$, and interconvertible trajectories showing transitions between $IF3_{0.24}$ and $IF3_{0.43}$, and between $IF3_{0.43}$ and $IF3_{0.84}$. In each trajectory, the irreversible transition to zero FRET efficiency is indicative of photobleaching.

2.3.8 A structural interpretation of IF3's three 30S subunit-bound interdomain configurations

It is already clear from these $30SIC^{-tRNA}_{-IF1,IF2}$ data that IF3 can adopt at least three distinct interdomain configurations on 30S ICs and that it can interconvert between these states,

confirming part of my hypothesis (section 2.1). Interpreting these IF3 interdomain configurations in a structural and molecular context leads to a few possible models (Figure 2.8). First, one can consider the possibility that these dynamics arise from motion between IF3's domains in which one domain is bound to the 30S subunit while the other domain is detached from the ribosome and freely mobile due to IF3's flexible interdomain linker. This scenario is unlikely, however, considering the sub-microsecond expected timescale for motions of a small, two-domain protein in which the two domains are connected by a flexible linker [60]. This is at least five orders of magnitude faster than the time resolution of the TIRF microscope employed here, thus this type of behavior would be detected as a single, average FRET efficiency rather than three distinct, slowly interconverting FRET efficiencies.

Another possible model to explain the dynamics of 30S subunit-bound IF3 is that these dynamics reflect conformational dynamics of the 30S subunit to which IF3 remains tightly bound (Figure 2.8A). It has been shown that the head domain of the 30S subunit is dynamic in the absence of a P-site codon-anticodon interaction and freely rotates about the neck helix in 16S rRNA [61]. Considering that IF3 has been crosslinked with a number of ribosomal proteins located in the head domain, including S7, S13 and S19, as well as S11 on the platform, near the neck (Figure 1.11), it is possible that IF3's interactions with these ribosomal proteins may be reflect these 30S subunit dynamics.

A third scenario to consider is the possible existence of three IF3 binding sites on the 30S subunit (Figure 2.8B). Although there are now a number of studies which agree that IF3's CTD binds on the platform near the P site [22, 27, 28, 32], many of these same studies disagree on the 30S subunit binding site for IF3's NTD [22, 27, 28]. The CTD binds with high affinity to rRNA even in the absence of the NTD [24, 62], while the NTD is thought to bind to ribosomal proteins and cannot bind to the 30S subunit in the absence of the linker and CTD [24]. Considering the number of 30S subunit ribosomal proteins that IF3 has been crosslinked with, and the differing ribosomal affinities of the two IF3 domains, it is possible that IF3's CTD remains tightly bound on the platform region of the 30S subunit while IF3's NTD samples at least three separate binding sites

spanning a broad area of the subunit. Support for the CTD's stable binding to the 30S subunit comes from another smFRET signal that I developed between IF1 and IF3's CTD (see Chapter 4). The CTD of IF3 remains anchored to the 30S subunit regardless of the presence or absence of other initiation components. It has yet to be shown with this IF1-IF3 smFRET signal if the NTD is able to sample three binding sites, in support of this model.

Lastly, it is possible, and even likely, that these dynamics reflect a combination of scenarios two and three (A and B in Figure 2.8), as well as others not mentioned here (see Figure 2.8C).

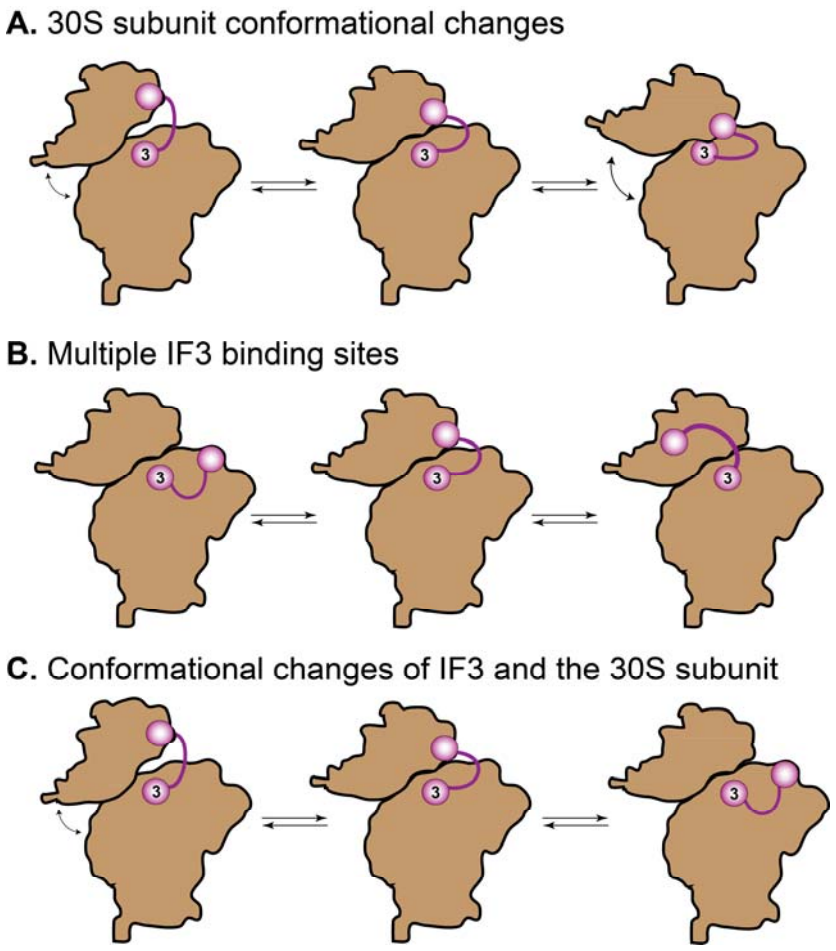

Figure 2.8 Possible models for the three smFRET efficiency values observed with 30S subunit-bound IF3. A. Dynamics of the 30S subunit head could result in relative repositioning of the tightly bound IF3 CTD and NTD. **B.** IF3 may have multiple binding sites, and dissociation from one site and rebinding to another site could result in different interdomain distances. **C.** The three interdomain configurations of IF3 may arise due to a combination of 30S subunit rearrangements and IF3 rearrangements on the 30S subunit.

2.3.9 IF3's 30S IC-bound interdomain configuration is influenced by the presence of the other IFs

Clearly IF3 can adopt at least three distinct interdomain configurations on 30S ICs and it can interconvert between these states, however it remained to be demonstrated if these dynamics are linked to IF3's function in substrate selection. IF3 has been shown to regulate initiator tRNA and start codon selection together with IF1 and IF2. Thus, a systematic approach was taken to first investigate the influence of these other IFs on the interdomain configurations and dynamics of 30S IC-bound IF3, before embarking on investigations of the tRNA and mRNA substrates. The effects of IF1 and IF2 on the interdomain configuration of 30S IC-bound IF3 were investigated by imaging 30S ICs similar to $30SIC^{-tRNA}_{-IF1,IF2}$ but prepared and imaged in the presence of saturating amounts of IF1 ($30SIC^{-tRNA}_{-IF2}$), IF2 ($30SIC^{-tRNA}_{-IF1}$), or both IF1 and IF2 ($30SIC^{-tRNA}$) (see Figure 2.9). The presence of IF1 shifts the conformational equilibria of 30S IC-bound IF3 further toward $IF3_{0.43}$ compared with $30SIC^{-tRNA}_{-IF1,IF2}$, yielding fractional occupancies of 16 ± 3% ($IF3_{0.24}$), 54 ± 1% ($IF3_{0.43}$) and 5 ± 1% ($IF3_{0.84}$). In contrast, the presence of IF2 dramatically shifts the conformational equilibria of 30S IC-bound IF3 toward $IF3_{0.84}$, yielding fractional occupancies of 5 ± 4% ($IF3_{0.24}$), 10 ± 5% ($IF3_{0.43}$), and 82 ± 9% ($IF3_{0.84}$) (see Figure 2.9). Notably, the $IF3_{0.84}$ state is largely inaccessible to 30SIC-bound IF3 in the presence of IF1, regardless of whether or not IF2 is present. This $IF3_{0.84}$ state predominates when IF2 is bound in the absence of IF1, however, and is also accessible to IF3 in the absence of both IF1 and IF2, on $30SIC^{-tRNA}_{-IF1,IF2}$. This suggests that a function of IF1 is to regulate the accessibility of this state, whether it is a conformation of the 30S subunit or an IF3 binding site (see model, Figure 2.8). The intracellular concentrations of the three IFs are tightly regulated and should be approximately equal under normal physiological conditions [63, 64]. Additionally, free 30S subunits should have all three IFs bound >99% of the time [63], so the most physiologically relevant 30S IC to consider is $30SIC^{-tRNA}$, which almost exclusively favors $IF3_{0.24}$ and $IF3_{0.43}$, yielding fractional occupancies of 27 ± 7% ($IF3_{0.24}$), 49 ± 3% ($IF3_{0.43}$), and only 5 ± 3% ($IF3_{0.84}$) (see Figure 2.9).

In considering the influence of the other initiation factors on the interdomain configuration of 30S IC-bound IF3, currently available data suggest that IF1 and IF3 do not directly contact each other within the 30S IC [65, 66], indicating that the influence exerted by IF1 over the conformational equilibria of 30S IC-bound IF3 is indirectly mediated via the 30S subunit. IF3 is known to increase the affinity of IF1 for the 30S subunit [64], and binding of IF1 triggers both local and global conformational changes of the 30S subunit, including disruption of the noncanonical A1413-G1487 base pair in helix 44, and other long-range conformational changes of 16S rRNA helix 44 and the 30S subunit neck and head domains [67]. These conformational changes may affect IF3's interactions with the 30S subunit since IF3 binding to the 30S subunit also triggers enhanced reactivity of G1487 toward the chemical probe kethoxal, suggesting that IF3 binding may affect helix 44 in a similar manner to IF1 [29]. Additionally, IF1's binding site neighbors that of ribosomal protein S12, a protein that has been crosslinked to IF3 in multiple studies [68-72] (see Figure 1.11). Finally, the eukaryotic counterparts of IF1 and IF3 (eIF1A and eIF1, respectively) have been shown to cooperatively trigger large-scale conformational changes of the 40S subunit [30], suggesting the possibility of a similar cooperation on the prokaryotic ribosome by IF1 and IF3.

In contrast, for $30SIC_{-1}^{-tRNA}$, IF2 has been crosslinked with IF3 [65], suggesting a direct physical interaction between these two factors. It remains unclear, however, whether IF2 exerts its influence over the conformational equilibria of 30S IC-bound IF3 through direct contacts with IF3, by indirectly modulating the conformation of the 30S subunit, or through a combination of both of these effects. The 30S IC containing all three IFs, $30SIC^{-tRNA}$, may have a network of interactions among the IFs. IF2 and IF1 have been crosslinked [65] and have been shown to directly contact each other in a recent 70S IC cryo-EM structure [66], though a 30S IC cryo-EM structure failed to identify a direct interaction between these two factors. IF1 and IF2 have been functionally replaced by a single factor, $IF2_{mt}$, in mammalian mitochondria, however, which can be taken as evidence for their cooperative roles in initiation.

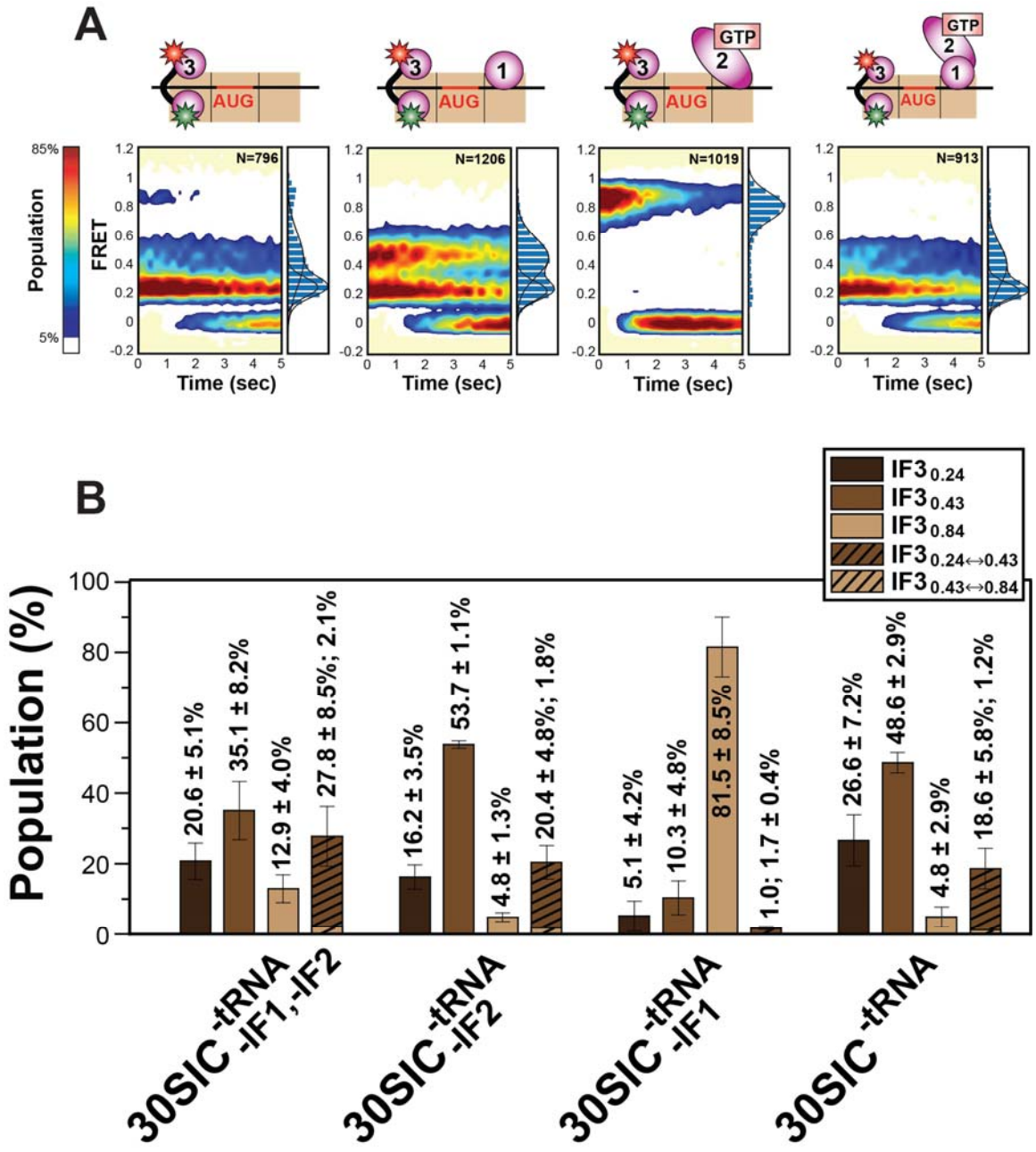

Figure 2.9 The interdomain configuration of 30S IC-bound IF3 is affected by IF1 and IF2. A. Time-evolution of population FRET histograms for $30SIC^{-tRNA}_{-IF1,IF2}$, $30SIC^{-tRNA}_{-IF2}$, $30SIC^{-tRNA}_{-IF1}$, and $30SIC^{-tRNA}$. The cartoon above each contour plot indicates the contents of each 30S IC. See Figure 1.2 for further details on the 30S IC cartoons. "N" represents the total number of FRET trajectories in each data set. The data are combined from three independent data sets. To the right of each contour plot is a normalized population FRET histogram of the first 0.5 seconds of data; bin size = 0.047 FRET. Each histogram was fit with Gaussian distributions (black line). The intensity at zero FRET arises due to Cy3 and Cy5 photobleaching. **B.** Subpopulation analysis of smFRET versus time trajectories for $30SIC^{-tRNA}_{-IF1,IF2}$, $30SIC^{-tRNA}_{-IF2}$, $30SIC^{-tRNA}_{-IF1}$, and $30SIC^{-tRNA}$. The mean and standard deviation of the percentage occupancy of each subpopulation were calculated from three independent data sets (see Appendix B).

2.3.10 IF3 undergoes an interdomain reconfiguration on a completely assembled 30S IC

To further test the hypothesis that IF3's interdomain configurations are linked to its role in substrate selection, a 30S IC was assembled which contained all the components of a correctly assembled 30S initiation complex in order to investigate IF3's interdomain configuration upon initiator tRNA and start codon selection. These complexes should have fMet-tRNA^{fMet} stably bound to the P site through interactions with 16S rRNA residues A1339, G1338, and A790, as well as Watson-Crick base pairing with the AUG codon on the mRNA. In these complexes, the 30S IC-bound IF3 predominantly (76 ± 10% of all trajectories) occupies IF3$_{0.84}$ (Figure 2.10 and 2.11B), with 4 ± 4% fractional occupancy in IF3$_{0.24}$ and 15 ± 7% in IF3$_{0.43}$. By comparing the behavior of 30S IC-bound IF3 on 30SIC^{fMet} with 30SIC^{-tRNA}, it is obvious that the presence of fMet-tRNA^{fMet} bound to its cognate AUG codon triggers a significant interdomain reconfiguration of 30S IC-bound IF3 (Figure 2.10). This striking comparison suggests that the interdomain configuration of the 30S IC-bound IF3 is responsive to the presence of a stable P-site codon-anticodon interaction.

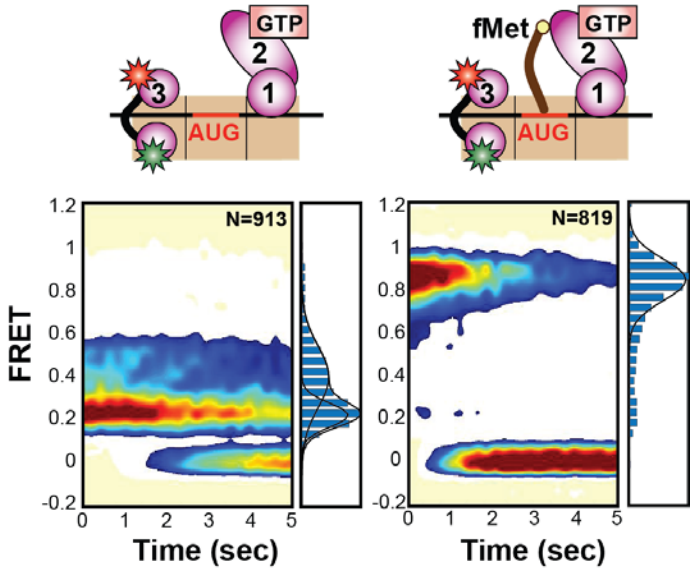

Figure 2.10 An IF3 interdomain reconfiguration signals proper initiator tRNA and start codon selection during translation initiation. Time-evolution of population FRET histograms for 30SIC^{-tRNA} and 30SIC^{fMet}. Normalized population FRET histograms for the first five frames of data are shown to the right of each contour plot. IF3 undergoes an interdomain reconfiguration in the presence of initiator tRNA and an AUG start codon.

The interdomain reconfiguration of 30S IC-bound IF3 does not occur uniquely in $30SIC^{fMet}$, however. Recall that the $IF3_{0.84}$ state favored here is also the state favored by IF3 on the incompletely assembled $30SIC^{-tRNA}_{-IF1}$ (Figure 2.9). Thus, the molecular implications of this IF3 interdomain reconfiguration, whether, for example, in signaling the complete assembly of a 30S IC or promoting rapid 50S subunit joining (see section 2.4 and Chapter 3), are that this reconfiguration may be a necessary but insufficient means to perform these roles. Instead, it is likely that numerous conformational rearrangements of the 30S subunit and its initiation components cooperatively signal 30S IC assembly. Whether or not the interdomain configurations favored by IF3 on these different complexes is exactly the same at the atomic level, however, cannot be known from these smFRET data. These experiments are limited by only revealing one distance constraint within these macromolecular complexes. Additionally, the FRET efficiency here, 0.84, is outside the optimal window of sensitivity to distance separations (see Figure 1.15), so small changes in distance would not be detected with this smFRET signal. Other smFRET signals could be designed which more precisely reveal the IF3 NTD-CTD separation on these complexes by moving the fluorophores closer to R_0 (55 Å) for these 30S ICs. Nevertheless, these $30SIC^{fMet}$ data reveal the responsiveness of IF3's interdomain configuration to the presence of correctly selected substrates.

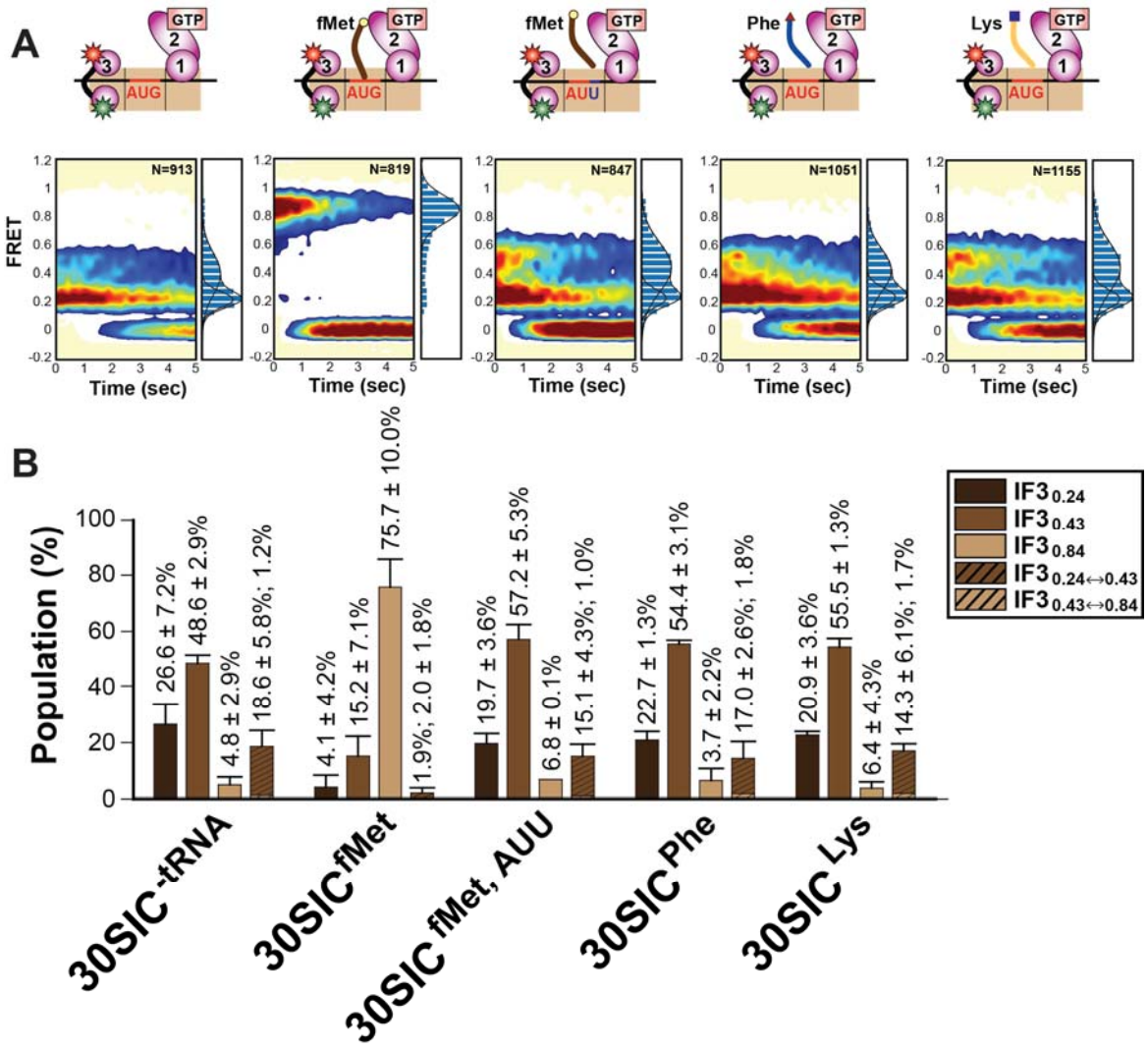

Figure 2.11 IF3 fails to undergo an interdomain reconfiguration on incorrectly assembled 30S ICs. A. Time-evolution of population FRET histograms for 30SIC^{-tRNA}, 30SIC^{fMet}, 30SIC^{fMet,AUU}, 30SIC^{Phe}, and 30SIC^{Lys}. **B.** Subpopulation analysis of smFRET versus time trajectories for 30SIC^{-tRNA}, 30SIC^{fMet}, 30SIC^{fMet,AUU}, 30SIC^{Phe}, and 30SIC^{Lys}. The mean and standard deviation of the percentage occupancy of each subpopulation were calculated from three independent data sets. The non-cognate codon-anticodon interaction that may be occurring in 30SIC^{fMet,AUU}, 30SIC^{Phe}, and 30SIC^{Lys} is indicated in the cartoons.

2.3.11 IF3 fails to undergo an interdomain reconfiguration on incorrectly assembled 30S ICs

If indeed this interdomain reconfiguration of 30S IC-bound IF3 occurs uniquely on properly assembled 30S ICs (i.e. 30SIC^{fMet}), then it follows that IF3 bound to an incorrectly assembled 30S IC, containing a non-canonical start codon or elongator tRNA, should fail to undergo this reconfiguration, and instead remain in another interdomain configuration. Thus, 30S

ICs were assembled with a 5'-biotinylated mRNA (mRNA #2, Table 2.1) containing the non-canonical AUU start codon (see section 1.2.2). IF3 has a known sensitivity to AUU as it is the start codon IF3 uses to autoregulate the expression of its own gene, *infC* (see section 1.3.3.8) [73, 74]. Confirming the substrate selection-dependence to the IF3 interdomain reconfiguration, the $IF3_{0.84}$ state of 30S IC-bound IF3 on these complexes, hereafter $30SIC^{fMet,AUU}$, was minimally occupied (7 ± 0.1%), unlike in $30SIC^{fMet}$. Instead, $30SIC^{fMet,AUU}$ showed high fractional occupancies of $IF3_{0.24}$ and $IF3_{0.43}$: 20 ± 4% and 57 ± 5%, respectively, of all smFRET trajectories (Figure 2.11).

Table 2.3 Subpopulation analysis of 30S IC-bound IF3(Cy3-Cy5) under various conditions. "30SIC" represents a complex containing IF3(Cy3-Cy5) and a biotinylated mRNA with an AUG start codon. The subscript following "30SIC" for all other complexes indicates which IFs were absent; the superscript indicates which tRNA, and mRNA start codon if not AUG, was used in the 30S IC. The mean and standard deviation of the percentage occupancy of each subpopulation were calculated from three independent data sets. All trajectories were idealized using vbFRET [48] and assigned to states as follows: $IF3_{0.24}$ (0-0.25 FRET), $IF3_{0.43}$ (0.3-0.65 FRET), and $IF3_{0.84}$ (0.7-1 FRET). These states were determined using the full-width at half maximum (FWHM) of Gaussian fits to population histograms of the first 0.5 sec of data. The shaded boxes indicate the most highly populated state within that 30S IC.

Complex	$IF3_{0.24}$ (%)	$IF3_{0.43}$ (%)	$IF3_{0.84}$ (%)	$IF3_{0.24\leftrightarrow0.43}$ (%)	$IF3_{0.43\leftrightarrow0.84}$ (%)
	Static			Dynamic	
$30SIC^{-tRNA}_{-1/2}$	20.6 ± 5.1	**35.1 ± 8.2**	12.9 ± 4.0	27.8 ± 8.5	2.1 ± 0
$30SIC^{-tRNA}_{-2}$	16.2 ± 3.5	**53.7 ± 1.1**	4.8 ± 1.3	20.4 ± 4.8	1.8 ± 0
$30SIC^{-tRNA}_{-1}$	5.1 ± 4.2	10.3 ± 4.8	**81.5 ± 8.5**	1.0 ± 0	1.7 ± 0.4
$30SIC^{-tRNA}$	26.6 ± 7.2	**48.6 ± 2.9**	4.8 ± 2.9	18.6 ± 5.8	1.2 ± 0
$30SIC^{fMet}$	4.1 ± 4.2	15.2 ± 7.1	**75.7 ± 10.0**	1.9 ± 0	2.0 ± 1.8
$30SIC^{fMet,AUU}$	19.7 ± 3.6	**57.2 ± 5.3**	6.8 ± 0.1	15.1 ± 4.3	1.0 ± 0
$30SIC^{Phe}$	22.7 ± 1.3	**54.4 ± 3.1**	3.7 ± 2.2	17.0 ± 2.6	1.8 ± 0
$30SIC^{Lys}$	20.9 ± 3.1	**55.5 ± 1.3**	6.4 ± 4.3	14.3 ± 6.1	1.7 ± 0

To further test the substrate dependence of IF3's interdomain reconfiguration, 30S ICs were formed on an mRNA containing an AUG start codon but with either Phe-$tRNA^{Phe}$ or Lys-$tRNA^{Lys}$, both of which are non-cognate with AUG. As seen with $30SIC^{fMet,AUU}$, $30SIC^{Phe}$ and

30SIC^{Lys} showed very low occupancy of IF3$_{0.84}$, the compact IF3 interdomain configuration favored in 30SIC^{fMet} (4 ± 2% and 6 ± 4%, respectively), instead favoring IF3$_{0.24}$ (23 ± 1% and 21 ± 3%) and IF3$_{0.43}$ (54 ± 3% and 56 ± 1% for 30SIC^{Phe} and 30SIC^{Lys}, respectively) (Figure 2.11 and Table 2.3). These results indicate that IF3's interdomain configuration is responsive to the presence of a cognate P-site codon-anticodon interaction, and that it fails to undergo an interdomain reconfiguration in its absence.

2.3.12 IF3 fails to undergo an interdomain reconfiguration on 30S ICs containing a cognate elongator tRNA codon-anticodon interaction

I was next interested in investigating whether the 30S IC-bound interdomain configuration of IF3 is uniquely responsive to a cognate AUG start codon-initiator tRNA anticodon interaction or if it responds to a stable, cognate P-site codon-anticodon interaction formed from any cognate codon and anticodon (i.e. initiator or elongator). 30S ICs were assembled with Phe-tRNA^{Phe} and an mRNA containing a UUC codon cognate for Phe-tRNA^{Phe} in place of an AUG codon at the location of the start codon (see Table 2.1 for mRNA #3 sequence). tRNA^{Phe} naturally has a higher propensity to bind to the P site than most elongator tRNAs since its anticodon stem contains the two GC base pairs (G29-C41 and G30-C40) needed for stable interaction with the P site 16S rRNA (bases A1339 and G1338) (see Figure 2.12) [61, 75, 76].

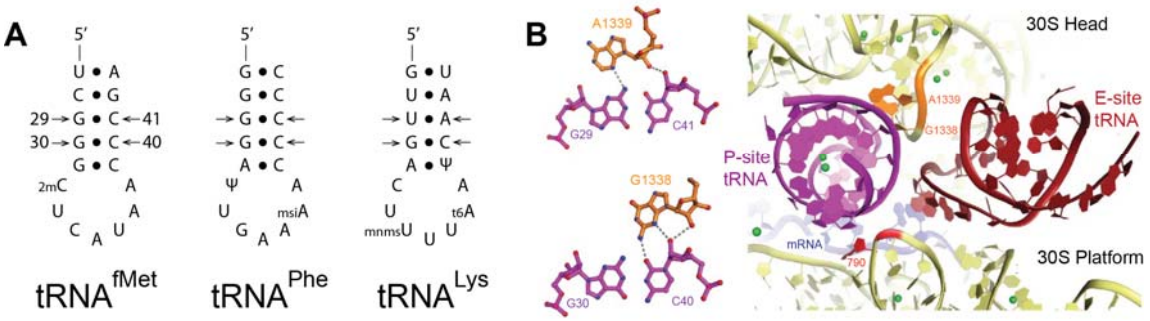

Figure 2.12 P-site tRNA-mRNA interactions. A. Anticodon stem-loop (ASL) sequences of tRNA^{fMet}, tRNA^{Phe}, and tRNA^{Lys} with anticodon stem base pairs critical for P site interactions indicated with arrows. Figure adapted from Ref. [75]. **B.** tRNA anticodon stem base pair G29-C41 interacts with A1339 to form a Type I A-minor interaction in the P site, while G30-C40 interacts with P site residue G1338 to form a Type II A-minor interaction. 16S rRNA base 790 also interacts with the opposite side of the ASL to prevent the tRNA's transit into the E site. Figure from Ref. [76].

Despite this structural advantage, the tRNA^{Phe} codon-anticodon interaction, if indeed stably present in 30SIC^{Phe,UUC}, failed to trigger the interdomain reconfiguration of 30S IC-bound IF3 that was triggered in 30SIC^{fMet} (Figure 2.12). This result is in line with previous work demonstrating that IF3 is responsive not just to codon-anticodon complementarity rules, but also to the unique features of tRNA^{fMet}'s anticodon stem, loop, and the anticodon itself [12]. A few studies have shown that IF3 only selects AUG-, GUG-, and UUG-tRNA^{fMet} codon-anticodon P-site interactions, never UUU-tRNA^{Phe} or AUU-tRNA^{fMet} interactions [8, 12].

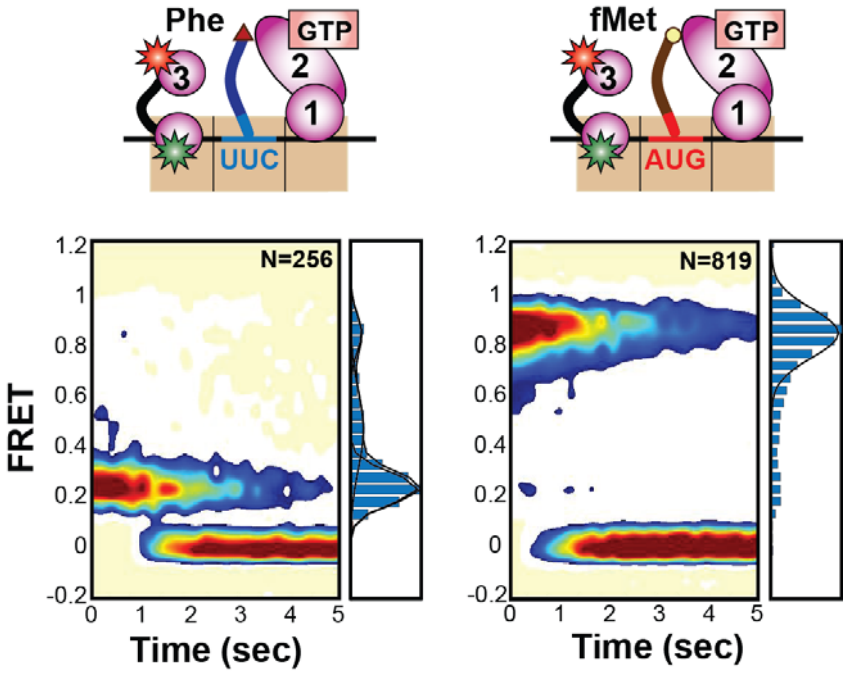

Figure 2.13 An IF3 interdomain reconfiguration requires a cognate AUG start codon-initiator tRNA anticodon interaction. A cognate elongator codon-anticodon interaction between Phe-tRNA^{Phe} and a UUC codon fails to trigger the IF3 interdomain reconfiguration that occurs with a cognate initiator AUG-fMet-tRNA^{fMet} codon-anticodon interaction. Time-evolution of population FRET histogram and 1D histogram for both complexes. The Phe-UUC 30S IC results need to be triplicated.

2.3.13 The loss-of-function point mutation, Y75N, has a unique smFRET signature

Lastly, the previously identified and characterized loss-of-function point mutant, IF3-Y75N [4, 5], which has a known substrate selection defect but maintains the ability to bind to 30S subunits was chosen to further test the connection between substrate selection and IF3's 30S IC-bound interdomain configuration. Y75N was introduced into the double Cys IF3 construct –

IF3(C65S-S38C-K97C-Y75N) – and dual labeled with Cy3 and Cy5 as described (see section 2.3.1). Use of this sample in fully assembled 30S ICs (i.e. $30SIC^{fMet}$) revealed a novel IF3 interdomain configuration favored by this IF3 variant (Figure 2.13) compared with others tested. Although the $IF3_{0.84}$ interdomain configuration is sampled by a sizable fraction of the sample population (19%), the $IF3_{0.43}$ state predominates in these 30S ICs. These results reveal a novel phenotype for this point mutation, that being a unique 30S IC-bound interdomain configuration compared with wild-type IF3 (with wild-type being the IF3(Cy3-Cy5) variant). Considering the phenotypes of this point mutant (loss of ability to select the initiator tRNA, inability to discriminate start codons, and inability to inhibit initiation on leaderless mRNAs), these results support the hypothesis that the interdomain configuration of IF3 on the 30S is critical for performing these functions. The proximity of Tyr75 to the linker suggests that its flexibility, or its interactions with the 30S subunit, may be impaired and restrict the protein's ability to sample its wild-type configurations.

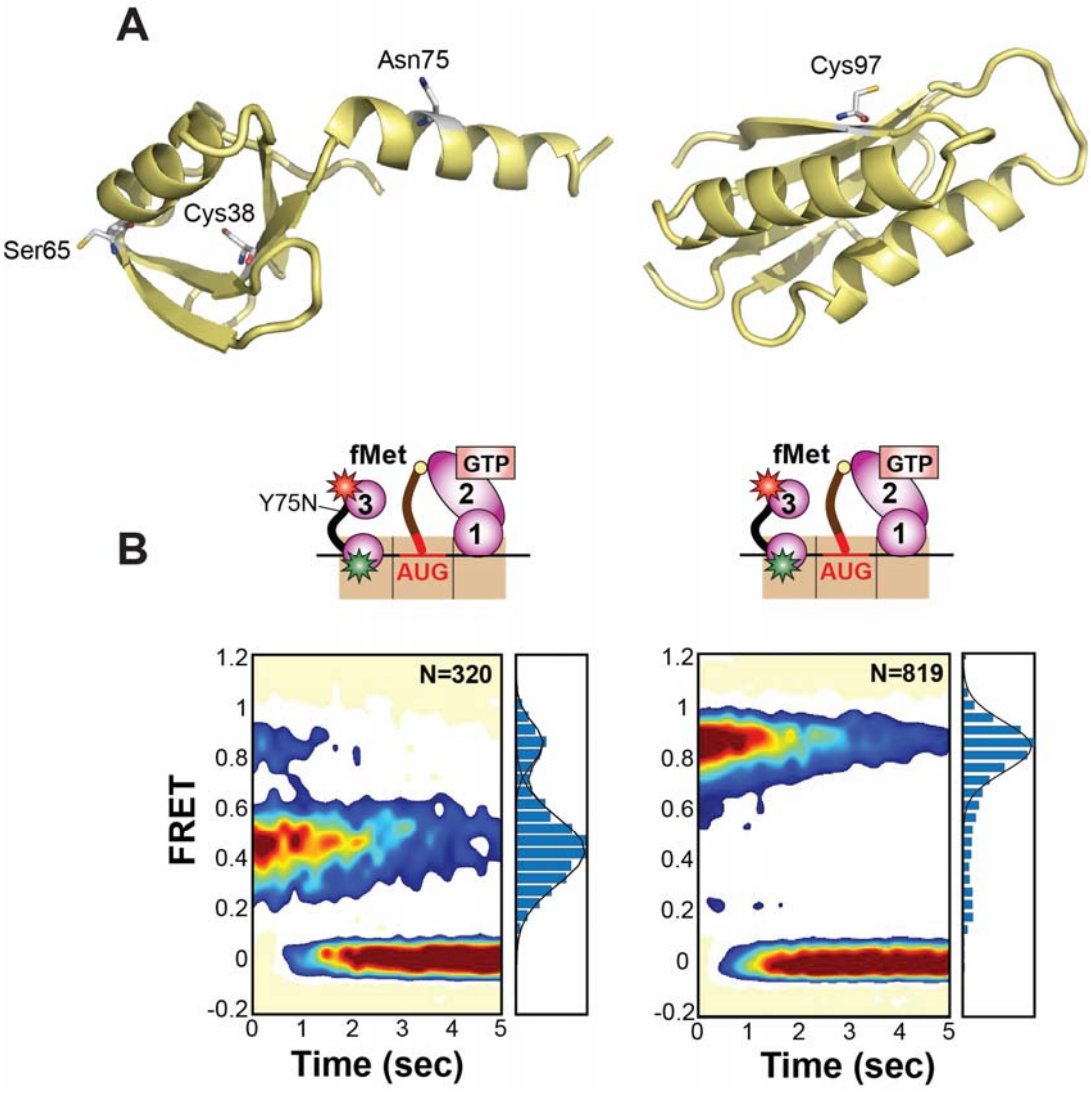

Figure 2.14 The loss-of-function IF3 point mutation Y75N exhibits a unique 30S IC-bound interdomain configuration. A. Structural model of IF3(C65S-S38C-K97C) containing the loss-of-function substitution mutation Y75N. Ribbon diagram of IF3 from *B. stearothermophilus* with *E. coli* IF3 numbering [77, 78]. Residues 79 to 82 are missing, and the protruding NTD α-helix (residues 61 to 78 (*B. stearo* numbering)) is flexible in *E. coli* [20]. Residues C65S, S38C, K97C, and Y75N are labeled. IF3 was modeled and rendered in PyMOL [79]. PDB codes 1TIF (NTD) and 1TIG (CTD). **B.** Time-evolution of population FRET histogram and normalized population histogram for the IF3-Y75N(Cy3-Cy5) 30S IC and IF3(Cy3-Cy5) 30S IC complexes. The IF3-Y75N(Cy3-Cy5) results need to be triplicated.

2.4 Conclusions and Future Directions

Taken together, these results reveal that 30S IC-bound IF3 can adopt at least three interdomain configurations, and the equilibrium between these configurations is regulated by the IFs and the P-site codon-anticodon interaction. This equilibrium converges on a single 30S IC-

bound IF3 interdomain configuration upon correct initiator tRNA and start codon selection. These results support the hypothesis that the interdomain configuration of 30S IC-bound IF3, and perhaps the 30S IC itself, is important in substrate selection and may signal complete 30S IC assembly. The existence of multiple 30S IC-bound IF3 interdomain configurations may explain the discrepancies among 30S-IF3 structural studies, as well as direct future structural investigations of 30S ICs. Considering the steps in the initiation pathway that follow 30S IC assembly, namely 50S subunit joining to form a 70S IC, it is possible that IF3's interdomain reconfiguration on completely and correctly assembled 30S ICs may occur to promote 50S subunit joining, and possibly its own dissociation from the ribosome. In contrast, IF3's other 30S IC interdomain configurations may restrict these events, perhaps by occupying binding sites on the 30S subunit which prevent intersubunit bridge formation, and thus also preventing rapid and stable 50S subunit joining. A mechanistic model relating IF3's 30S IC-bound interdomain configuration to 50S subunit joining is developed and presented in Chapter 3.

The role of IF3's interdomain configurations and dynamics in substrate selection can be tested further with IF3 mutants in which the linker length or flexibility is altered toward the aim of disrupting IF3's ability to sample all of its accessible interdomain configurations. It would also be interesting to test the less frequently used start codons GUG and UUG (see section 1.2.2), in addition to the non-canonical CUG codon. Lastly, chimeric tRNAs, such as fMet-tRNA^{Phe} or Phe-tRNA^{fMet}, or tRNA^{fMet} containing alterations in its anticodon stem or loop, can be employed to tested the structural features of the tRNA substrate which are discriminated by IF3 [80].

2.5 References

1. Sonenberg, N. and A.G. Hinnebusch, *New modes of translational control in development, behavior, and disease.* Mol Cell, 2007. **28**(5): p. 721-9.

2. Sonenberg, N. and A.G. Hinnebusch, *Regulation of translation initiation in eukaryotes: mechanisms and biological targets.* Cell, 2009. **136**(4): p. 731-45.

3. Boelens, R. and C.O. Gualerzi, *Structure and function of bacterial initiation factors.* Curr Protein Pept Sci, 2002. **3**(1): p. 107-19.

4. Sussman, J.K., E.L. Simons, and R.W. Simons, *Escherichia coli translation initiation factor 3 discriminates the initiation codon in vivo.* Mol Microbiol, 1996. **21**(2): p. 347-60.

5. Maar, D., D. Liveris, J.K. Sussman, S. Ringquist, I. Moll, N. Heredia, A. Kil, U. Blasi, I. Schwartz, and R.W. Simons, *A single mutation in the IF3 N-terminal domain perturbs the fidelity of translation initiation at three levels.* J Mol Biol, 2008. **383**(5): p. 937-44.

6. Haggerty, T.J. and S.T. Lovett, *IF3-mediated suppression of a GUA initiation codon mutation in the recJ gene of Escherichia coli.* J Bacteriol, 1997. **179**(21): p. 6705-13.

7. Sacerdot, C., C. Chiaruttini, K. Engst, M. Graffe, M. Milet, N. Mathy, J. Dondon, and M. Springer, *The role of the AUU initiation codon in the negative feedback regulation of the gene for translation initiation factor IF3 in Escherichia coli.* Mol Microbiol, 1996. **21**(2): p. 331-46.

8. Hartz, D., J. Binkley, T. Hollingsworth, and L. Gold, *Domains of initiator tRNA and initiation codon crucial for initiator tRNA selection by Escherichia coli IF3.* Genes Dev, 1990. **4**(10): p. 1790-800.

9. Hartz, D., D.S. McPheeters, and L. Gold, *Selection of the initiator tRNA by Escherichia coli initiation factors.* Genes Dev, 1989. **3**(12A): p. 1899-912.

10. Risuleo, G., C. Gualerzi, and C. Pon, *Specificity and properties of the destabilization, induced by initiation factor IF-3, of ternary complexes of the 30-S ribosomal subunit, aminoacyl-tRNA and polynucleotides.* Eur J Biochem, 1976. **67**(2): p. 603-13.

11. Berkhout, B., C.J. van der Laken, and P.H. van Knippenberg, *Formylmethionyl-tRNA binding to 30 S ribosomes programmed with homopolynucleotides and the effect of translational initiation factor 3.* Biochim Biophys Acta, 1986. **866**(2-3): p. 144-53.

12. Meinnel, T., C. Sacerdot, M. Graffe, S. Blanquet, and M. Springer, *Discrimination by Escherichia coli initiation factor IF3 against initiation on non-canonical codons relies on complementarity rules.* J Mol Biol, 1999. **290**(4): p. 825-37.

13. Antoun, A., M.Y. Pavlov, M. Lovmar, and M. Ehrenberg, *How initiation factors maximize the accuracy of tRNA selection in initiation of bacterial protein synthesis.* Mol Cell, 2006. **23**(2): p. 183-93.

14. Antoun, A., M.Y. Pavlov, M. Lovmar, and M. Ehrenberg, *How initiation factors tune the rate of initiation of protein synthesis in bacteria.* EMBO J, 2006. **25**(11): p. 2539-50.

15. Antoun, A., M.Y. Pavlov, T. Tenson, and M. Ehrenberg, *Ribosome formation from subunits studied by stopped-flow and Rayleigh light scattering.* Biol Proced Online, 2004. **6**(1): p. 35-54.

16. Milon, P., A.L. Konevega, C.O. Gualerzi, and M.V. Rodnina, *Kinetic checkpoint at a late step in translation initiation.* Mol Cell, 2008. **30**(6): p. 712-20.

17. Grigoriadou, C., S. Marzi, D. Pan, C.O. Gualerzi, and B.S. Cooperman, *The translational fidelity function of IF3 during transition from the 30 S initiation complex to the 70 S initiation complex.* J Mol Biol, 2007. **373**(3): p. 551-61.

18. Kycia, J.H., V. Biou, F. Shu, S.E. Gerchman, V. Graziano, and V. Ramakrishnan, *Prokaryotic translation initiation factor IF3 is an elongated protein consisting of two crystallizable domains.* Biochemistry, 1995. **34**(18): p. 6183-7.

19. Fortier, P.L., J.M. Schmitter, C. Garcia, and F. Dardel, *The N-terminal half of initiation factor IF3 is folded as a stable independent domain.* Biochimie, 1994. **76**(5): p. 376-83.

20. Moreau, M., E. de Cock, P.L. Fortier, C. Garcia, C. Albaret, S. Blanquet, J.Y. Lallemand, and F. Dardel, *Heteronuclear NMR studies of E. coli translation initiation factor IF3. Evidence that the inter-domain region is disordered in solution.* J Mol Biol, 1997. **266**(1): p. 15-22.

21. de Cock, E., M. Springer, and F. Dardel, *The interdomain linker of Escherichia coli initiation factor IF3: a possible trigger of translation initiation specificity.* Mol Microbiol, 1999. **32**(1): p. 193-202.

22. Fabbretti, A., C.L. Pon, S.P. Hennelly, W.E. Hill, J.S. Lodmell, and C.O. Gualerzi, *The real-time path of translation factor IF3 onto and off the ribosome.* Mol Cell, 2007. **25**(2): p. 285-96.

23. Sette, M., R. Spurio, P. van Tilborg, C.O. Gualerzi, and R. Boelens, *Identification of the ribosome binding sites of translation initiation factor IF3 by multidimensional heteronuclear NMR spectroscopy.* RNA, 1999. **5**(1): p. 82-92.

24. Petrelli, D., A. LaTeana, C. Garofalo, R. Spurio, C.L. Pon, and C.O. Gualerzi, *Translation initiation factor IF3: two domains, five functions, one mechanism?* EMBO J, 2001. **20**(16): p. 4560-9.

25. Shapkina, T.G., M.A. Dolan, P. Babin, and P. Wollenzien, *Initiation factor 3-induced structural changes in the 30 S ribosomal subunit and in complexes containing tRNA(f)(Met) and mRNA.* J Mol Biol, 2000. **299**(3): p. 615-28.

26. Muralikrishna, P. and E. Wickstrom, *Escherichia coli initiation factor 3 protein binding to 30S ribosomal subunits alters the accessibility of nucleotides within the conserved central region of 16S rRNA.* Biochemistry, 1989. **28**(19): p. 7505-10.

27. McCutcheon, J.P., R.K. Agrawal, S.M. Philips, R.A. Grassucci, S.E. Gerchman, W.M. Clemons, Jr., V. Ramakrishnan, and J. Frank, *Location of translational initiation factor IF3 on the small ribosomal subunit.* Proc Natl Acad Sci U S A, 1999. **96**(8): p. 4301-6.

28. Dallas, A. and H.F. Noller, *Interaction of translation initiation factor 3 with the 30S ribosomal subunit.* Mol Cell, 2001. **8**(4): p. 855-64.

29. Moazed, D., R.R. Samaha, C. Gualerzi, and H.F. Noller, *Specific protection of 16 S rRNA by translational initiation factors.* J Mol Biol, 1995. **248**(2): p. 207-10.

30. Passmore, L.A., T.M. Schmeing, D. Maag, D.J. Applefield, M.G. Acker, M.A. Algire, J.R. Lorsch, and V. Ramakrishnan, *The eukaryotic translation initiation factors eIF1 and eIF1A induce an open conformation of the 40S ribosome.* Mol Cell, 2007. **26**(1): p. 41-50.

31. Lomakin, I.B., N.E. Shirokikh, M.M. Yusupov, C.U. Hellen, and T.V. Pestova, *The fidelity of translation initiation: reciprocal activities of eIF1, IF3 and YciH.* EMBO J, 2006. **25**(1): p. 196-210.

32. Rabl, J., M. Leibundgut, S.F. Ataide, A. Haag, and N. Ban, *Crystal structure of the eukaryotic 40S ribosomal subunit in complex with initiation factor 1.* Science, 2011. **331**(6018): p. 730-6.

33. Hasenohrl, D., D. Benelli, A. Barbazza, P. Londei, and U. Blasi, *Sulfolobus solfataricus translation initiation factor 1 stimulates translation initiation complex formation.* RNA, 2006. **12**(4): p. 674-82.

34. Lomakin, I.B., V.G. Kolupaeva, A. Marintchev, G. Wagner, and T.V. Pestova, *Position of eukaryotic initiation factor eIF1 on the 40S ribosomal subunit determined by directed hydroxyl radical probing.* Genes Dev, 2003. **17**(22): p. 2786-97.

35. Pioletti, M., F. Schlunzen, J. Harms, R. Zarivach, M. Gluhmann, H. Avila, A. Bashan, H. Bartels, T. Auerbach, C. Jacobi, T. Hartsch, A. Yonath, and F. Franceschi, *Crystal structures of complexes of the small ribosomal subunit with tetracycline, edeine and IF3.* EMBO J, 2001. **20**(8): p. 1829-39.

36. Ehresmann, C., H. Moine, M. Mougel, J. Dondon, M. Grunberg-Manago, J.P. Ebel, and B. Ehresmann, *Cross-linking of initiation factor IF3 to Escherichia coli 30S ribosomal subunit by trans-diamminedichloroplatinum(II): characterization of two cross-linking sites in 16S rRNA; a possible way of functioning for IF3.* Nucleic Acids Res, 1986. **14**(12): p. 4803-21.

37. Liiv, A. and M. O'Connor, *Mutations in the intersubunit bridge regions of 23 S rRNA.* J Biol Chem, 2006. **281**(40): p. 29850-62.

38. Yusupov, M.M., G.Z. Yusupova, A. Baucom, K. Lieberman, T.N. Earnest, J.H. Cate, and H.F. Noller, *Crystal structure of the ribosome at 5.5 A resolution.* Science, 2001. **292**(5518): p. 883-96.

39. Jelenc, P.C. and C.G. Kurland, *Nucleoside triphosphate regeneration decreases the frequency of translation errors.* Proc Natl Acad Sci U S A, 1979. **76**(7): p. 3174-8.

40. Pavlov, M.Y. and M. Ehrenberg, *Rate of translation of natural mRNAs in an optimized in vitro system.* Arch Biochem Biophys, 1996. **328**(1): p. 9-16.

41. Wagner, E.G., P.C. Jelenc, M. Ehrenberg, and C.G. Kurland, *Rate of elongation of polyphenylalanine in vitro.* Eur J Biochem, 1982. **122**(1): p. 193-7.

42. Wang, J., *Regulation of IF2 Binding Kinetics and 30S IC Conformational Dynamics during Translation Initiation*, 2010, Columbia University.

43. Fei, J., J. Wang, S.H. Sternberg, D.D. MacDougall, M.M. Elvekrog, D.K. Pulukkunat, M.T. Englander, and R.L. Gonzalez, Jr., *A highly purified, fluorescently labeled in vitro translation system for single-molecule studies of protein synthesis.* Methods Enzymol, 2010. **472**: p. 221-59.

44. Blanchard, S.C., H.D. Kim, R.L. Gonzalez, Jr., J.D. Puglisi, and S. Chu, *tRNA dynamics on the ribosome during translation.* Proc Natl Acad Sci U S A, 2004. **101**(35): p. 12893-8.

45. Fei, J., P. Kosuri, D.D. MacDougall, and R.L. Gonzalez, Jr., *Coupling of ribosomal L1 stalk and tRNA dynamics during translation elongation.* Mol Cell, 2008. **30**(3): p. 348-59.

46. Fei, J., J.E. Bronson, J.M. Hofman, R.L. Srinivas, C.H. Wiggins, and R.L. Gonzalez, Jr., *Allosteric collaboration between elongation factor G and the ribosomal L1 stalk directs tRNA movements during translation.* Proc Natl Acad Sci U S A, 2009. **106**(37): p. 15702-7.

47. Sternberg, S.H., J. Fei, N. Prywes, K.A. McGrath, and R.L. Gonzalez, Jr., *Translation factors direct intrinsic ribosome dynamics during translation termination and ribosome recycling.* Nat Struct Mol Biol, 2009. **16**(8): p. 861-8.

48. Bronson, J.E., J. Fei, J.M. Hofman, R.L. Gonzalez, Jr., and C.H. Wiggins, *Learning rates and states from biophysical time series: a Bayesian approach to model selection and single-molecule FRET data.* Biophys J, 2009. **97**(12): p. 3196-205.

49. Altschul, S.F., W. Gish, W. Miller, E.W. Myers, and D.J. Lipman, *Basic local alignment search tool.* J Mol Biol, 1990. **215**(3): p. 403-10.

50. Bastiaens, P.I. and T.M. Jovin, *Microspectroscopic imaging tracks the intracellular processing of a signal transduction protein: fluorescent-labeled protein kinase C beta I.* Proc Natl Acad Sci U S A, 1996. **93**(16): p. 8407-12.

51. Hohng, S., C. Joo, and T. Ha, *Single-molecule three-color FRET.* Biophys J, 2004. **87**(2): p. 1328-37.

52. Koradi, R., M. Billeter, and K. Wuthrich, *MOLMOL: a program for display and analysis of macromolecular structures.* J Mol Graph, 1996. **14**(1): p. 51-5, 29-32.

53. Milon, P., A.L. Konevega, F. Peske, A. Fabbretti, C.O. Gualerzi, and M.V. Rodnina, *Transient kinetics, fluorescence, and FRET in studies of initiation of translation in bacteria.* Methods Enzymol, 2007. **430**: p. 1-30.

54. Pon, C., S. Cannistraro, A. Giovane, and C. Gualerzi, *Structure-function relationship in Escherichia coli initiation factors. Environment of the Cys residue and evidence for a hydrophobic region in initiation factor IF3 by fluorescence and ESR spectroscopy.* Arch Biochem Biophys, 1982. **217**(1): p. 47-57.

55. Hua, Y. and D.P. Raleigh, *On the global architecture of initiation factor IF3: a comparative study of the linker regions from the Escherichia coli protein and the Bacillus stearothermophilus protein.* J Mol Biol, 1998. **278**(4): p. 871-8.

56. Gell, C., D. Brockwell, and A. Smith, *Handbook of Single Molecule Fluorescence Spectroscopy*2006, Oxford: Oxford University Press.

57. Qin, D., N.M. Abdi, and K. Fredrick, *Characterization of 16S rRNA mutations that decrease the fidelity of translation initiation.* RNA, 2007. **13**(12): p. 2348-55.

58. Qin, D. and K. Fredrick, *Control of translation initiation involves a factor-induced rearrangement of helix 44 of 16S ribosomal RNA.* Mol Microbiol, 2009. **71**(5): p. 1239-49.

59. Hartz, D., D.S. McPheeters, R. Traut, and L. Gold, *Extension inhibition analysis of translation initiation complexes.* Methods Enzymol, 1988. **164**: p. 419-25.

60. Adcock, S.A. and J.A. McCammon, *Molecular dynamics: survey of methods for simulating the activity of proteins.* Chem Rev, 2006. **106**(5): p. 1589-615.

61. Berk, V., W. Zhang, R.D. Pai, and J.H. Cate, *Structural basis for mRNA and tRNA positioning on the ribosome.* Proc Natl Acad Sci U S A, 2006. **103**(43): p. 15830-4.

62. Garcia, C., P.L. Fortier, S. Blanquet, J.Y. Lallemand, and F. Dardel, *Solution structure of the ribosome-binding domain of E. coli translation initiation factor IF3. Homology with the U1A protein of the eukaryotic spliceosome.* J Mol Biol, 1995. **254**(2): p. 247-59.

63. Howe, J.G. and J.W. Hershey, *Initiation factor and ribosome levels are coordinately controlled in Escherichia coli growing at different rates.* J Biol Chem, 1983. **258**(3): p. 1954-9.

64. Zucker, F.H. and J.W. Hershey, *Binding of Escherichia coli protein synthesis initiation factor IF1 to 30S ribosomal subunits measured by fluorescence polarization.* Biochemistry, 1986. **25**(12): p. 3682-90.

65. Boileau, G., P. Butler, J.W. Hershey, and R.R. Traut, *Direct cross-links between initiation factors 1, 2, and 3 and ribosomal proteins promoted by 2-iminothiolane.* Biochemistry, 1983. **22**(13): p. 3162-70.

66. Allen, G.S., A. Zavialov, R. Gursky, M. Ehrenberg, and J. Frank, *The cryo-EM structure of a translation initiation complex from Escherichia coli.* Cell, 2005. **121**(5): p. 703-12.

67. Carter, A.P., W.M. Clemons, Jr., D.E. Brodersen, R.J. Morgan-Warren, T. Hartsch, B.T. Wimberly, and V. Ramakrishnan, *Crystal structure of an initiation factor bound to the 30S ribosomal subunit.* Science, 2001. **291**(5503): p. 498-501.

68. Heimark, R.L., L. Kahan, K. Johnston, J.W. Hershey, and R.R. Traut, *Cross-linking of initiation factor IF3 to proteins of the Escherichia coli 30 S ribosomal subunit.* J Mol Biol, 1976. **105**(2): p. 219-30.

69. Hawley, D.A., L.I. Slobin, and A.J. Wahba, *The mechanism of action of initiation factor 3 in protein synthesis. II. Association of the 30S ribosomal protein S12 with IF-3.* Biochem Biophys Res Commun, 1974. **61**(2): p. 544-50.

70. Gualerzi, C. and C.L. Pon, *Nature of the ribosomal binding site for initiation factor 3 (IF-3).* Biochem Biophys Res Commun, 1973. **52**(3): p. 792-9.

71. MacKeen, L.A., L. Kahan, A.J. Wahba, and I. Schwartz, *Photochemical cross-linking of initiation factor-3 to Escherichia coli 30 S ribosomal subunits.* J Biol Chem, 1980. **255**(21): p. 10526-31.

72. Cooperman, B.S., A. Expert-Bezancon, L. Kahan, J. Dondon, and M. Grunberg-Manago, *IF-3 crosslinking to Escherichia coli ribosomal 30 S subunits by three different light-dependent procedures: identification of 30 S proteins crosslinked to IF-3--utilization of a new two-stage crosslinking reagent, p-nitrobenzylmaleimide.* Arch Biochem Biophys, 1981. **208**(2): p. 554-62.

73. Sacerdot, C., G. Fayat, P. Dessen, M. Springer, J.A. Plumbridge, M. Grunberg-Manago, and S. Blanquet, *Sequence of a 1.26-kb DNA fragment containing the structural gene for*

E.coli initiation factor IF3: presence of an AUU initiator codon. EMBO J, 1982. **1**(3): p. 311-5.

74. Gold, L., G. Stormo, and R. Saunders, *Escherichia coli translational initiation factor IF3: a unique case of translational regulation.* Proc Natl Acad Sci U S A, 1984. **81**(22): p. 7061-5.

75. Lancaster, L. and H.F. Noller, *Involvement of 16S rRNA Nucleotides G1338 and A1339 in discrimination of initiator tRNA.* Molecular Cell, 2005. **20**(4): p. 623-632.

76. Selmer, M., C.M. Dunham, F.V.t. Murphy, A. Weixlbaumer, S. Petry, A.C. Kelley, J.R. Weir, and V. Ramakrishnan, *Structure of the 70S ribosome complexed with mRNA and tRNA.* Science, 2006. **313**(5795): p. 1935-42.

77. Larkin, M.A., G. Blackshields, N.P. Brown, R. Chenna, P.A. McGettigan, H. McWilliam, F. Valentin, I.M. Wallace, A. Wilm, R. Lopez, J.D. Thompson, T.J. Gibson, and D.G. Higgins, *Clustal W and Clustal X version 2.0.* Bioinformatics, 2007. **23**(21): p. 2947-8.

78. Goujon, M., H. McWilliam, W. Li, F. Valentin, S. Squizzato, J. Paern, and R. Lopez, *A new bioinformatics analysis tools framework at EMBL-EBI.* Nucleic Acids Res, 2010. **38**(Web Server issue): p. W695-9.

79. DeLano, W., *The PyMOL Molecular Graphics System, Version 1.3, Schrödinger, LLC.* 2008.

80. Effraim, P.R., J. Wang, M.T. Englander, J. Avins, T.S. Leyh, R.L. Gonzalez, Jr., and V.W. Cornish, *Natural amino acids do not require their native tRNAs for efficient selection by the ribosome.* Nat Chem Biol, 2009. **5**(12): p. 947-53.

Chapter 3
Initiation factor 3-mediated 50S subunit joining

3.1 Introduction

50S subunit joining to a completely assembled 30S IC is a pivotal step toward transit into translation elongation and rapid mRNA decoding (Figure 3.1). For this reason, subunit joining to 30S ICs is highly regulated by the IFs. IF2(GTP) catalyzes rapid 50S subunit joining to initiator tRNA-bound 30S subunits [1, 2], while IF1 affects subunit association by inducing large scale conformational changes in the 30S subunit [3-5]. IF3 is especially critical in regulating 50S subunit joining by preventing the efficient formation of 70S ribosomes lacking initiator tRNA (Figure 3.2) [2, 6] or containing either a non-initiator tRNA [7] (see Table 3.1) or a non-canonical AUU start codon [4, 8, 9]. In the absence of IF3, this substrate discrimination, especially for an AUG versus AUU start codon, is largely lost and 50S subunits associate rapidly with 30S ICs formed on AUU- and AUG-containing mRNAs [8]. Association of the 50S subunit with the 30S IC involves the sequential, multi-step formation of twelve intersubunit bridges (see Table 1.1) [10]. Recalling that the binding site of IF3 on the 30S subunit coincides with the binding sites for a number of intersubunit bridges (Figure 1.12) [11], it follows that the presence of IF3 on a 30S IC may either prevent 50S subunit joining entirely or else hinder the formation of a subset of these bridges, thus preventing the ribosome's conversion from an unstable to a stable 70S IC [8, 12]; these lines of reasoning form the foundation for two opposing models for how IF3 regulates subunit joining.

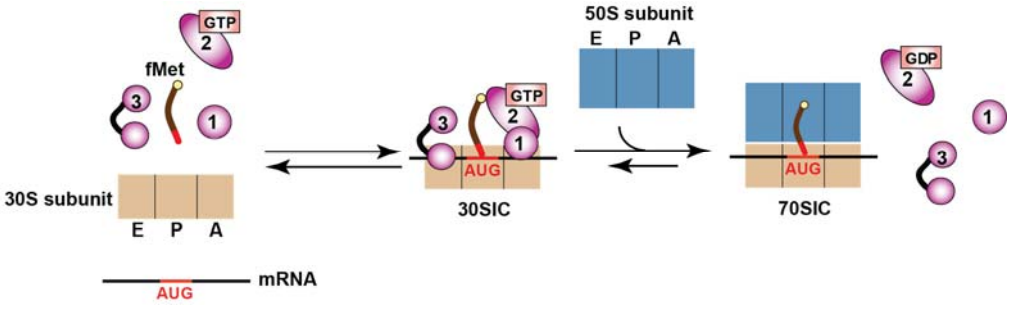

Figure 3.1 A minimal mechanistic model for translation initiation. The two main steps in translation initiation are the formation of the 30S IC and 70S IC. These steps are regulated by the initiation factors.

Table 3.1 Effects of initiation factors, and the presence and identity of tRNA on the first-order rate constant, k_c, of 50S joining to 30S-mRNA initiation complexes. From Ref. [7]; measured by light scattering, as depicted in Figure 3.2. The shaded cells highlight the effect of omitting or including IF3 in otherwise completely assembled 30S ICs.

IF1	IF2	IF3	k_c (s^{-1})		
			—	fMet-tRNA^{fMet}	Phe-tRNA^{Phe}
+	+	+	0	2.9	0.0071
+	-	+	0	0.0098	~ 0
-	+	+	0.020	8.7	0.015
-	-	+	0.020	0.038	0.0032
+	+	-	1.4	42	4.0
+	-	-	0.43	0.24	0.16
-	+	-	0.63	36	0.79
-	-	-	0.56	0.30	0.19

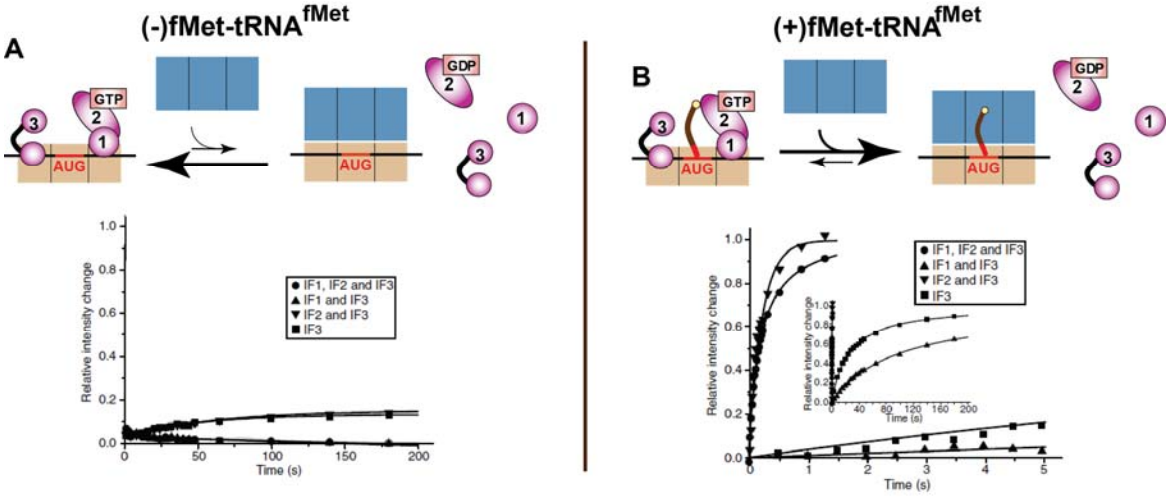

Figure 3.2 The kinetics of 50S subunit joining to 30S ICs formed in the absence or presence of initiator tRNA and initiation factors. Data from Antoun et al. [2]. **A.** Cartoon of the 50S subunit joining reaction in the absence of initiator tRNA. The raw light scattering data is shown in black. 30S ICs were formed in the presence of the IFs indicated in the legend. The calculated rate constants, k_c, for 50S subunit joining are found in Table 3.1. **B.** Cartoon of the 50S subunit joining reaction in the presence of initiator tRNA, with raw light scattering data shown below.

The two opposing mechanistic models for how IF3 regulates 50S subunit joining to 30S ICs are schematized in Figure 3.3. In the first model, IF3 must spontaneously dissociate from a correctly assembled 30S IC before 50S subunit joining [2, 7]. This model is supported by data collected using stopped-flow techniques with detection of Rayleigh-scattered light to estimate the rate constant for 50S subunit joining in the presence and absence of IF3. In the absence of IF3, it was shown, through a 50S subunit concentration series, that 70S ribosome formation obeyed

second-order kinetics. In the presence of IF3, however, the rate of 50S subunit joining, k_c, became a hyperbolic function with a clear plateau. When IF3 was titrated into fixed concentrations of mRNA-programmed 30S ICs, subunit association was inhibited with a linear dependence of $1/k_c$ on 1/[IF3]. These data therefore imply that IF3 is a competitive inhibitor of 50S subunit joining, indicating that the presence of IF3 and a 50S subunit on a 30S IC is mutually exclusive. In this model, the rate of 50S subunit joining would be limited by the rate of IF3 dissociation.

In the second model, IF3 remains bound to a correctly assembled 30S IC upon 50S subunit joining and dissociates either during or after the 50S subunit has joined [4, 8, 12, 13]. This model was developed based on results from rapid kinetics, ensemble FRET, and time-resolved hydroxyl radical probing experiments indicating that IF3 is present, at least transiently, on 50S subunit-associated 30S ICs [4, 8, 12, 13]. Attempts to explain the discrepancies between these two models include suggestions by Gualerzi and co-workers that the Ehrenberg study was performed with an N-terminally degraded IF3 [12]. Others posited that differing buffer compositions and the use of mRNAs differing in SD strength and spacer length have a significant effect on the rate of IF3 dissociation [4], which may explain differences in the timing of IF3 dissociation with respect to 50S subunit joining. Also important to keep in mind are the inherent limitations of a given biophysical technique. For example, the detection of light scattering to measure subunit association may be limited to detection of just stable 70S ribosome formation, giving limited or no information on transient subunit association, which may be a necessary intermediate on the pathway of, or an impetus for, IF3 dissociation (D. MacDougall, personal communication). Thus, the mutual exclusivity of these two models may not be as pronounced as their authors suggest.

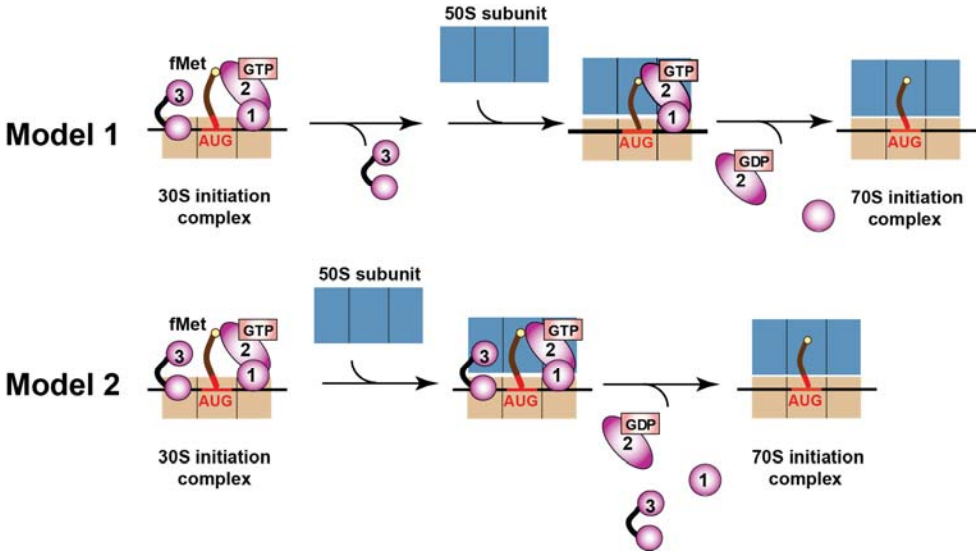

Figure 3.3 Two opposing models for the timing of IF3's dissociation during translation initiation. In Model 1, IF3 spontaneously dissociates from a completely assembled 30S IC before 50S subunit joining occurs. In Model 2, IF3 remains bound to the 30S IC during 50S subunit association and dissociates at some point during or after 50S subunit joining.

In Chapter Two, it was shown that 30S IC-bound IF3 is conformationally dynamic and that its conformational dynamics are linked to IF3's role in substrate selection (Figure 3.4). These results already support the latter of these two mechanistic models due to the long-lived presence of IF3 on a completely assembled 30S IC (see Figure 3.4). IF3 remains stably bound on surface-immobilized 30S subunits over the course of 20-30 minutes of smFRET data acquisition. Contrasting with this, if IF3 spontaneously dissociated from completely assembled 30S ICs on a reasonable timescale, as described by Model #1, smFRET data acquisition would not be possible due to the dissociation of the fluorescent observable. The new evidence for IF3 being conformationally dynamic on 30S ICs, however, suggests additional layers of complexity beyond the relative timing of IF3's dissociation from the ribosome to explain IF3's role in regulating subunit joining.

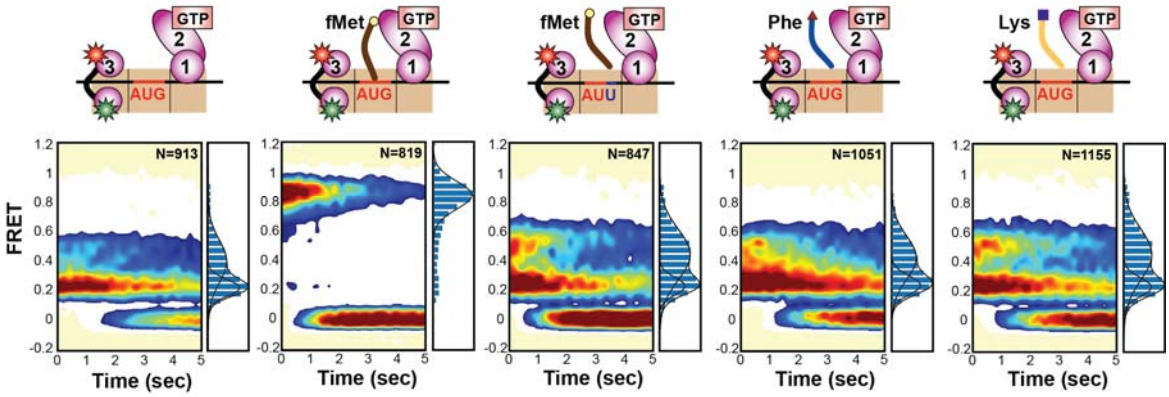

Figure 3.4 An IF3 interdomain reconfiguration signals proper initiator tRNA and start codon selection during translation initiation. Time-evolution of population FRET histograms for 30SIC^{-tRNA}, 30SIC^{fMet}, 30SIC^{fMet,AUU}, 30SIC^{Phe}, and 30SIC$^{Lys.}$ Normalized population FRET histograms, containing the first five frames of data, are shown to the right of each contour plot. The contents of each 30S IC are depicted in the cartoon above each FRET histogram. The non-cognate codon-anticodon interaction in the three plots on the right is depicted by a loose interaction between the tRNA and the 30S subunit P site.

As presented in Chapter Two, and depicted in Figure 3.4, IF3 undergoes an interdomain reconfiguration to a single, unique interdomain configuration only upon proper initiator tRNA and start codon selection. Ensemble biochemical studies of 50S subunit joining (Table 3.1, Figure 3.2, and Refs. [2, 4, 6-8]) have shown that, in the presence of IF3, subunit joining occurs most rapidly with completely assembled 30S ICs, containing an AUG start codon and fMet-tRNA^{fMet}. Perturbation of the identity of the tRNA or start codon can slow down subunit joining by up to 400-fold [7], while complete omission of tRNA slows subunit joining by more than 1200-fold [7] (Table 3.1). Correlating these previous kinetic results with this new structural finding on the interdomain configuration of IF3 on comparable correctly versus incorrectly assembled 30S ICs leads to a new mechanistic model which incorporates the interdomain configuration of IF3 into its ability to regulate 50S subunit joining (Figure 3.5). Here, it is both the presence and interdomain configuration of IF3 that determines the efficiency of subunit joining. IF3's switch to a single compact interdomain configuration on completely and correctly assembled 30S ICs may be more conducive to 50S subunit joining, the formation of intersubunit bridges, and/or IF3's dissociation rate from the 70S ribosome, whereas IF3's other 30S IC-bound interdomain configurations may

occlude a subset of these intersubunit bridges, thus preventing rapid and stable subunit joining, and perhaps also slowing down IF3's dissociation from 70S ribosomes.

It is unlikely that IF3's interdomain configuration is the sole determinant of rapid 50S subunit joining, however. This is evident by the stable sampling of $IF3_{0.84}$, the presumably "subunit joining ready" IF3 interdomain configuration, in the presence of just IF2 and the absence of IF1 or fMet-tRNA^{fMet} (see Figure 2.9). These 30S ICs ($30SIC^{-tRNA}_{-IF1}$) do not undergo rapid subunit joining (see Table 3.1 row 3), suggesting that IF3's interdomain reconfiguration may be necessary, but not sufficient, for subunit joining. Instead, the interdomain configuration of 30S IC-bound IF3 is likely to be one of a number of conditions, including the presence of IF2(GTP) and fMet-tRNA^{fMet} on the 30S IC, which must be met for rapid, stable 70S IC formation.

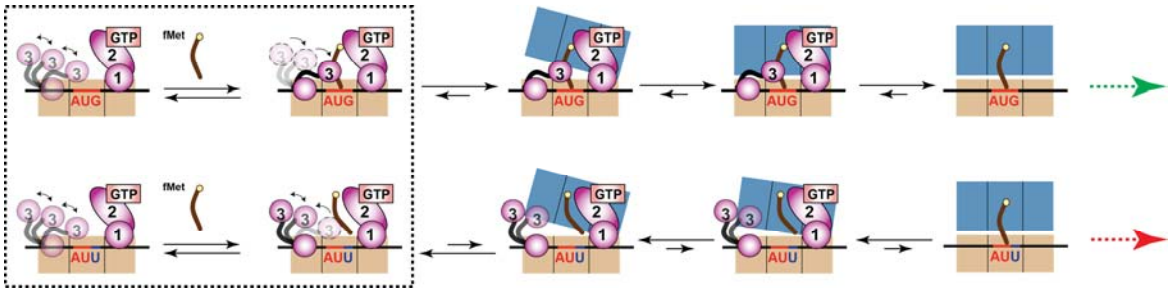

Figure 3.5 A mechanistic model for the role of 30S IC-bound IF3's interdomain configuration in regulating the fidelity of 70S IC formation. 30S IC-bound IF3 samples at least three distinct and interconvertible interdomain configurations on the various 30S ICs it encounters along the initiation pathway, as revealed in Chapter Two. IF3 converges on a single interdomain configuration within completely and correctly assembled 30S ICs containing the full array of initiation components: IF1, IF2, an AUG start codon, and fMet-tRNA^{fMet}. On incompletely or incorrectly assembled 30S ICs, containing, for example, a non-canonical start codon or non-initiator tRNA, IF3 predominantly samples two other interdomain configurations. These results are found in Chapter Two and cartooned here in the dashed box. At the right is a possible model for how IF3's interdomain configuration on the 30S IC regulates subunit joining. It is proposed that only in one specific interdomain configuration does IF3 allow rapid subunit joining, while in its other interdomain configurations, IF3 may occlude intersubunit bridge formation, thus slowly down or preventing 50S subunit joining. IF3's conformational rearrangements can result from either (i) dynamics of the 30S subunit, (ii) multiple IF3 binding sites, and (iii) a combination of these scenarios (see Figure 2.8).

3.2 Preliminary results and discussion

Although this new mechanistic model remains to be rigorously tested, I have taken steps toward developing an smFRET signal development between IF3 and the 50S subunit – a signal that can indicate the timing of the 50S subunit's joining with an IF3-bound 30S IC. The interdomain configuration of 30S IC-bound IF3 during this subunit joining step can be measured

with IF3(Cy3-Cy5) (see Chapter Two), while observation of both IF3's interdomain configuration, as well as the presence of a joined 50S subunit on the 30S IC, will necessitate a third fluorophore on the 50S subunit (see Figure 3.7).

3.2.1 Development of a smFRET signal between IF3 and the 50S subunit

50S subunits reconstituted with mutageneized and labeled ribosomal protein L9(Q18C)-Cy5 were prepared by Dr. Jingyi Fei during the development of a smFRET signal between L9 and ribosomal protein L1 [14]. Measurements of the distance between the Cys18 residue of L9 and the 16S rRNA nucleotides protected by IF3 [15] enabled rough estimation of the separation between L9 and 30S-bound IF3. These distances, measured in PyMOL [16] with PDB files 2J00 and 2J01, are: L9(Q18C) to G685 of rRNA helix 23 = 49 Å, L9(Q18C) to G703 of helix 23 = 52 Å, and L9(Q18C) to U793 of helix 24 = 89 Å. According to the location of IF3's NTD and CTD modeled using constraints from directed hydroxyl radical probing [15], L9 is closer to IF3's NTD than its CTD.

3.2.2 IF3 remains bound to the 30S subunit following 50S subunit joining

A proof-of-principle experiment has been performed to verify that IF3-NTD is within FRET distance of L9 (see Figure 3.6). Cy5-labeled 50S subunits were stopped-flow delivered in realtime to surface immobilized 30S ICs that were assembled with a complete set of initiation components (i.e. $30SIC^{fMet}$). The arrival of 50S subunits to IF3-bound 30S ICs was visualized by the onset of FRET. Based on ~25 smFRET trajectories, the measured FRET efficiency is ~0.45, a FRET efficiency corresponding to a distance of ~57 Å, assuming an R_0 value of 55 Å, and in line with the distances measured using a 70S ribosome crystal structure (see section 3.2.1). It is clear from these acquired smFRET trajectories that IF3 remains stably bound to 70S ribosomes for at least several seconds, possibly limited by photobleaching. These data support the second model described in section 3.1 (Figure 3.3) and further suggest that IF3 does not immediately dissociate upon 50S subunit association, but can remain relatively stably bound to a 70S ribosome. These results also support a cryo-EM structure of a 70S IC containing the initiator tRNA, IF1,

IF2(GDPNP), and IF3 [17], as well as time-resolved chemical probing investigations of the timing of IF3's dissociation with respect to 50S subunit arrival [12].

Numerous questions arise from these results, however, such as whether all intersubunit bridges are formed on these IF3-bound 70S ribosomes (i.e. have the 70S ribosomes undergone a transition from an "unstable" to a "stable" intersubunit configuration [8, 12]), as well as IF3's precise lifetime on the 70S ribosome, and what stimulus is required for its eventual release, assuming it is not simply the presence of a successfully joined 50S subunit. Another member of the Gonzalez research group, Mr. Daniel MacDougall, is investigating the role of IF3 in stable 70S ribosome formation with an smFRET signal between IF2 and the 50S subunit ribosomal protein L11, and his preliminary results indicate that IF3 significantly affects the frequency of stable subunit joining, as well as the stability of the interaction between the joined subunits. The L9-IF3 smFRET signal described here can be used to answer questions about IF3-NTD's lifetime on the ribosome, however it is limited by the lack of a physical observable that reports on the relative interdomain distance of IF3's NTD and CTD, as well as the location of IF3's CTD upon 50S subunit binding – data that may demonstrate a mechanistically important conformational rearrangement of IF3.

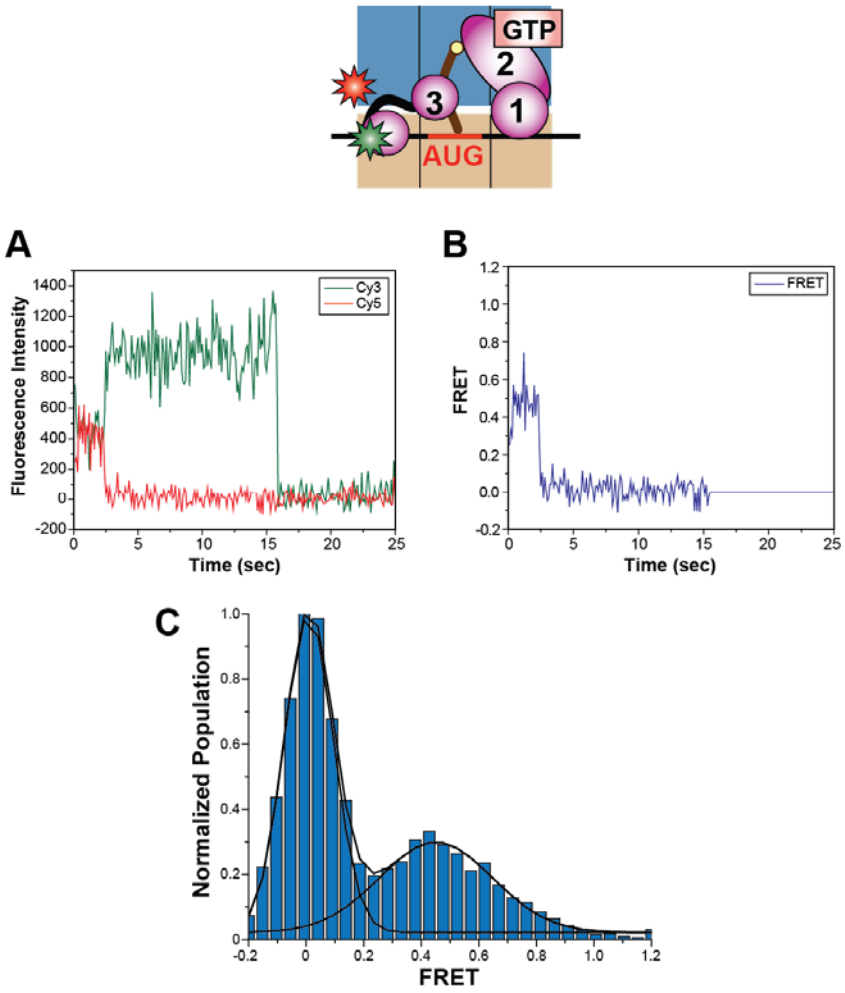

Figure 3.6 smFRET between IF3 and the 50S subunit. 50S subunits were delivered to surface-immobilized 30S ICs and FRET was observed between IF3(C65S-S38C)-Cy3 and ribosomal protein L9(Q18C)-Cy5. The cartoon represents a 70S ribosome carrying IF1, IF2(GTP), IF3(Cy3), initiator tRNA, and a Cy5-labeled 50S subunit. **A**. A representative fluorescence versus time trajectory. Cy3 fluorescence emission intensity in green and Cy5 fluorescence emission in red. **B**. The corresponding FRET versus time trajectory, calculated as FRET = $I_{Cy5}/(I_{Cy3}+I_{Cy5})$. **C**. The normalized population FRET histogram from all FRET trajectories (N=25) plotted with all 1200 frames of data. The peak at zero FRET is due to Cy3 and Cy5 photobleaching.

3.2.3 Future directions: the interdomain configuration of ribosome-bound IF3 during 50S subunit joining

Future experiments can be done to test the interdomain configuration of IF3 during 50S subunit joining. In order to probe both the relative interdomain distance between IF3's NTD and CTD to get insight into IF3's interdomain configuration, as well as identify the timing of 50S subunit joining to an IF3-bound 30S subunit, experiments can be designed as depicted in Figure

3.7. Two separate observables will be needed for these experiments. First, 30S IC-bound IF3(Cy3-Cy5) will enable measurements on the interdomain configuration of 30S- and 70S-bound IF3. Secondly, fluorescently labeled 50S subunits will be needed to indicate the arrival of the 50S subunit. This arrival event can be determined from fluorescence co-localization experiments in which one can monitor the appearance and disappearance of fluorophores within a diffraction-limited region of the TIRFM field-of-view (see, for example, Ref. [18]). Co-localization experiments would require careful controls to ensure that non-specific binding of the fluorophore-labeled components to the TIRF-illuminated surface of the microfluidic flowcell is minimal, as well as careful optimization of the TIRFM optics and data analysis procedures to minimize user bias when identifying and co-localizing regions of fluorescence intensity. Alternatively, the arrival of the 50S subunit can be determined by FRET between the 50S subunit and IF3(Cy3-Cy5). In this case, a fluorophore needs to be chosen that is both photostable and has spectral overlap with either Cy3 or Cy5. Possible fluorophores include Cy7 or Atto488. The specific labeling position could be L9(Q18C) of the 50S subunit due to this residue's proximity to IF3's NTD, or if the fluorescence co-localization approach is taken, the 50S subunits could be non-specifically labeled on peripheral lysine residues following the procedure developed and described by Dr. Jiangning Wang [19]. Though technically challenging, these experiments have the potential to reveal a wealth of information on the interdomain configuration of IF3 during 50S subunit joining, as well as report on any 70S IC-bound IF3 conformational dynamics and the timing of IF3's dissociation. Importantly, these experiments will enable direct testing of the new mechanistic model presented in section 3.1 (Figure 3.5), namely that IF3's ability to regulate 50S subunit joining stems from the interdomain configuration of the ribosome-bound IF3.

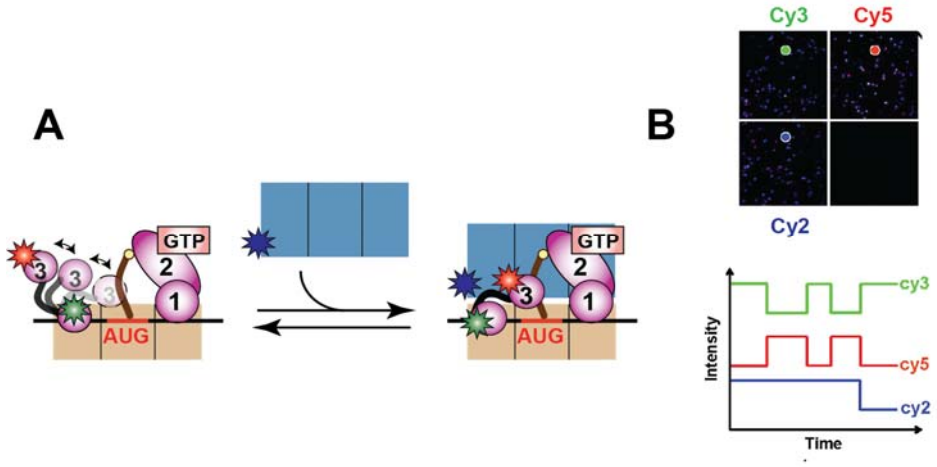

Figure 3.7 A possible experiment to monitor the interdomain configuration of IF3 during 50S subunit joining. The relative distance between IF3's NTD and CTD can be measured with 30S subunit-bound IF3(Cy3-Cy5), and the joining of a 50S subunit to an IF3-bound 30S subunit can be monitored either via FRET between IF3-NTD and the L9 protein of the 50S subunit, or via fluorescence co-localization between an IF3-bound fluorophore (either Cy3 or Cy5) and 50S-bound fluorophore (e.g. Atto488 or Cy2).

3.3 References

1. Antoun, A., M.Y. Pavlov, K. Andersson, T. Tenson, and M. Ehrenberg, *The roles of initiation factor 2 and guanosine triphosphate in initiation of protein synthesis.* EMBO J, 2003. **22**(20): p. 5593-601.

2. Antoun, A., M.Y. Pavlov, M. Lovmar, and M. Ehrenberg, *How initiation factors tune the rate of initiation of protein synthesis in bacteria.* EMBO J, 2006. **25**(11): p. 2539-50.

3. Carter, A.P., W.M. Clemons, Jr., D.E. Brodersen, R.J. Morgan-Warren, T. Hartsch, B.T. Wimberly, and V. Ramakrishnan, *Crystal structure of an initiation factor bound to the 30S ribosomal subunit.* Science, 2001. **291**(5503): p. 498-501.

4. Milon, P., A.L. Konevega, C.O. Gualerzi, and M.V. Rodnina, *Kinetic checkpoint at a late step in translation initiation.* Mol Cell, 2008. **30**(6): p. 712-20.

5. Belotserkovsky, J.M., E.R. Dabbs, and L.A. Isaksson, *Mutations in 16S ribosomal RNA that suppress a cold sensitive initiation factor 1 affect ribosomal subunit association.* FEBS J, 2011.

6. Antoun, A., M.Y. Pavlov, T. Tenson, and M. Ehrenberg, *Ribosome formation from subunits studied by stopped-flow and Rayleigh light scattering.* Biol. Proced. Online, 2004. **6**(1): p. 35-45.

7. Antoun, A., M.Y. Pavlov, M. Lovmar, and M. Ehrenberg, *How initiation factors maximize the accuracy of tRNA selection in initiation of bacterial protein synthesis.* Mol Cell, 2006. **23**(2): p. 183-93.

8. Grigoriadou, C., S. Marzi, D. Pan, C.O. Gualerzi, and B.S. Cooperman, *The translational fidelity function of IF3 during transition from the 30 S initiation complex to the 70 S initiation complex.* J Mol Biol, 2007. **373**(3): p. 551-61.

9. Betney, R., E. de Silva, J. Krishnan, and I. Stansfield, *Autoregulatory systems controlling translation factor expression: thermostat-like control of translational accuracy.* RNA, 2010. **16**(4): p. 655-63.

10. Hennelly, S.P., A. Antoun, M. Ehrenberg, C.O. Gualerzi, W. Knight, J.S. Lodmell, and W.E. Hill, *A time-resolved investigation of ribosomal subunit association.* J Mol Biol, 2005. **346**(5): p. 1243-58.

11. Ali, I.K., L. Lancaster, J. Feinberg, S. Joseph, and H.F. Noller, *Deletion of a conserved, central ribosomal intersubunit RNA bridge.* Molecular Cell, 2006. **23**(6): p. 865-874.

12. Fabbretti, A., C.L. Pon, S.P. Hennelly, W.E. Hill, J.S. Lodmell, and C.O. Gualerzi, *The real-time path of translation factor IF3 onto and off the ribosome.* Mol Cell, 2007. **25**(2): p. 285-96.

13. Pon, C.L. and C.O. Gualerzi, *Mechanism of translational initiation in prokaryotes. IF3 is released from ribosomes during and not before 70 S initiation complex formation.* FEBS Lett, 1986. **195**(1-2): p. 215-9.

14. Fei, J., J.E. Bronson, J.M. Hofman, R.L. Srinivas, C.H. Wiggins, and R.L. Gonzalez, Jr., *Allosteric collaboration between elongation factor G and the ribosomal L1 stalk directs tRNA movements during translation.* Proc Natl Acad Sci U S A, 2009. **106**(37): p. 15702-7.

15. Dallas, A. and H.F. Noller, *Interaction of translation initiation factor 3 with the 30S ribosomal subunit.* Mol Cell, 2001. **8**(4): p. 855-64.

16. DeLano, W., *The PyMOL Molecular Graphics System, Version 1.3, Schrödinger, LLC.* 2008.

17. Allen, G.S., A. Zavialov, R. Gursky, M. Ehrenberg, and J. Frank, *The cryo-EM structure of a translation initiation complex from Escherichia coli.* Cell, 2005. **121**(5): p. 703-12.

18. Marshall, R.A., M. Dorywalska, and J.D. Puglisi, *Irreversible chemical steps control intersubunit dynamics during translation.* Proc Natl Acad Sci U S A, 2008. **105**(40): p. 15364-9.

19. Wang, J., *Regulation of IF2 Binding Kinetics and 30S IC Conformational Dynamics during Translation Initiation*, 2010, Columbia University.

Chapter 4

Real-time observation of IF1 binding to the small ribosomal subunit during translation initiation*

4.1 Introduction

Translation initiation involves formation of a 30S IC that, upon 50S subunit joining, ultimately forms an elongation-competent 70S IC (Figure 1.2) (Reviewed in Refs. [1-5]). Assembly of a 30S IC involves the association of IF1, IF2(GTP), IF3, initiator tRNA, and mRNA with a 30S subunit. The efficiency with which a particular mRNA is assembled within a 30S IC is a major determinant of its expression level [6]. There can be numerous *cis* or *trans*-acting regulators that affect an mRNA's interaction with the 30S subunit, leading either to its activation or repression [6-8]. One of these regulators is the strength of the mRNA's translation initiation region (TIR) [9-11]. Considering that the efficiency of the overall 30S IC assembly is likely to vary across mRNAs depending on the details of each mRNA's TIR, 30S IC assembly may serve as a critical point of regulation for the translational control of gene expression. The molecular details of 30S IC assembly remain poorly defined, however, thus preventing a full understanding of the assembly pathway(s) and its regulation.

Due to its role in ribosome recycling [12-16] and the high affinity with which it binds to the 30S subunit [17], current models for 30S IC assembly propose that IF3 is the first initiation component to bind to the 30S subunit and that it remains bound to the 30S subunit throughout the 30S IC assembly process. The order and relative timing of IF1, IF2(GTP), mRNA, and tRNA binding to the IF3-bound 30S subunit, however, are still debated [2]. In fact, virtually every possible permutation of the binding order for these components has been proposed [3, 18-27],

* *Elvekrog, MM and Gonzalez Jr, RL (2011) Real-time observation of IF1 binding to the small ribosomal subunit during translation initiation (In preparation)*

leading to some models in which the 30S IC assembly pathway is described as being random [4, 28].

In general, single-molecule methods provide a powerful means for dissecting the assembly of macromolecular complexes such as the 30S IC, as well as for characterizing the mechanistically revealing kinetics of each step in the assembly pathway(s). As a good example, single-molecule fluorescence was used recently to reveal the assembly pathway of the spliceosome, which is a macromolecular machine containing over 100 protein and RNA components [29]. That study elucidated the order and reversibility of the main steps in the assembly pathway, revealing that splicing can be regulated at any step during spliceosome assembly. As a first step toward developing a complete kinetic scheme describing 30S IC assembly, I have developed an smFRET signal between IF3 and IF1. Due to its stable and early association with the 30S subunit, IF3 is an appropriate static landmark from which to monitor the kinetics of IF1 binding to the 30S subunit during 30S IC assembly.

IF1 is thought to have roles in initiator tRNA and start codon selection [30-32], as well as an influence on the stability of IF2 binding to the 30S subunit [33]. Thus, determining the timing of IF1 binding to the 30S subunit relative to tRNA, mRNA, and IF2 binding to the 30S subunit will provide previously inaccessible insight into the function of IF1 in regulating substrate selection and IF2's stability on the 30S subunit. IF1 may also have functions in later steps in initiation, specifically in regulating the rate of subunit joining [30, 34, 35], as well as possibly stabilizing IF2's interaction with fMet-tRNA^{fMet} for optimal subunit joining [36, 37]. The relative timing of IF1's dissociation either from the 30S IC or 70S IC has important mechanistic implications for its function in initiation.

The results described in this chapter reveal that IF1 binds to the 30S subunit reversibly in the presence of just IF3, as well as in the presence of IF3 and initiator tRNA. Analogous experiments performed in the presence of both IF2 and IF3 demonstrate that IF2 virtually traps IF1 onto the 30S subunit, though additional experiments are needed to probe the reversibility or irreversibility of IF1's binding to the 30S subunit in the presence of IF2. Future experiments also

include measurements of the timing of IF1's dissociation from the 30S IC with respect to 50S subunit joining.

4.2 Experimental Methods

4.2.1 Preparation of 30S Initiation Complexes (30S ICs)

The nomenclature used throughout this chapter will follow the style introduced in Chapter 2 with the exception that all 30S ICs here necessarily contain 30S ribosomal subunits, IF1(Cy5), IF3(Cy3), and a 64 nucleotide 5'-biotin-mRNA containing either an AUG or AUU start codon (mRNAs #1 or #2 in Table 2.1). Thus, for example, an initiation complex referred to as $30SIC^{fMet}_{-IF2}$ is composed of a 30S subunit, IF1(Cy5), IF3(Cy3), an AUG-containing mRNA, and fMet-$tRNA^{fMet}$, but is lacking IF2. Details on the design, preparation, labeling, and biochemical activity testing of IF1(Cy5) and IF3(Cy3) are found in sections 3.3.1→3.3.3 and 2.3.1→2.3.5, respectively. Low-salt Tris-polymix buffer was used for all smFRET and ensemble biochemical experiments, with the concentration of $Mg(OAc)_2$ determined by the experiment. 30S IC components (ribosomes, aa-tRNAs, initiation factors, mRNAs) were all from *E. coli* and prepared as described in sections 2.2, 5.1 and 5.2, and detailed in Ref. [38].

4.2.2 TIRF microscope imaging of surface-immobilized 30S ICs

30S ICs were assembled at micromolar concentrations of all initiation components with the ribosomes and fluorescently-labeled components at limiting concentrations [1.8 μM 5'-bi-mRNA, 0.6 μM 30S ribosomal subunits, 0.9 μM fMet-$tRNA^{fMet}$ (when included), 0.9 μM IF2 (when included), 1 mM GTP, 0.6 μM IF1(Cy5), 0.6 μM IF3(Cy3)] and incubated for 10 minutes at 37°C. 30S ICs were then aliquoted, flash frozen with liquid nitrogen, and stored at -80°C until further use. For TIRFM imaging, 30S ICs were diluted to 100-300 pM concentrations in 5 mM $Mg(OAc)_2$ low-salt Tris-polymix buffer containing 50 nM IF1(Cy5), 0.9 μM IF2 (when included), and 0.9 μM fMet-$tRNA^{fMet}$ (when included), and delivered to a PEG-biotin-streptavidin passivated quartz microscope flowcell (see Figure 1.17 and section 2.2.7). The final imaging buffer contained an oxygen scavenging system (glucose oxidase and catalase) and triplet state quenchers (1,3,5,7-

cyclooctatetraene and *p*-nitrobenzyl alcohol) at the concentrations optimized for IF3(Cy3-Cy5) TIRFM work (see section 2.2.7). All data were collected under 7 mW 532 nm laser power excitation (measured at the prism), and at 10 frames per second time resolution for 1200 frames (two minutes).

4.2.3 smFRET data analysis

Trajectory selection and histogram generation were performed as described in section 1.4.3. The average non-zero FRET value (0.80 ± 0.09) was determined from fitting each normalized population FRET histogram (first ten seconds of data) with two Gaussian curves in Origin 8.0 (OriginLab). The error in the average FRET value (0.09) is a measure of the broadness of the FRET state, which is a function of the signal-to-noise ratio, the range of inter-fluorophore distances sampled by the FRET pair, and the number of trajectories in the data set, and was calculated from the average ½(FWHM) of the non-zero peaks. The standard deviation of the value of the peak center from nine randomly chosen data sets is 0.02, indicating the reproducibility of the FRET value both day-to-day and among different 30S ICs. The rates of transitions between zero and non-zero FRET states were calculated as described in Figure 4.4.

4.3 Results and Discussion

4.3.1 Design of an IF1-IF3 smFRET signal

The lack of structural information on fully assembled 30S ICs hindered the rational design of an IF1-IF3 smFRET signal. Instead, the signal was developed serendipitously with reagents originally designed and prepared for other research projects. Cy3-labeled IF3 was prepared en route to development of dual-labeled IF3 for investigations of IF3's intramolecular dynamics (see Chapter 2). Both IF3(C65S-S38C)-Cy3 and IF3(C65S-K97C)-Cy3 were available and biochemically active (see section 2.3.5). IF1(Cy5), on the other hand, was originally developed as a FRET partner for both IF2 and fMet-tRNA^{fMet}. Early efforts in my Ph.D. work were put toward the development of labeled IF1 for investigations into its role in positioning of IF2 and initiator tRNA on 30S ICs and 70S ICs. This work was motivated by two different 70S IC cryo-EM

structures, prepared in the absence and presence of IF1, which revealed distinctly different IF2 conformations on the 70S IC [36, 37, 39]. Although this original hypothesis remains untested, another member of the Gonzalez research group, Dr. Jiangning Wang, used these triangulated smFRET signals (IF1-IF2, IF2-tRNA, tRNA-IF1) to investigate the conformational dynamics of the 30S initiation complex [40]. The approach taken to label IF1 followed the same route used to label other translation factors [38], that is, the site-specific labeling of a single Cys residue in the protein using maleimide-conjugated Cy5. Wild-type *E. coli* IF1 does not contain any Cys residues, so multiple sequence alignment coupled with structural analysis and previous IF1 mutagenesis studies identified three positions in the IF1 amino acid sequence where replacement with Cys could be made: Q10C, L14C, and T33C [41-43]. IF1-Q10C (see Figure 4.1B) was the first positive hit during the site-directed mutagenesis stage, and all subsequent work was performed with this mutant. Multiple attempts at generating IF1-L14C and IF1-T33C proved unsuccessful.

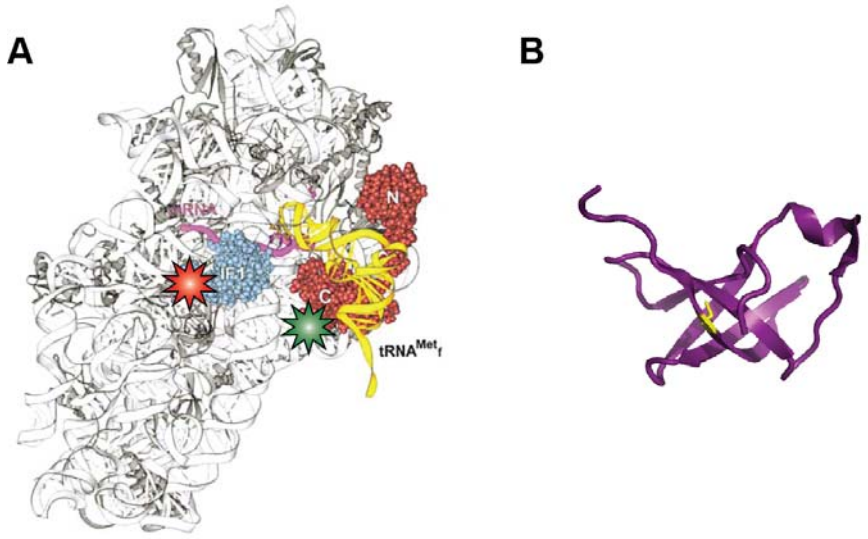

Figure 4.1 An IF1(Cy5)-IF3(Cy3) smFRET signal. A. The binding site of IF3 was modeled into the 30S-IF1 crystal structure [35] based on constraints from directed hydroxyl radical probing [44]. IF3's NTD is placed near the E site and its CTD is close to the P site. The mRNA and P-site tRNA binding sites are indicated. Figure adapted from Ref. [44]. The red and green stars indicate Cy5 and Cy3, respectively. **B.** Ribbon diagram of IF1(Q10C) with the Gln-to-Cys mutation indicated in yellow. This residue was labeled with maleimide-conjugated Cy5. IF1 was modeled and rendered in PyMOL [45] using the PDB file 1AH9.

4.3.2 Preparation of fluorescently-labeled and biochemically-active IF1

As described in section 2.2.5 and Ref. [38], the cloning strategy for wild-type IF1 and the pProEX-HTb plasmid (Invitrogen) employed necessitated that the *N*-terminal end of IF1 differ slightly from the wild-type *E. coli* gene sequence. Specifically, the *N*-terminal end is G-A-M1, where the underlined amino acid and sequence position denote the beginning of the wild-type gene sequence. The mutant IF1(Q10C) was generated by QuickChange II-E site-directed mutagenesis (Stratagene). After positive gene sequence identification (Genewiz, South Plainfield, NJ), the pProEX-HTb plasmid containing the IF1(Q10C) gene was transformed into BL21-DE3 cells for optimal protein overexpression with Isopropyl β-D-1-thiogalactopyranoside (IPTG). His_6-tagged IF1 was overexpressed and purified as previously described [38] using a standard metal-chelate affinity chromatography approach involving two Ni^{2+}-NTA columns with an intermediate TEV cleavage step. Lastly, the protein was purified to homogeneity using gel filtration (Superdex 75, GE Healthcare). IF1 was labeled with 20-fold excess of Cy5 maleimide following one hour at room temperature with ten-fold molar excess of tris(2-carboxyethyl)phosphine (TCEP) to reduce any disulfide bonds. Unreacted dye was removed by gel filtration, and unlabeled IF1(Q10C) was separated from IF1(Q10C)-Cy5 by hydrophobic interaction chromatography (HIC) (see Figure 4.2). The collected HIC fractions were 100% Cy5-labeled, enabling efficient 30S IC complex formation and TIRFM imaging. Protein concentration was measured by the Bradford method [46].

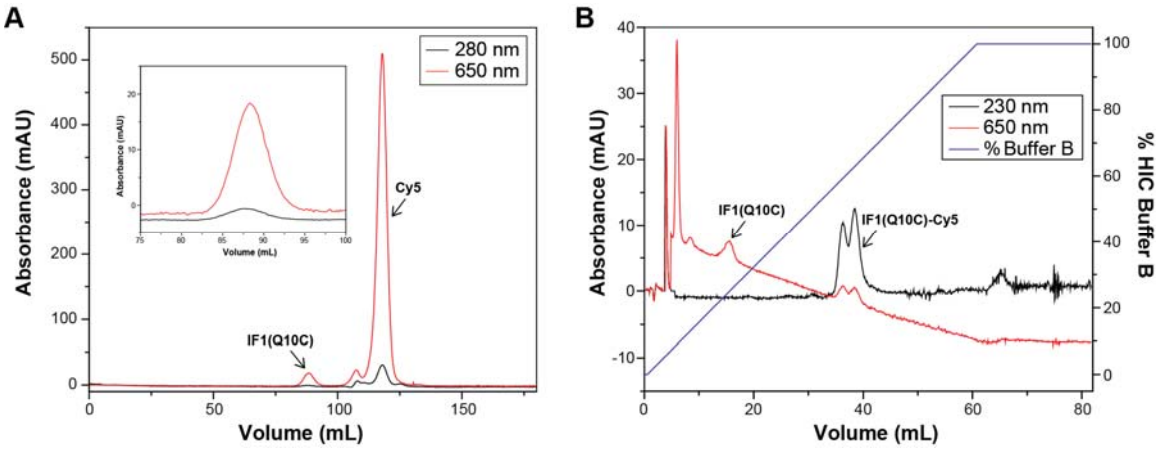

Figure 4.2 Purification of IF1(Cy5). A. Gel filtration (Superdex 75, GE Healthcare) was used to separate IF1(Cy5) from unreacted dye following the labeling reaction. The peak at 90 mL (see also the figure inset) is IF1(Cy5), while the dominant peak at 120 mL is unreacted dye. **B.** Hydrophobic interaction column chromatography (Tosoh Bioscience) was used to further purify IF1(Cy5) from unlabeled IF1. The peak at 35

mL is labeled IF1 and the peak at 15 mL is unlabeled IF1, as confirmed by injection of only unlabeled IF1 (not shown). The protein was monitored at 230 nm due to its low extinction coefficient at 280 nm. The signal at 650 nm is indicative of Cy5. This sample was concentrated and tested for biochemical activity with the Toeprinting assay.

The biochemical activity of IF1(Q10C)-Cy5 (hereafter IF1(Cy5)) was verified by a primer-extension inhibition assay ("toeprinting") [47]. Specifically, this assay tests IF1's ability to enhance the formation of correctly initiated 70S ICs in the presence of IF2 and IF3. 30S and 50S subunits are mixed with ^{32}P-primer-annealed mRNA, IF1, IF2, IF3, and an equal mixture of fMet-tRNA^{fMet} and tRNA^{Phe} in a buffer containing ATP and dNTPs for reverse transcription. After allowing the reaction mixture to equilibrate, primer-extension was initiated by addition of avian myeloblastosis virus (AMV) reverse transcriptase (New England Biolabs). Reactions performed in the presence of IF1 generally show an approximately three-fold increase in the intensity at the +15 (fMet-tRNA^{fMet}) toeprint compared with reactions performed in the absence of IF1 (Figure 4.3).

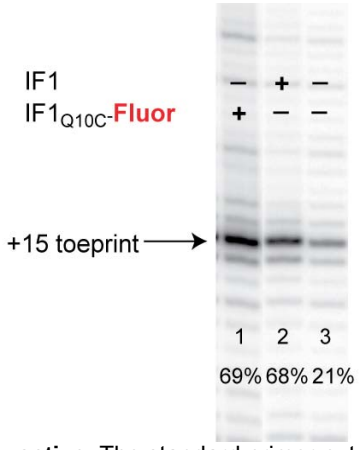

Figure 4.3 IF1(Cy5) is biochemically active. The standard primer-extension inhibition ("Toeprinting") assay was performed to test the activity of IF1(Cy5) in enhancing the formation of a correctly initiated 70S IC in the presence of IF2 and IF3. IF2(GTP), IF3, fMet-tRNA^{fMet}, tRNA^{Phe}, 30S and 50S ribosomal subunits were present in all three reactions. The presence of IF1 enhances the percent of correctly initiated 70S ICs by 3-fold (lanes 2 vs. 3). IF1(Cy5) has activity comparable to unlabeled IF1 (lanes 3 vs. 2). The extent of 70S IC formation was calculated by taking the ratio of signal intensity at the +15 toeprint over the signal intensity of the full length cDNA (not shown) [47].

4.3.3 Preparation of fluorescently-labeled and biochemically-active IF3 NTD and CTD variants

As described in Chapter 2 (section 2.3.1), numerous residues on IF3's NTD and CTD were chosen for conversion to Cys in the development of a double Cys-containing IF3 for dual Cy3- and Cy5-labeling (see Table 2.2). Along the way, single-Cys IF3 variants were also purified

and Cy3- or Cy5-labeled in order to determine the least perturbing locations for fluorophore labeling. IF3(C65S-S39C)-Cy3 and IF3(C65S-K97C)-Cy3 are single NTD and CTD Cys IF3 variants, respectively, which still show activities comparable to wild-type IF3. Their activities were tested with toeprinting as described in sections 2.3.5, 5.4, and Ref. [38] (see Figure 2.6).

4.3.4 IF3's CTD is within FRET distance of IF1 on $30SIC^{-tRNA}_{-IF2}$

The precise location of IF3's NTD and CTD on the 30S subunit is not known, but directed hydroxyl radical probing placed the CTD on the platform near the P site using a number of constraints [44]. The location of the NTD was tentatively placed near the E site on the 30S subunit, though this placement was based on just a single NTD residue's cleavage pattern. This placement of the NTD contradicted the findings from a low-resolution cryo-EM structure of IF3-30S, however [48]. Considering the uncertainties in the placement of IF3's NTD and CTD on the 30S subunit, my initial attempts at developing an IF1-IF3 smFRET signal involved testing both IF3(C65S-S39C)-Cy3 and IF3(C65S-K97C)-Cy3 in $30SIC^{-tRNA}_{-IF2}$ to determine which IF3 variant, if either, was close enough to IF1(Cy5) for FRET to occur (see section 1.4). IF3(C65S-K97C)-Cy3 reproducibly failed to cause FRET with IF1(Cy5) in $30SIC^{-tRNA}_{-IF2}$, indicating that IF3'S NTD is greater than ~90 Å from IF1's Q10C residue (assuming R_0 for Cy3/5 of 55 Å [49, 50]), at least on $30SIC^{-tRNA}_{-IF2}$ (recall from Chapter 2 that IF3 adopts a different conformational state on 30S ICs +/- IF1, IF2, and initiator tRNA). IF3(C65S-S39C)-Cy3, on the other hand, is within FRET range of IF1 on all 30S ICs tested (+/- IF2 and/or initiator tRNA). IF1-Cy5 and IF3(C65S-S39C)-Cy3 (hereafter IF3-Cy3) produce a FRET signal with an efficiency of 0.80 ± 0.09, which corresponds to approximately 44 Å (assuming R_0 for Cy3/5 of 55 Å [49, 50]) separation between the two fluorophores. These results are in line with the placement of IF3's CTD near the P site [44, 51, 52] and the NTD binding site distal to IF1's binding site, perhaps closer to the E site (see Figure 4.1), at least on $30SIC^{-tRNA}_{-IF2}$. Future experiments can be done to test whether IF3(C65S-K97C)-Cy3 is within FRET distance of IF1(Q10C)-Cy5 on $30SIC^{fMet}$, which is a 30S IC in which IF3 undergoes a large-scale conformational change relative to $30SIC^{-tRNA}_{-IF2}$ (see Chapter 2). These

experiments would give credence to the model that IF3's NTD samples different binding sites on the 30S subunit depending on the presence and identity of other initiation components (see section 2.4 and Figure 2.8).

4.3.5 The binding kinetics of IF1 on $30SIC^{-tRNA}_{-IF2}$

smFRET trajectories recorded of IF3(Cy3)-IF1(Cy5) on $30SIC^{-tRNA}_{-IF2}$ with IF1(Cy5) at 15 nM concentration in solution were divided into two distinct subpopulations: fluctuating and static (Figure 4.6). The majority (94%) are fluctuating trajectories that show frequent transitions between zero and non-zero FRET (0.80 $\pm$ 0.09 FRET) (see Figure 4.4A). This type of behavior can arise in single-molecule FRET experiments due to three main phenomena: (1) association/dissociation events of the acceptor (Cy5)-labeled ligand; (2) a reversible large-scale conformational change of the ribosomal subunit which brings the Cy3- and Cy5-labeled proteins into and out of FRET distance range; (3) it could result from the photophysical phenomenon known as "blinking" which occurs when Cy5 transits into a non-fluorescent, triplet (dark) state for a period of time, resulting in zero FRET efficiency [53, 54]. In order to determine the source of this FRET signal, I performed an IF1(Cy5) concentration series, monitoring the IF1-IF3 FRET signal of $30SIC^{-tRNA}_{-IF2}$ over a range of IF1(Cy5) concentrations from 15 nM to 75 nM. This concentration range was chosen based on the previously reported K_d of 28 nM for IF1 binding to 30S-IF3 complexes [28]. Also, I was restricted to an upper limit of 75 nM IF1(Cy5) due to the decrease in the signal-to-noise ratio that comes with increased concentrations of Cy5 in solution. One of the benefits of TIRFM is its restricted penetration depth, which limits direct excitation to those fluorophores in the evanescent field near the quartz-water interface; nevertheless increased concentrations of Cy5 in solution do result in increased background fluorescence due to indirect excitation of Cy5 by the 532 nm excitation laser. This increase in background signal results in additional noise in the Cy3-Cy5 intensity versus time trajectories and FRET versus time trajectories (see, for example, Figure 4.5).

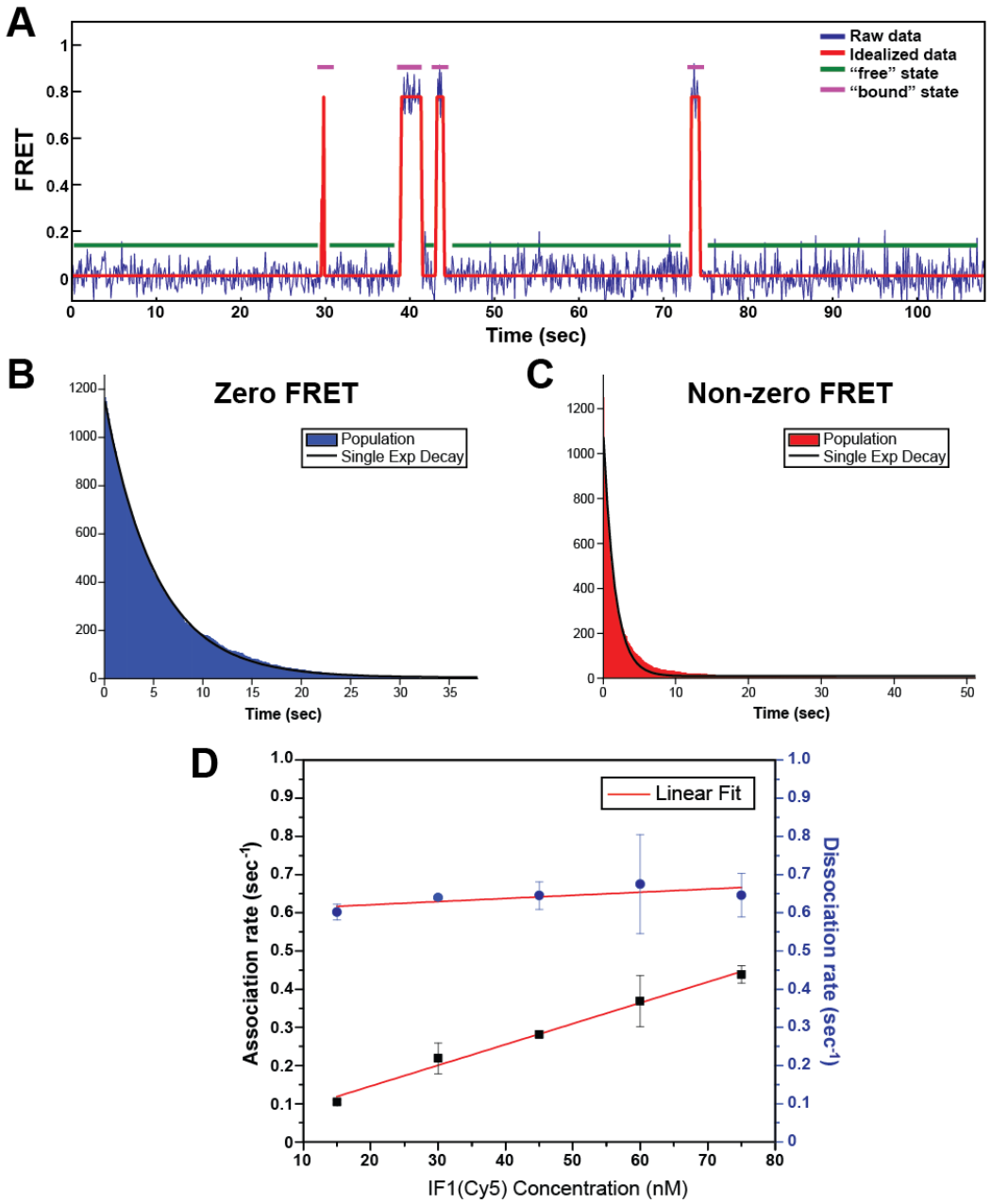

Figure 4.4 Kinetic analysis of smFRET versus time data. A. Sample raw (blue) and idealized (red) FRET trajectory from 30SIC$^{-tRNA}_{-IF2}$ data. The trajectory was idealized as a Hidden Markov model using the vbFRET software package [55]. The "bound" state includes all data points greater than or equal to 0.2 FRET, while the "free" state includes all points below 0.2 FRET. **B.** Sample dwell time histogram of zero FRET state data. **C.** Sample dwell time histogram of data from the non-zero FRET state. **D.** A plot of the k_a and k_d values versus IF1(Cy5) concentration. k_a and k_d were calculated as: $k_a = 1/\tau_{off}$ and $k_d = 1/\tau_{on}$, where τ_{off} and τ_{on} were measured from fitting each dwell time histogram (see **B** and **C**) with a single exponential decay. Error bars represent the standard deviation of the mean rate from duplicate data sets.

The resulting FRET trajectories display an IF1(Cy5) concentration dependent increase in the rate of transitions from the zero to non-zero state. The concentration dependence of the IF1 association rate, and the concentration independence of its dissociation rate, clearly indicates that this FRET signal is a read-out of IF1(Cy5) association and dissociation, not Cy5 blinking or a large-scale conformational change of the ribosomal subunit. This concentration dependent behavior is consistent with a pseudo first order bimolecular binding process. A plot of the zero to non-zero FRET transition rate as a function of IF1(Cy5) concentration (Figure 4.4D) yields a second order rate constant for the association of IF1 (k_a) with $30SIC^{-tRNA}_{-IF2}$ of 5.45 $\pm$ 0.18 $\mu M^{-1}s^{-1}$, determined from the slope of a linear fit of the k_a versus IF1 concentration data. The transition rate from the non-zero to zero FRET states, on the other hand, is independent of the IF1(Cy5) concentration, consistent with a unimolecular dissociation process. This first order rate constant for IF1(Cy5) dissociation (k_d) from $30SIC^{-tRNA}_{-IF2}$ is 0.82 $\pm$ 0.42 s^{-1}. An equilibrium dissociation constant, K_d, can be calculated from these rate constants using the relationship $K_d = k_d/k_a$, resulting in a value of 150 $\pm$ 102 nM.

Though kinetic measurements on IF1's binding to 30S subunits have not been previously measured, thermodynamic measurements have been made of its association with 30S subunits in the presence of IF3 and/or IF2. The K_d value measured here is comparable to the previously measured K_d (28 nM) for IF1 binding to 30S subunits in the presence of IF3, though lacking mRNA. The 5-fold difference in these values may arise from the presence versus absence of mRNA in the 30S ICs, but is more likely to result from differences in the buffer conditions used in the two experiments. IF1 K_d measurements have previously been shown to be very sensitive to ionic strength, especially NH_4^+ concentrations [28]. Another source of error in these K_d measurements may be from the route used here to calculate k_a and k_d. It is possible that k_a was underestimated by excluding single data point transitions between FRET states. Exclusion of these data points would result in an increase in the value of the zero FRET dwell times, leading to smaller k_a values. Alternatively, or additionally, k_d may have been overestimated since a

correction for the photobleaching rate was not performed here. The overestimation of k_d and underestimation of k_a results in an overestimated K_d. Thus, the ~150 nM K_d measured here represents an upper limit of the actual K_d value for IF1's binding to 30S-IF3-mRNA complexes. In the future, this experiment will be repeated to obtain triplicate results, and a rigorous photobleaching correction will be employed to correct the k_d values.

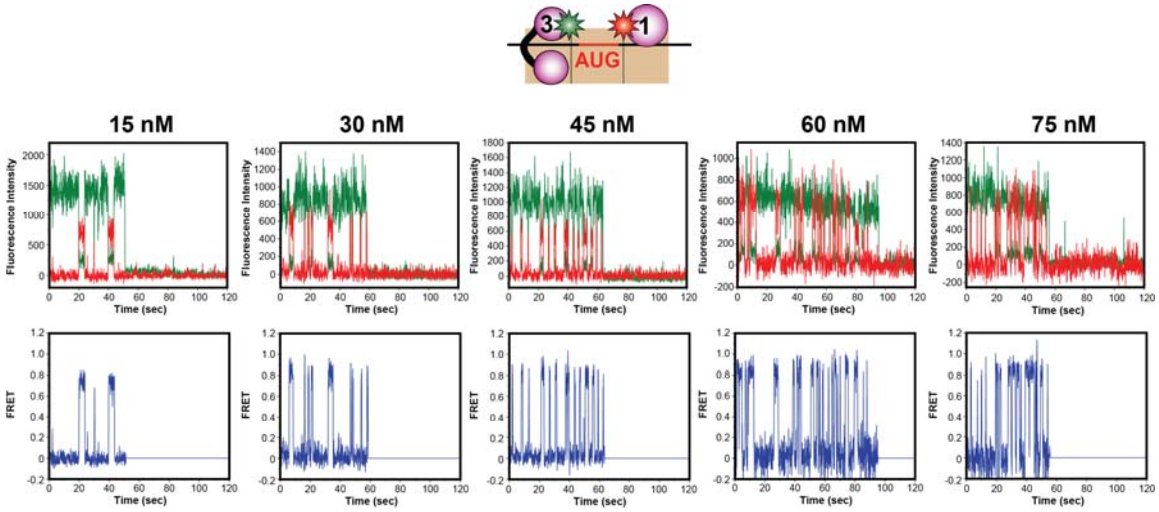

Figure 4.5 An IF1(Cy5) concentration dependent increase in the population of the non-zero FRET state. $30SIC^{-tRNA}_{-IF2}$ complexes imaged with 15 nM, 30 nM, 45 nM, 60 nM, and 75 nM IF1(Cy5) in solution. The concentration dependence to the FRET signal indicates that the signal is due to reversible IF1(Cy5) dissociation and re-binding behavior.

An important feature of the IF3(Cy3)-IF1(Cy5) FRET trajectories not just from $30SIC^{-tRNA}_{-IF2}$, but from all the collected data sets, is the existence of two main subpopulations: (1) trajectories that display transitions between the zero and non-zero FRET states, and (2) trajectories in which only the non-zero FRET state is sampled before signal loss due to Cy3 or Cy5 photobleaching, which is an irreversible photochemical destruction of the fluorophores (Figure 4.6). As will be discussed in later sections, the percentage of trajectories within each of these two subpopulations changes as a function of the initiation components present in the 30S IC (i.e. +/- tRNA, IF2). The former subpopulation arises from multiple, reversible IF1(Cy5) binding and dissociation events during the observation period (two minutes) prior to Cy3 photobleaching whereas the latter subpopulation reports just a single IF1(Cy5) binding event, whose lifetime in the non-zero FRET state is mainly limited by photobleaching. The percentage of trajectories within each of these two

subpopulations is a function of the stability of IF1 on the particular 30S IC and is indicative of the reversibility or irreversibility of IF1's binding to the particular 30S IC. As is clear from the data here, IF1 binds reversibly to 30S subunits carrying IF3.

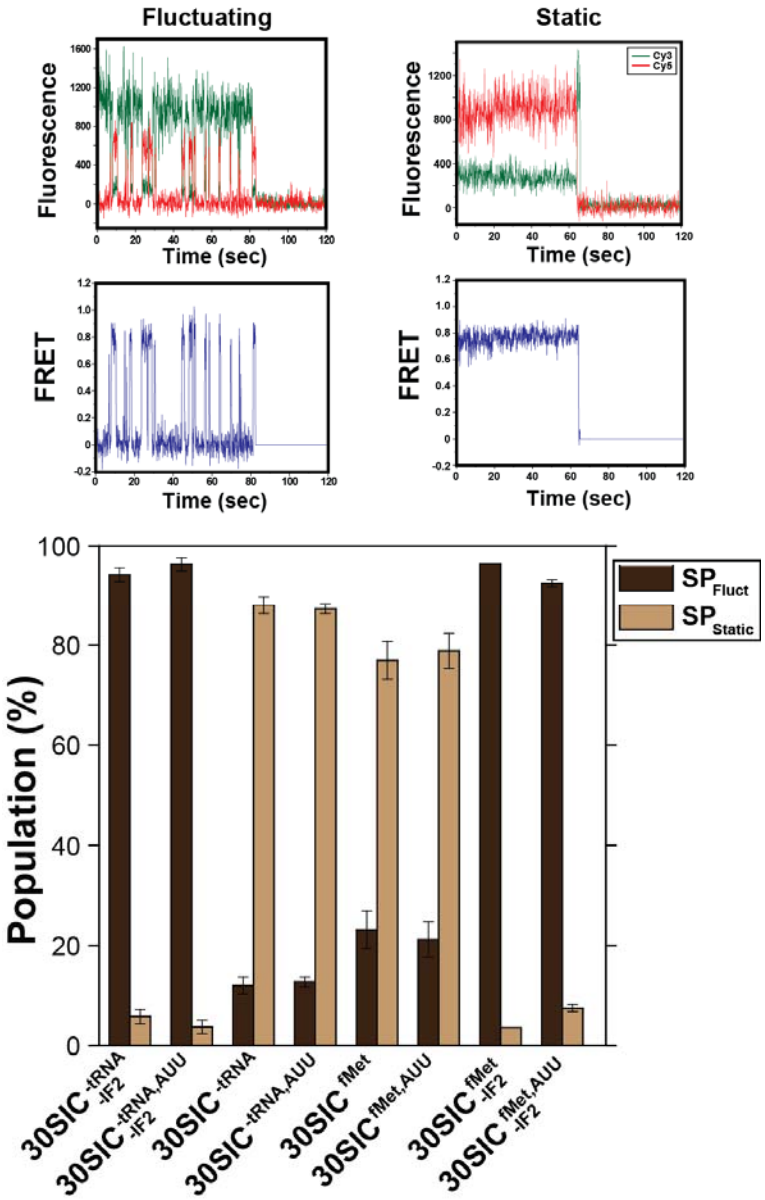

Figure 4.6 Subpopulation analysis of IF1(Cy5)-IF3(Cy3) smFRET trajectories. SP_{Fluct} refers to the trajectories are those which exhibit at least one IF1(Cy5) dissociation and re-binding event. SP_{Static} refers to the trajectories that remain in a non-zero FRET state until irreversible dissociation or photobleaching. Errors bars, when present, are from duplicate data sets and represent the standard deviation of the mean population percentage. All experiments were performed at 50 nM IF1(Cy5). All experiments need to be triplicated. See Table 4.1 for the exact values of each subpopulation and its error term.

Table 4.1: IF1(Cy5) association and dissociation rate constants, equilibrium dissociation constants, and FRET trajectory subpopulation analysis

	k_a (μM^{-1} s^{-1})	k_d (s^{-1})	K_d (nM)	$SP_{Fluctuating}$ (%)	SP_{Static} (%)
$30SIC^{-tRNA}_{-IF2}$ [a]	5.5 ± 0.2 [c]	0.8 ± 0.4 [c]	150 ± 102 [c]	94.2 ± 1.4	5.8 ± 1.4
$30SIC^{-tRNA,AUU}_{-IF2}$ [b]	5.5 ± 2.9 [d]	0.7 ± 0.2 [d]	140.1 ± 34.7 [d]	96.3 ± 1.3	3.7 ± 1.3
30S IC^{-tRNA}	(N.D.)	(N.D.)	(N.D.)	12 ± 1.7 [d]	88 ± 1.7 [d]
30S $IC^{-tRNA,AUU}$	(N.D.)	(N.D.)	(N.D.)	12.7 ± 1 [d]	87.3 ± 1 [d]
30S IC^{fMet}	(N.D.)	(N.D.)	(N.D.)	23.1 ± 3.8 [d]	76.9 ± 3.8 [d]
30S $IC^{fMet,AUU}$	(N.D.)	(N.D.)	(N.D.)	21.2 ± 3.5 [d]	78.8 ± 3.5 [d]
$30SIC^{fMet}_{-IF2}$ [b]	6.0	0.4	67.7	96.4	3.6
$30SIC^{fMet,AUU}_{-IF2}$ [b]	7.9 ± 4.6 [d]	1.1 ± 0.9 [d]	129.6 ± 37.3 [d]	92.5	7.5

[a] k_a calculated from the slope, m, of a linear fit (y=mx+b) of ($1/(\tau_{zero})$) vs. [IF1-Cy5]; $k_d = 1/(\tau_{nonzero})$; $K_d = k_d/k_a$

[b] $k_a = 1/((\tau_{zero})*([IF1\text{-}Cy5]))$; [IF1-Cy5] = 50 nM; $k_d = 1/\tau_{nonzero}$; $K_d = k_d/k_a$

[c] Error calculated as the average of the slope of linear fits from two separate concentration data series

[d] Error calculated as the standard deviation of the average of two data sets

N.D. = not determined due to the small number of fluctuating trajectories. See section 4.3.6 for a description of how these values can be measured or estimated in the future.

4.3.6 IF2 increases the stability of IF1 in $30SIC^{-tRNA}$

Inclusion of IF2(GTP) within a 30S IC containing IF1(Cy5) and IF3(Cy3) (i.e. $30SIC^{-tRNA}$) results in distinctly different FRET behavior compared with those 30S ICs imaged in the absence of IF2 (see Figure 4.7), however the mean FRET efficiency value of the IF1-IF3 signal remains at 0.80 ± 0.09. In the presence of IF2, IF1(Cy5) now rarely exhibits multiple reversible binding and dissociation events, but instead remains stably bound to the $30SIC^{-tRNA}$ for the duration of Cy5's lifetime before photobleaching in 88% of all trajectories (see Figure 4.7). In the future, an IF2 concentration dependent experiment will be performed to test if the percentage of traces in the fluctuating subpopulation decreases under increasing IF2 concentrations. This experiment will test if the IF2 concentration employed here, 0.9 μM, was saturating. If these experiments were performed under sub-saturating conditions, it is possible that the 12% of the total trajectories that showed at least one dissociation and re-binding event (i.e. "fluctuating") may have arisen from 30S ICs lacking IF2. These IF2 concentration dependent experiments will confirm either the reversibility or the irreversibility of IF1's binding to 30S subunits in the presence of IF2, which is a mechanistically important aspect to both the overall 30S IC assembly pathway, as well as the role of IF1 in regulating the assembly pathway.

The large percentage (88%) of static trajectories for IF1 binding to 30SIC^{-tRNA} complexes indicates IF2's stabilizing effect on IF1. Although the binding kinetics for IF1's interaction with a 30S-IF2-IF3 complex have not been previously investigated, the K_d for IF1's interaction with a 30S-IF3-IF2 complex has been measured using fluorescence anisotropy and shown to be 4 nM [28]. This value is three orders of magnitude lower than IF1's K_d for binding with just an empty 30S subunit, which is 2 μM [28], clearly demonstrating IF2's stabilizing effect on IF1.

The small percentage of reversible IF1 binding events within these 30SIC^{-tRNA} complexes requires alternate approaches to estimate IF1's binding kinetics. Careful measurement of IF1's k_a will require stopped-flow delivery of IF1 to 30SIC^{-tRNA} in order to accurately measure its initial binding to the 30S subunit. A lower limit for the k_d can be estimated from the IF1(Cy5) photobleaching rate. For these experiments, a laser shuttering strategy can be employed to extend the fluorophores' lifetime and more accurately measure IF1's residence time on these 30SIC^{-tRNA} complexes. These experiments, in combination with an IF2 concentration series, will enable measurements of IF1's binding kinetics and determine the reversibility or irreversibility of this binding pathway.

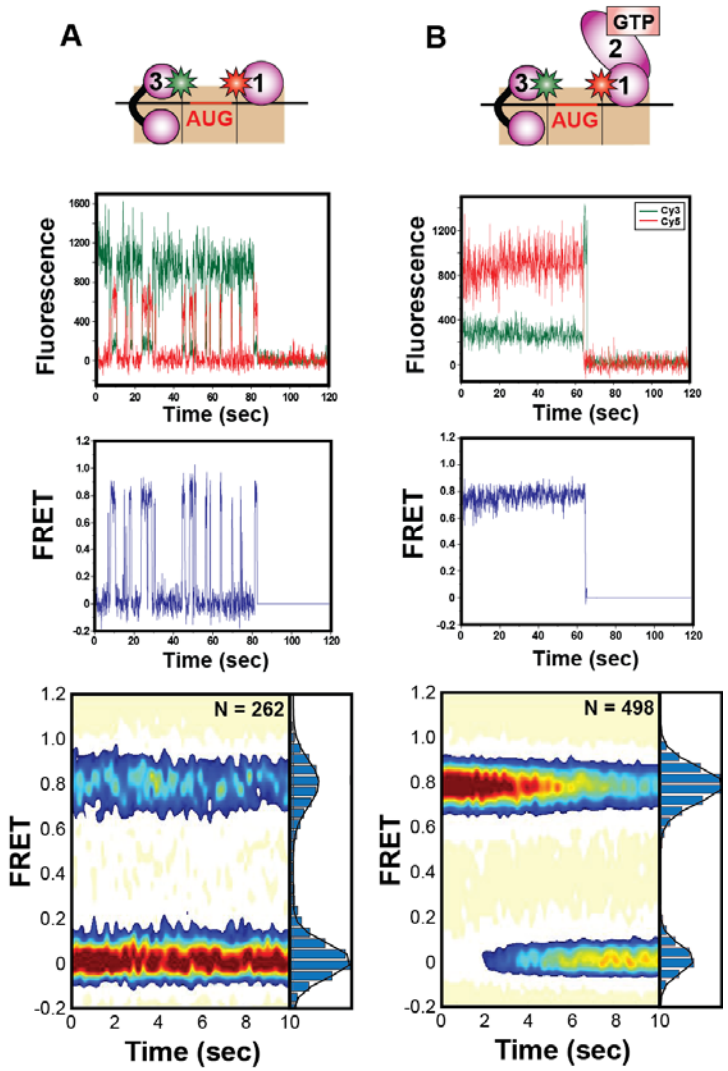

Figure 4.7 IF2 stabilizes IF1 on the 30S IC. Top row: cartoon representation of the contents of the 30S IC (**A.** $30SIC^{-tRNA}_{-IF2}$ **B.** $30SIC^{-tRNA}$). The red star indicates Cy5 and the green star indicates Cy3. Second row: representative fluorescence versus time trajectories. Third row: the corresponding FRET trajectories calculated as FRET = $I_{Cy5}/(I_{Cy3}+I_{Cy5})$. Bottom row: Time-evolution of population FRET histograms. The normalized population histogram to the right of each 2D FRET histogram is plotted with the first 100 frames (10 seconds) of data to minimize contamination due to photobleaching. The number of trajectories in each plot is indicated in the upper right hand corner ("N"). IF1(Cy5) was included in solution at 50 nM concentration.

IF2's stabilizing effect on IF1 is reasonable in light of known functions of IF1 and IF2 during translation initiation. The interplay between IF1 and IF2 on 30S ICs, and even 70S ICs, is well established and it has been suggested that IF1's main function in translation is to position or stabilize IF2 so that it can effectively perform its roles in initiator tRNA selection and 50S subunit

joining [56]. IF1's dissociation may be required for IF2 release though this remains speculative and the timing of IF1's dissociation during initiation is unknown [33]. IF1 is a highly conserved protein across all domains with an exception being that no IF1 homolog is present in mammalian mitochondria. Interestingly, recent findings reveal that the sequence of mitochondrial IF2 ($IF2_{mt}$) contains a 37 amino acid insertion that biochemical and molecular modeling data show substitutes for the function of IF1 [57]. A cryo-EM structure of $IF2_{mt}$ in complex with initiator tRNA on a bacterial 30S subunit reveals that the insertion region of $IF2_{mt}$ binds in the same location that is normally occupied by IF1, confirming the orthology between IF1 and this short domain on $IF2_{mt}$ [58]. Cryo-EM structures of a 70S IC and a pseudo 30S IC reveal the close proximity of IF1 and IF2 on initiation complexes, though discrepancies exist regarding whether or not IF1 and IF2 are in direct contact on the 30S subunit [1, 36].

4.3.7 IF1 is insensitive to the presence of fMet-tRNA^{fMet} on 30SIC$^{fMet}_{-IF2}$

When fMet-tRNA^{fMet} is included at 0.9 μM concentration during 30SIC$^{fMet}_{-IF2}$ preparation, dilution, and imaging, the IF1(Cy5)-IF3(Cy3) FRET signal is largely unaffected by its inclusion compared with 30SIC$^{-tRNA}_{-IF2}$ (see Figure 4.7 and Table 1). The FRET efficiency remains centered at ~0.80, and IF1(Cy5) exhibits reversible association and dissociation to the 30S ICs in 96% of the trajectories. The calculated equilibrium dissociation constant, K_d = 68 nM, is within error of the value for the same 30S IC prepared with an AUG start codon (Table 1), though this result needs to be reproduced. The insensitivity of the FRET signal to the presence or absence of tRNA may be due to the tRNA's short residence time on the 30S IC. It has been shown that IF1 and IF3, in the absence of IF2, play a cooperative role in indiscriminately destabilizing P-site tRNAs by increasing their k_d's; an effect that is counteracted on fully assembled 30S ICs by IF2's selective stabilization of fMet-tRNA^{fMet} [21]. The previously measured K_d for fMet-tRNA^{fMet} on comparable complexes (30SIC$^{fMet}_{-IF2}$) is 460 nM, while in the presence of IF2 it is decreased to 2.7 nM [32].

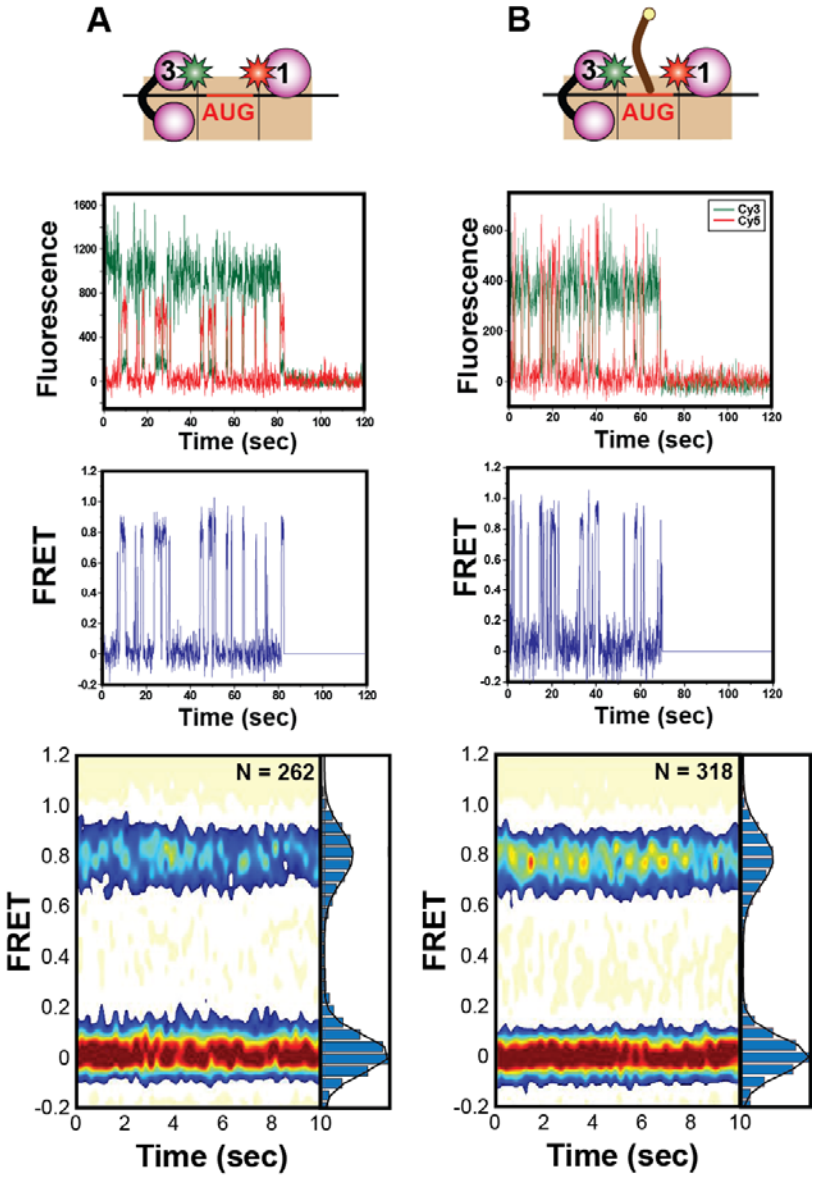

Figure 4.8 IF1's 30S subunit binding kinetics are unaffected by the presence of initiator tRNA. Top row: cartoon representation of the contents of the 30S IC (**A.** $30SIC^{\cdot tRNA}_{\cdot IF2}$ **B.** $30SIC^{fMet}$). The red star indicates Cy5 and the green star indicates Cy3. Second row: representative fluorescence versus time trajectories. Third row: the corresponding FRET trajectories calculated as FRET = $I_{Cy5}/(I_{Cy3}+I_{Cy5})$. Bottom row: Time-evolution of population FRET histograms. The normalized population histogram to the right of each 2D FRET histogram is plotted with the first 100 frames (10 seconds) of data to minimize contamination due to photobleaching. The number of trajectories in each plot is indicated in the upper right hand corner ("N"). IF1(Cy5) was included in solution at 50 nM concentration.

4.3.8 IF1 is stabilized on the completely assembled initiation complex ($30SIC^{fMet}$)

When both fMet-$tRNA^{fMet}$ and IF2 are included in the 30S IC with IF1(Cy5) and IF3(Cy3), IF1(Cy5) is stabilized on the 30S IC, exhibiting reversible binding and dissociation behavior in only 23 $\pm$ 4% of the total trajectories (see Figure 4.9, and Table 1). Comparison of the time-evolution of population FRET histograms and the percentage of trajectories in each subpopulation between $30SIC^{fMet}$ data and those from $30SIC^{-tRNA}_{-IF2}$, $30SIC^{-tRNA}$ and $30SIC^{fMet}_{-IF2}$ (see Figure 4.9 and Table 4.1) suggests that IF2 has the greatest effect on the stability of IF1 in $30SIC^{fMet}$, with the presence or absence of P-site tRNA largely undetected by this IF1-IF3 FRET signal at the concentrations of IF2 and tRNA employed.

Although IF1's kinetic behavior on $30SIC^{-tRNA}$ and $30SIC^{fMet}$ appears similar, the percentage of trajectories that show reversible dissociation and re-binding events in $30SIC^{fMet}$ is reproducibly larger than the percentage of fluctuating trajectories in $30SIC^{-tRNA}$ (23.1 $\pm$ 3.8% vs. 12 $\pm$ 1.7%). This difference may be mechanistically meaningful, though future experiments are needed to confirm this. For one, both IF2 and fMet-$tRNA^{fMet}$ concentration dependent experiments are needed to test if these data were collected under saturating conditions (here: 0.9 μM each of tRNA and IF2). As described in section 4.3.6, the apparent reversibility of IF1's binding to $30SIC^{fMet}$ may arise due to sub-saturating conditions of IF2 and fMet-$tRNA^{fMet}$. These concentration series experiments also have the potential to reveal the preferred order of binding for IF1, IF2, and fMet-$tRNA^{fMet}$, if indeed these components follow an ordered assembly pathway. It was shown in sections 4.3.6 and 4.3.7 that IF1 is relatively insensitive to the presence of fMet-$tRNA^{fMet}$, but is significantly stabilized by IF2. If the binding of IF2 and fMet-$tRNA^{fMet}$ is ordered, then this may be observed by changes in IF1's 30S subunit binding kinetics during a titration of fMet-$tRNA^{fMet}$ or IF2. Future experiments can also be done to test the effect of the identity of the tRNA on the stability of IF1 on the 30S IC. Phe-$tRNA^{Phe}$, for example, has been shown to destabilize IF2 [40], which would likely lead to reversible IF1 binding to the 30S IC.

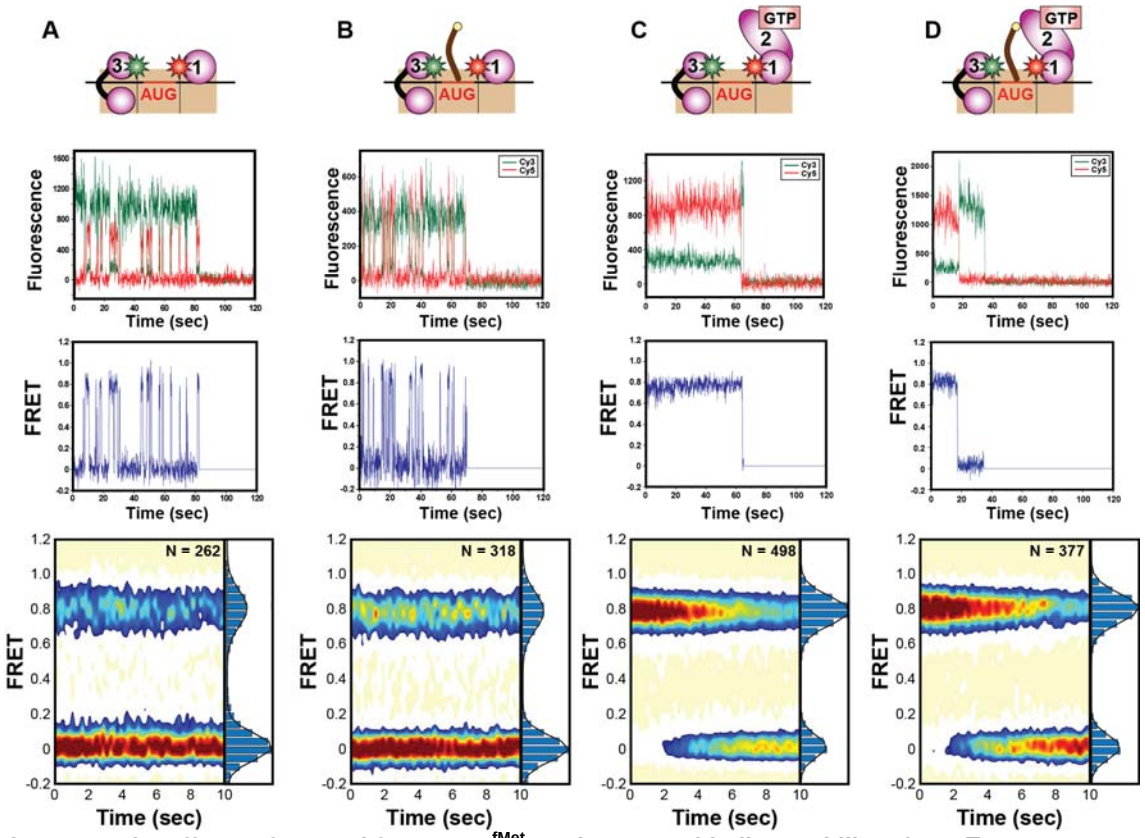

Figure 4.9 The effects of IF2 and fMet-tRNA^{fMet} on the 30S IC binding stability of IF1. Top row: cartoon representation of the contents of each 30S IC (**A** = 30SIC$^{-tRNA}_{-IF2}$; **B** = 30SIC$^{fMet}_{-IF2}$; **C** = 30S IC^{-tRNA}; **D** = 30S IC^{fMet}). The red star indicates Cy5 and the green star indicates Cy3. Second row: Representative fluorescence versus time trajectories. Third row: The corresponding FRET trajectories calculated as FRET = $I_{Cy5}/(I_{Cy5}+I_{Cy3})$. Bottom row: Time-evolution of population FRET histograms. The normalized population histogram to the right of each 2D FRET histogram is plotted with the first 100 frames (10 seconds) of data to minimize contamination due to photobleaching. The number of trajectories in each plot is indicated in the upper right hand corner ("N"). The signal that grows in at zero FRET in the 2D histogram (in panels C and D) is due to fluorophore photobleaching. The signal at zero FRET in panels A and B indicates either the absence of IF1(Cy5) from the 30S IC (an "off state" dwell) or photobleaching. For all three 30S ICs, IF1(Cy5) was included in solution at 50 nM concentration.

4.3.9 The IF1(Cy5)-IF3(Cy3) smFRET signal employed here is insensitive to the identity of the P-site start codon

In line with the insensitivity of the IF1-IF3 FRET signal to the presence of the tRNA, this signal is also unable to distinguish between 30S ICs assembled on mRNAs carrying a canonical AUG versus the non-canonical AUU start codon (see Figure 4.10), though a subset of these experiments need to be reproduced and all need to be triplicated. It is expected that sensitivity to the start codon would arise only when tRNA is present in the 30S IC, and that it is the stable codon-anticodon interaction, or lack thereof, that may be detectable by the initiation factors or

30S subunit itself. Any difference in the fMet-tRNA^{fMet}-AUG vs. -AUU interaction, however, is undetected by this IF1-IF3 smFRET signal in either 30SIC^{fMet,AUU} or 30SIC$^{fMet,AUU}_{-IF2}$ when compared with their AUG counterparts (Figure 4.10), which correlates well with the signal's insensitivity to the presence of the tRNA. Other 30S IC smFRET signals, including the IF3(Cy3-Cy5) signal discussed in Chapter 2 and an IF2(Cy5)-tRNA(Cy3) signal developed by Dr. Jiangning Wang [40], are highly sensitive to differences in the P-site codon, however, reinforcing the importance of cautious data interpretation and the limitations of a single distance constraint from FRET. This limitation can easily be overcome by designing multiple smFRET signals to probe different aspects of the system of interest.

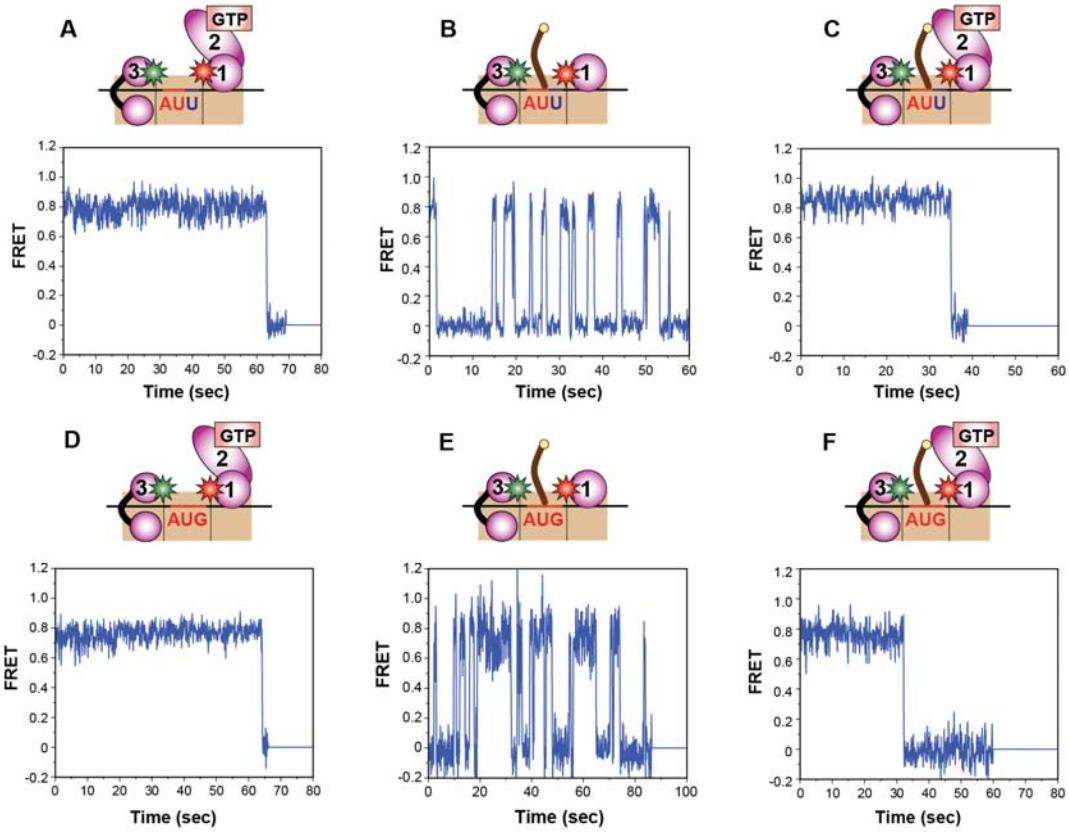

Figure 4.10 IF1(Cy5)-IF3(Cy3) smFRET data collected on 30S ICs containing mRNA with the non-canonical AUU start codon is comparable to data collected on 30S ICs containing mRNA with the canonical AUG start codon. FRET versus time trajectories. For all six 30S ICs, IF1(Cy5) was included at 50 nM concentration. **(A-C)** smFRET data collected on 30S ICs formed with mRNA containing the non-canonical AUU start codon. **(D-E)** smFRET data collected on 30S ICs formed with mRNA containing the non-canonical AUG start codon. The cartoons above each FRET trajectory indicate the contents of the 30S IC.

4.4 Conclusions and Future Directions

The development of an smFRET signal between IF1 and IF3, and preliminary investigations into IF1's 30S subunit binding kinetics, mark the first steps toward a detailed kinetic scheme for the 30S IC assembly pathway. The data presented here reveal that IF1 binds reversibly to the 30S subunit in the presence of just IF3, as well as in the presence of IF3 and initiator tRNA. IF1 is stabilized on the 30S subunit by IF2, and future experiments will be performed to determine the reversibility of IF1's binding to these 30S ICs. The results here also show the insensitivity of IF1's binding kinetics to the presence of fMet-tRNA^{fMet} and the start codon on 30S ICs.

Also notable from this work is the lack of a FRET signal between IF1(Cy5) and IF3-NTD-(Cy3), and the presence of a high FRET efficiency (0.80 ± 0.09) signal between IF1(Cy5) and IF3-CTD-(Cy3). This result lends support to the placement of IF3's CTD at the P site, and IF3's NTD distal from IF1's binding site at the A site, perhaps closer to the E site as proposed by Dallas and Noller [44] (see Chapter 2). Notably, this result comes only from a test for IF1(Cy5) FRET to IF3's CTD versus NTD on $30SIC^{-tRNA}_{-IF2}$, and it remains to be seen if an IF1-IF3(NTD) FRET signal is detectable on other 30S ICs if IF3 undergoes an interdomain reconfiguration (see Figure 2.8).

Future experiments will be designed to probe the timing of IF1's dissociation from the ribosome during the initiation pathway, especially with respect to the timing of 50S subunit joining and 70S IC assembly. These experiments could resolve a decades-old debate about the timing of IF1's release with respect to 50S subunit joining and IF2 release [33, 43, 59]. Technically, these experiments are somewhat challenging since they will involve a third fluorescent dye on the 50S subunit in order to observe the timing of 50S subunit joining in a real-time subunit delivery experiment (see section 3.2.3). Labeling of the 50S subunit is straightforward, however, as shown by previous and ongoing work in our research group and others' [60-63]. It may also be possible to develop an smFRET signal between IF1 and the 50S subunit that would eliminate the need for a three-fluorophore experiment. Other interesting experiments could be performed to probe the

proposed anti-subunit association role of IF1 [30, 34, 35], especially through the use of 16S rRNA mutants displaying unusual subunit joining kinetics, such as the helix 44 mutants A1413C, A1410G, or A1408G [31].

4.5 References

1. Simonetti, A., S. Marzi, A.G. Myasnikov, A. Fabbretti, M. Yusupov, C.O. Gualerzi, and B.P. Klaholz, *Structure of the 30S translation initiation complex.* Nature, 2008. **455**(7211): p. 416-20.

2. Gualerzi, C.O. and C.L. Pon, *Initiation of mRNA translation in prokaryotes.* Biochemistry, 1990. **29**(25): p. 5881-9.

3. Boelens, R. and C.O. Gualerzi, *Structure and function of bacterial initiation factors.* Curr Protein Pept Sci, 2002. **3**(1): p. 107-19.

4. Laursen, B.S., H.P. Sorensen, K.K. Mortensen, and H.U. Sperling-Petersen, *Initiation of protein synthesis in bacteria.* Microbiol Mol Biol Rev, 2005. **69**(1): p. 101-23.

5. Schmeing, T.M. and V. Ramakrishnan, *What recent ribosome structures have revealed about the mechanism of translation.* Nature, 2009. **461**(7268): p. 1234-42.

6. Mathews, M.B., N. Sonenberg, and J.W. Hershey, *Origins and Principles of Translational Control*, in *Translational Control in Biology and Medicine*, M.B. Mathews, N. Sonenberg, and J.W. Hershey, Editors. 2007, Cold Spring Harbor Laboratory Press. p. 1-40.

7. Marzi, S., P. Fechter, C. Chevalier, P. Romby, and T. Geissmann, *RNA switches regulate initiation of translation in bacteria.* Biol Chem, 2008. **389**(5): p. 585-98.

8. Schlax, P.J. and D.J. Worhunsky, *Translational repression mechanisms in prokaryotes.* Mol Microbiol, 2003. **48**(5): p. 1157-69.

9. Lee, K., C.A. Holland-Staley, and P.R. Cunningham, *Genetic analysis of the Shine-Dalgarno interaction: selection of alternative functional mRNA-rRNA combinations.* RNA, 1996. **2**(12): p. 1270-85.

10. Komarova, A.V., L.S. Tchufistova, E.V. Supina, and I.V. Boni, *Protein S1 counteracts the inhibitory effect of the extended Shine-Dalgarno sequence on translation.* RNA, 2002. **8**(9): p. 1137-47.

11. Skorski, P., P. Leroy, O. Fayet, M. Dreyfus, and S. Hermann-Le Denmat, *The highly efficient translation initiation region from the Escherichia coli rpsA gene lacks a shine-dalgarno element.* J Bacteriol, 2006. **188**(17): p. 6277-85.

12. Subramanian, A.R., E.Z. Ron, and B.D. Davis, *A factor required for ribosome dissociation in Escherichia coli.* Proc Natl Acad Sci U S A, 1968. **61**(2): p. 761-7.

13. Subramanian, A.R. and B.D. Davis, *Activity of initiation factor F3 in dissociating Escherichia coli ribosomes.* Nature, 1970. **228**(5278): p. 1273-5.

14. Miall, S.H. and T. Tamaoki, *Dissociation of Escherichia coli ribosomes. Role of initiation factors.* Biochemistry, 1972. **11**(25): p. 4826-30.

15. Gottleib, M., B.D. Davis, and R.C. Thompson, *Mechanism of dissociation of ribosomes of Escherichia coli by initiation factor IF-3.* Proc Natl Acad Sci U S A, 1975. **72**(11): p. 4238-42.

16. Chaires, J.B., C. Pande, and A. Wishnia, *The effect of initiation factor IF-3 on Escherichia coli ribosomal subunit association kinetics.* J Biol Chem, 1981. **256**(13): p. 6600-7.

17. Weiel, J. and J.W. Hershey, *Fluorescence polarization studies of the binding of fluorescein-labeled initiation factor IF3 to 30 S ribosomal subunits from Escherichia coli.* FEBS Lett, 1978. **87**(1): p. 103-6.

18. Gualerzi, C.O., A. Fabbretti, L. Brandi, P. Milon, and C.L. Pon, *Role of the Initiation Factors in mRNA Start Site Selection and fMet-tRNA Recruitment by Bacterial Ribosomes.* Israel Journal of Chemistry, 2010. **50**(1): p. 80-94.

19. Pon, C.L. and C.O. Gualerzi, *Mechanism of protein biosynthesis in prokaryotic cells. Effect of initiation factor IF1 on the initial rate of 30 S initiation complex formation.* FEBS Lett, 1984. **175**(2): p. 203-7.

20. Milon, P., M. Carotti, A.L. Konevega, W. Wintermeyer, M.V. Rodnina, and C.O. Gualerzi, *The ribosome-bound initiation factor 2 recruits initiator tRNA to the 30S initiation complex.* EMBO Rep, 2010. **11**(4): p. 312-6.

21. Antoun, A., M.Y. Pavlov, M. Lovmar, and M. Ehrenberg, *How initiation factors tune the rate of initiation of protein synthesis in bacteria.* EMBO J, 2006. **25**(11): p. 2539-50.

22. Tomsic, J., L.A. Vitali, T. Daviter, A. Savelsbergh, R. Spurio, P. Striebeck, W. Wintermeyer, M.V. Rodnina, and C.O. Gualerzi, *Late events of translation initiation in bacteria: a kinetic analysis.* EMBO J, 2000. **19**(9): p. 2127-36.

23. Gualerzi, C., G. Risuleo, and C.L. Pon, *Initial rate kinetic analysis of the mechanism of initiation complex formation and the role of initiation factor IF-3.* Biochemistry, 1977. **16**(8): p. 1684-9.

24. Gold, L., G. Stormo, and R. Saunders, *Escherichia coli translational initiation factor IF3: a unique case of translational regulation.* Proc Natl Acad Sci U S A, 1984. **81**(22): p. 7061-5.

25. Pon, C.L. and C.O. Gualerzi, *Mechanism of translational initiation in prokaryotes. IF3 is released from ribosomes during and not before 70 S initiation complex formation.* FEBS Lett, 1986. **195**(1-2): p. 215-9.

26. La Teana, A., C.L. Pon, and C.O. Gualerzi, *Translation of mRNAs with degenerate initiation triplet AUU displays high initiation factor 2 dependence and is subject to initiation factor 3 repression.* Proc Natl Acad Sci U S A, 1993. **90**(9): p. 4161-5.

27. Grigoriadou, C., S. Marzi, S. Kirillov, C.O. Gualerzi, and B.S. Cooperman, *A quantitative kinetic scheme for 70 S translation initiation complex formation.* J Mol Biol, 2007. **373**(3): p. 562-72.

28. Zucker, F.H. and J.W. Hershey, *Binding of Escherichia coli protein synthesis initiation factor IF1 to 30S ribosomal subunits measured by fluorescence polarization.* Biochemistry, 1986. **25**(12): p. 3682-90.

29. Hoskins, A.A., L.J. Friedman, S.S. Gallagher, D.J. Crawford, E.G. Anderson, R. Wombacher, N. Ramirez, V.W. Cornish, J. Gelles, and M.J. Moore, *Ordered and dynamic assembly of single spliceosomes.* Science, 2011. **331**(6022): p. 1289-95.

30. Milon, P., A.L. Konevega, C.O. Gualerzi, and M.V. Rodnina, *Kinetic checkpoint at a late step in translation initiation.* Mol Cell, 2008. **30**(6): p. 712-20.

31. Qin, D. and K. Fredrick, *Control of translation initiation involves a factor-induced rearrangement of helix 44 of 16S ribosomal RNA.* Mol Microbiol, 2009. **71**(5): p. 1239-49.

32. Antoun, A., M.Y. Pavlov, M. Lovmar, and M. Ehrenberg, *How initiation factors maximize the accuracy of tRNA selection in initiation of bacterial protein synthesis.* Mol Cell, 2006. **23**(2): p. 183-93.

33. Stringer, E.A., P. Sarkar, and U. Maitra, *Function of initiation factor 1 in the binding and release of initiation factor 2 from ribosomal initiation complexes in Escherichia coli.* J Biol Chem, 1977. **252**(5): p. 1739-44.

34. Belotserkovsky, J.M., E.R. Dabbs, and L.A. Isaksson, *Mutations in 16S ribosomal RNA that suppress a cold sensitive initiation factor 1 affect ribosomal subunit association.* FEBS J, 2011.

35. Carter, A.P., W.M. Clemons, Jr., D.E. Brodersen, R.J. Morgan-Warren, T. Hartsch, B.T. Wimberly, and V. Ramakrishnan, *Crystal structure of an initiation factor bound to the 30S ribosomal subunit.* Science, 2001. **291**(5503): p. 498-501.

36. Allen, G.S., A. Zavialov, R. Gursky, M. Ehrenberg, and J. Frank, *The cryo-EM structure of a translation initiation complex from Escherichia coli.* Cell, 2005. **121**(5): p. 703-12.

37. Allen, G.S. and J. Frank, *Structural insights on the translation initiation complex: ghosts of a universal initiation complex.* Mol Microbiol, 2007. **63**(4): p. 941-50.

38. Fei, J., J. Wang, S.H. Sternberg, D.D. MacDougall, M.M. Elvekrog, D.K. Pulukkunat, M.T. Englander, and R.L. Gonzalez, Jr., *A highly purified, fluorescently labeled in vitro translation system for single-molecule studies of protein synthesis.* Methods Enzymol, 2010. **472**: p. 221-59.

39. Myasnikov, A.G., S. Marzi, A. Simonetti, A.M. Giuliodori, C.O. Gualerzi, G. Yusupova, M. Yusupov, and B.P. Klaholz, *Conformational transition of initiation factor 2 from the GTP- to GDP-bound state visualized on the ribosome.* Nat Struct Mol Biol, 2005. **12**(12): p. 1145-9.

40. Wang, J., *Regulation of IF2 Binding Kinetics and 30S IC Conformational Dynamics during Translation Initiation*, 2010, Columbia University.

41. Croitoru, V., M. Bucheli-Witschel, P. Hagg, F. Abdulkarim, and L.A. Isaksson, *Generation and characterization of functional mutants in the translation initiation factor IF1 of Escherichia coli.* Eur J Biochem, 2004. **271**(3): p. 534-44.

42. Spurio, R., M. Paci, R.T. Pawlik, A. La Teana, B.V. DiGiacco, C.L. Pon, and C.O. Gualerzi, *Site-directed mutagenesis and NMR spectroscopic approaches to the elucidation of the structure-function relationships in translation initiation factors IF1 and IF3.* Biochimie, 1991. **73**(7-8): p. 1001-6.

43. Gualerzi, C.O., R. Spurio, A. La Teana, R. Calogero, B. Celano, and C.L. Pon, *Site-directed mutagenesis of Escherichia coli translation initiation factor IF1. Identification of the amino acid involved in its ribosomal binding and recycling.* Protein Eng, 1989. **3**(2): p. 133-8.

44. Dallas, A. and H.F. Noller, *Interaction of translation initiation factor 3 with the 30S ribosomal subunit.* Mol Cell, 2001. **8**(4): p. 855-64.

45. DeLano, W., *The PyMOL Molecular Graphics System, Version 1.3, Schrödinger, LLC.* 2008.

46. Bradford, M.M., *A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding.* Anal Biochem, 1976. **72**: p. 248-54.

47. Hartz, D., D.S. McPheeters, and L. Gold, *Selection of the initiator tRNA by Escherichia coli initiation factors.* Genes Dev, 1989. **3**(12A): p. 1899-912.

48. McCutcheon, J.P., R.K. Agrawal, S.M. Philips, R.A. Grassucci, S.E. Gerchman, W.M. Clemons, Jr., V. Ramakrishnan, and J. Frank, *Location of translational initiation factor IF3 on the small ribosomal subunit.* Proc Natl Acad Sci U S A, 1999. **96**(8): p. 4301-6.

49. Bastiaens, P.I. and T.M. Jovin, *Microspectroscopic imaging tracks the intracellular processing of a signal transduction protein: fluorescent-labeled protein kinase C beta I.* Proc Natl Acad Sci U S A, 1996. **93**(16): p. 8407-12.

50. Hohng, S., C. Joo, and T. Ha, *Single-molecule three-color FRET.* Biophys J, 2004. **87**(2): p. 1328-37.

51. Lomakin, I.B., N.E. Shirokikh, M.M. Yusupov, C.U. Hellen, and T.V. Pestova, *The fidelity of translation initiation: reciprocal activities of eIF1, IF3 and YciH.* EMBO J, 2006. **25**(1): p. 196-210.

52. Rabl, J., M. Leibundgut, S.F. Ataide, A. Haag, and N. Ban, *Crystal structure of the eukaryotic 40S ribosomal subunit in complex with initiation factor 1.* Science, 2011. **331**(6018): p. 730-6.

53. Roy, R., S. Hohng, and T. Ha, *A practical guide to single-molecule FRET.* Nat Methods, 2008. **5**(6): p. 507-16.

54. Rasnik, I., S.A. McKinney, and T. Ha, *Nonblinking and longlasting single-molecule fluorescence imaging.* Nature Methods, 2006. **3**(11): p. 891-893.

55. Bronson, J.E., J. Fei, J.M. Hofman, R.L. Gonzalez, Jr., and C.H. Wiggins, *Learning rates and states from biophysical time series: a Bayesian approach to model selection and single-molecule FRET data.* Biophys J, 2009. **97**(12): p. 3196-205.

56. Choi, S.K., D.S. Olsen, A. Roll-Mecak, A. Martung, K.L. Remo, S.K. Burley, A.G. Hinnebusch, and T.E. Dever, *Physical and functional interaction between the eukaryotic orthologs of prokaryotic translation initiation factors IF1 and IF2.* Mol Cell Biol, 2000. **20**(19): p. 7183-91.

57. Gaur, R., D. Grasso, P.P. Datta, P.D. Krishna, G. Das, A. Spencer, R.K. Agrawal, L. Spremulli, and U. Varshney, *A single mammalian mitochondrial translation initiation factor functionally replaces two bacterial factors.* Mol Cell, 2008. **29**(2): p. 180-90.

58. Yassin, A.S., M.E. Haque, P.P. Datta, K. Elmore, N.K. Banavali, L.L. Spremulli, and R.K. Agrawal, *Insertion domain within mammalian mitochondrial translation initiation factor 2 serves the role of eubacterial initiation factor 1.* Proc Natl Acad Sci U S A, 2011. **108**(10): p. 3918-23.

59. Celano, B., R.T. Pawlik, and C.O. Gualerzi, *Interaction of Escherichia coli translation-initiation factor IF-1 with ribosomes.* Eur J Biochem, 1988. **178**(2): p. 351-5.

60. Fei, J., P. Kosuri, D.D. MacDougall, and R.L. Gonzalez, Jr., *Coupling of ribosomal L1 stalk and tRNA dynamics during translation elongation.* Mol Cell, 2008. **30**(3): p. 348-59.

61. Fei, J., J.E. Bronson, J.M. Hofman, R.L. Srinivas, C.H. Wiggins, and R.L. Gonzalez, Jr., *Allosteric collaboration between elongation factor G and the ribosomal L1 stalk directs tRNA movements during translation.* Proc Natl Acad Sci U S A, 2009. **106**(37): p. 15702-7.

62. Marshall, R.A., C.E. Aitken, and J.D. Puglisi, *GTP hydrolysis by IF2 guides progression of the ribosome into elongation.* Mol Cell, 2009. **35**(1): p. 37-47.

63. Cornish, P.V., D.N. Ermolenko, H.F. Noller, and T. Ha, *Spontaneous intersubunit rotation in single ribosomes.* Mol Cell, 2008. **30**(5): p. 578-88.

Chapter 5
Materials and Methods

Data were collected and analyzed for this dissertation following numerous established methods, as well as methods that were developed and/or optimized along the way. This chapter contains detailed protocols for all techniques performed in the collection and analysis of this dissertation's results. Sections 5.1 and 5.2 contain detailed protocols for the preparation of most of the components within the *in vitro* translation system employed for the work in this dissertation. The methods used in the preparation and purification of ribosomes and ribosomal subunits can be found in Dr. Jiangning Wang's dissertation [1]. Details on all biochemical activity assays performed are found in sections 5.3 and 5.4. smFRET data collection and analysis procedures are included in sections 5.5 and 5.6.

5.1 Preparation of tRNA reagents

The following sections describe the aminoacylation and formylation of $tRNA^{fMet}$, and the aminoacylation of $tRNA^{Phe}$. The other tRNA reagents used in the work described in this dissertation, Lys-$tRNA^{Lys}$ and fMet-(Cy3)$tRNA^{fMet}$, were generously prepared by Drs. Michael Englander and Jiangning Wang, respectively. Please refer to their dissertations [1, 2] for details on the preparation of these reagents.

5.1.1 Aminoacylation and formylation of $tRNA^{fMet}$

5.1.1.1 Preparation of ^{10}N-formyltetrahydrofolate from folinic acid

The formyl donor substrate for formylmethionyl-tRNA formyltransferase, ^{10}N-formyltetrahydrofolate, is chemically prepared starting from the calcium salt of folinic acid (Acros Organics) as previously described [3]. The following procedure was originally obtained from Prof. Ruben Gonzalez's laboratory notebook titled "tRNA Labeling" from his postdoctoral research work at Stanford University. The reaction chemistry is depicted in Figure 5.1.

1. Obtain folinic acid (calcium salt, 100 mg) from Acros Organics. CAS #: 1492-18-8. Calcium salt pentahydrate 95%-105%

2. Dissolve 25 mg of folinic acid in 2 mL of 50 mM BME (prepare BME stock as follows: 6.99 μL BME + 1993.01 μL nanopure water). Run reaction in a 15 mL Falcon tube.

3. Treat the solution with 220 μL of 1 M HCl (prepare 1 M HCl stock by diluting 12.1 M stock). Precipitated material will need to be resuspended by warming in 37°C water bath.

4. After 3 hours at room temperature, cyclization to 5:10-methenyltetrahydrofolate (MTHF) should be complete. Cyclization is monitored via absorbance at 355 nm. $\varepsilon_{355\ nm}$ = 25 x 10^3 $M^{-1}cm^{-1}$ for 5:10-MTHF (see Figure 5.2).

5. Adjust concentration of 5:10-MTHF to 15 mM using 0.1 M HCl as a diluent and the above extinction coefficient.

6. Prepare 200 μL aliquots and store at -20°C.

7. Before use, warm aliquot to room temperature, neutralize with 20 μmol KOH and 10 μmol Tris-HCl, pH_{rt}=7.8 (10 μL 5:10 MTHF + 1 μL 1 M KOH + 0.5 μL Tris-HCl) and incubate at room temperature for 15 minutes to generate ^{10}N-formyltetrahydrofolate.

Reaction chemistry:

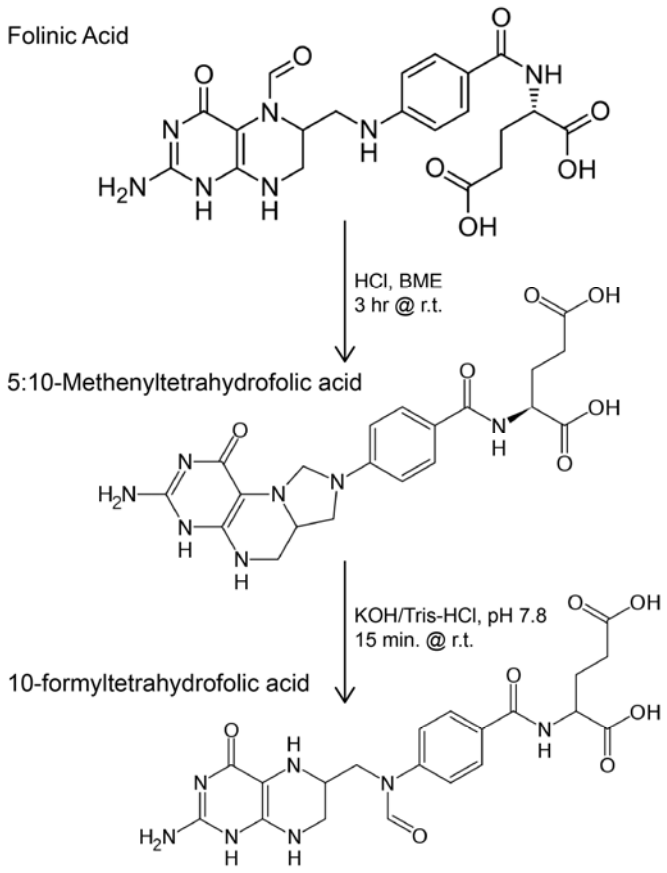

Figure 5.1 Synthetic pathway for the preparation of 10-formyltetrahydrofolic acid from folinic acid.

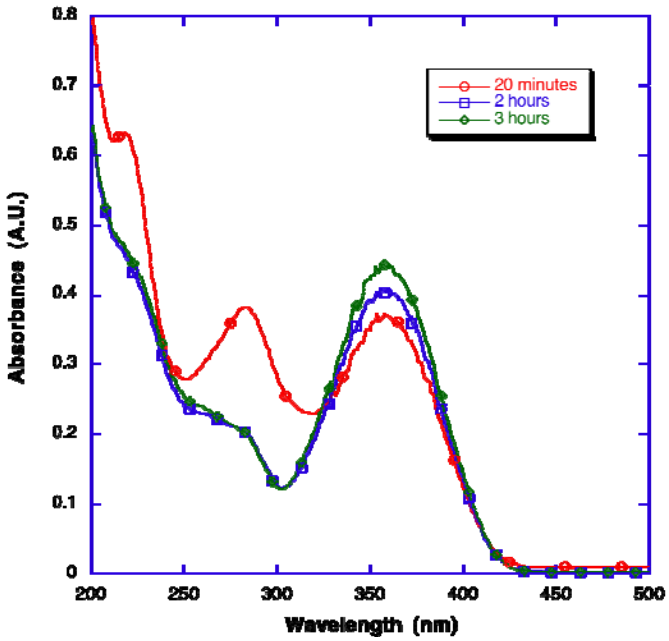

Figure 5.2 Preparation of 5:10-methenyltetrahydrofolic acid. UV-Vis spectra of the conversion of folinic acid to 5:10-methenyltetrahydrofolic acid in the presence of BME and HCl. Peak at 355 nm increases over time. Reaction was complete at 3 hours.

5.1.1.2 Aminoacylation of $tRNA^{fMet}$

Methionyl tRNA synthetase and formylmethionyl-tRNA formyltransferase, both from *E. coli*, were prepared as previously described [4].

Procedure

1. Prepare 5 mL of 5x Aminoacylation/Formylation (AF) buffer using the following recipe:
   - 625 µL 1 M Tris-HCl, $pH_{37°C}$=7.5
   - 35.7 µL 4.9 M $MgCl_2$
   - 1875 µL 2 M KCl
   - 5 µL 0.5 M EDTA
   - 25 µL 1 M DTT
   - 1250 µL 50 mM ATP
   - 1184 µL nanopure water

2. Prepare ^{10}N-formyltetrahydrofolate as described in section 5.1.1.1.

3. Prepare a mixture with the following components, adding the enzymes last and adjusting the volume with nanopure water:
   - 20 µM $tRNA^{fMet}$
   - 80 µM methionine
   - 300 µM ^{10}N-formyltetrahydrofolate
   - 1x AF buffer
   - 0.02 µM methionyl tRNA synthetase
   - 0.20 µM formylmethionyl-tRNA formyltransferase

4. Incubate the tRNA reaction mixture for 10 minutes at 37ºC. Quench the reaction by adding 0.1x total reaction volume of 3 M NaOAc, pH 5.2.
   *Note: Steps 5 and 6 are performed in a cold (4ºC) room.

5. Extract the mixture twice with cold (4ºC) phenol (1x reaction volume), and twice with 1x reaction volume of cold chloroform.

6. Add 3x reaction volume of -20ºC 100% ethanol. Mix well by pipetting.

7. Incubate the reaction mixture at -80ºC for at least one hour.

8. Centrifuge the mixture for 20 minutes at 18,000 x g at 4ºC.

9. Carefully remove the supernatant, taking care not to disturb the white pellet.

10. Resuspend the pellet in 25-50 µL of 10 mM KOAc (pH 5) and apply to a Micro Bio-Spin 6 gel filtration spin column (Biorad) that has been equilibrated with 10 mM KOAc (pH 5).

11. Immediately prepare aliquots (3-5 µL), flash freeze with liquid nitrogen, and store at -80ºC.

12. Use a 1:100 dilution of a tRNA aliquot to calculate the final tRNA concentration by measuring absorbance at 260 nm with a UV-Vis spectrometer. The extinction coefficient for $tRNA^{fMet}$ is 726,700 $M^{-1}cm^{-1}$ at 260 nm.

5.1.1.3 Assessing the aminoacylation and formylation yields for fMet-$tRNA^{fMet}$

The aminoacylation and formylation yields are assessed by hydrophobic interaction chromatography (HIC). The HIC buffers are:

tRNA HIC Buffer A: 1.7 M NH_4SO_4
10 mM NH_4OAc, pH 6.3*

tRNA HIC Buffer B: 10% Methanol
10 mM NH_4OAc, pH 6.3*

*Note that this is the pH of the 1 M NH_4OAc stock solution, not the final HIC Buffer pH. The final solution does not need to be pH adjusted.

The tRNA HIC buffers should be filtered and chilled prior to use. Dilute 0.1 nmoles of fMet-$tRNA^{fMet}$ (one or two aliquots, depending on the final concentration) with 60-70 µL of tRNA HIC Buffer A and inject the sample onto the HIC column that has been equilibrated with tRNA HIC Buffer A. Elute the tRNA over a linear gradient of 0-100% tRNA HIC Buffer B over 25 column volumes. The three tRNA species ($tRNA^{fMet}$, Met-$tRNA^{fMet}$, and fMet-$tRNA^{fMet}$) elute from the HIC

column sequentially due to their increasing hydrophobicities (see Figure 5.3). Typically, fMet-tRNA^{fMet} is obtained with a >90% yield, as assessed by chromatogram peak areas.

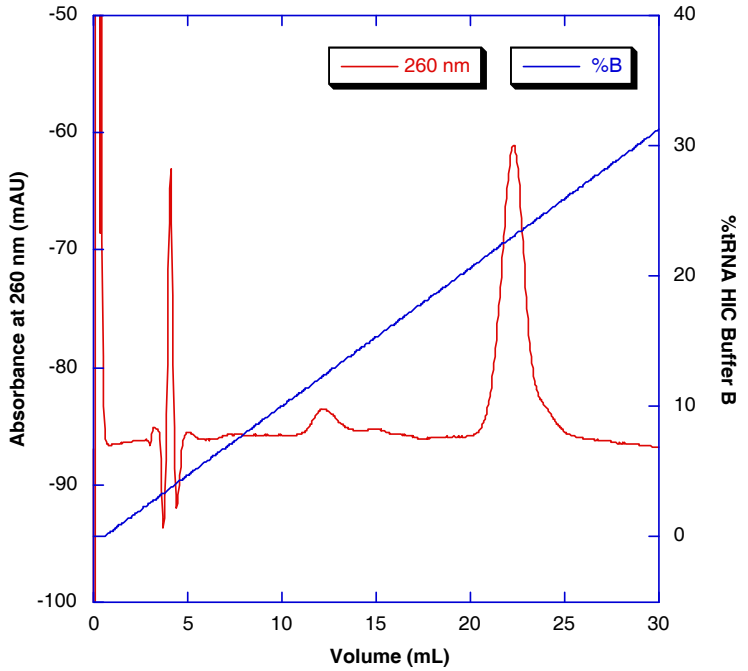

Figure 5.3 Assessment of fMet-tRNA^{fMet} aminoacylation and formylation yield. HIC chromatogram of 0.1 nmol fMet-tRNA^{fMet} eluted with a 0-100% Buffer B over 25 column volumes gradient. The major peak at ~22 mL is fMet-tRNA^{fMet} and the minor peaks at ~12 and 15 mL are tRNA^{fMet} and Met-tRNA^{fMet}, respectively.

5.1.2 Aminoacylation of tRNA^{Phe}

Phenylalanyl tRNA synthetase from *E. coli* was prepared as previously described [4].

Procedure

1. Prepare a mixture with the following components, adding the synthetase last and adjusting the volume with nanopure water:
   - 15 µM tRNA^{Phe}
   - 55 µM phenylalanine
   - 200 mM Tris-HCl, $pH_{37°C}$=7.5
   - 15 mM $MgCl_2$
   - 25 mM KCl
   - 2 mM BME
   - 5 mM ATP
   - 10 mM phosphoenolpyruvate (PEP)
   - 30 U mL^{-1} pyruvate kinase
   - 0.75 µM phenylalanyl tRNA synthetase

2. Incubate the tRNA reaction mixture for 10 minutes at 37ºC. Quench the reaction by adding 0.1x total reaction volume of 3 M NaOAc, pH 5.2.
   *Note: Steps 3 and 4 are performed in a cold (4ºC) room.

3. Extract the mixture twice with cold (4ºC) phenol (1x reaction volume), and twice with 1x reaction volume of cold chloroform.

4. Add 3x reaction volume of -20ºC 100% ethanol. Mix well by pipetting.

5. Incubate the reaction mixture at -80ºC for at least one hour.

6. Centrifuge the mixture for 20 minutes at 18,000 x g at 4ºC.

7. Carefully remove the supernatant, taking care not to disturb the white pellet.

8. Resuspend the pellet in 25-50 µL of 10 mM KOAc (pH 5) and apply to a Micro Bio-Spin 6 gel filtration spin column that was equilibrated with 10 mM KOAc (pH 5).

9. Immediately prepare aliquots (3-5 µL), flash freeze with liquid nitrogen, and store at -80ºC.

10. Use a 1:100 dilution of the tRNA aliquot to calculate the final tRNA concentration by measuring absorbance at 260 nm with a UV-Vis spectrometer. The extinction coefficient for $tRNA^{Phe}$ is 760,000 $M^{-1}cm^{-1}$ at 260 nm.

11. Assess the extent of aminoacylation by HIC chromatography as described in section 5.1.1.3. See Figure 5.4 for a sample chromatogram.

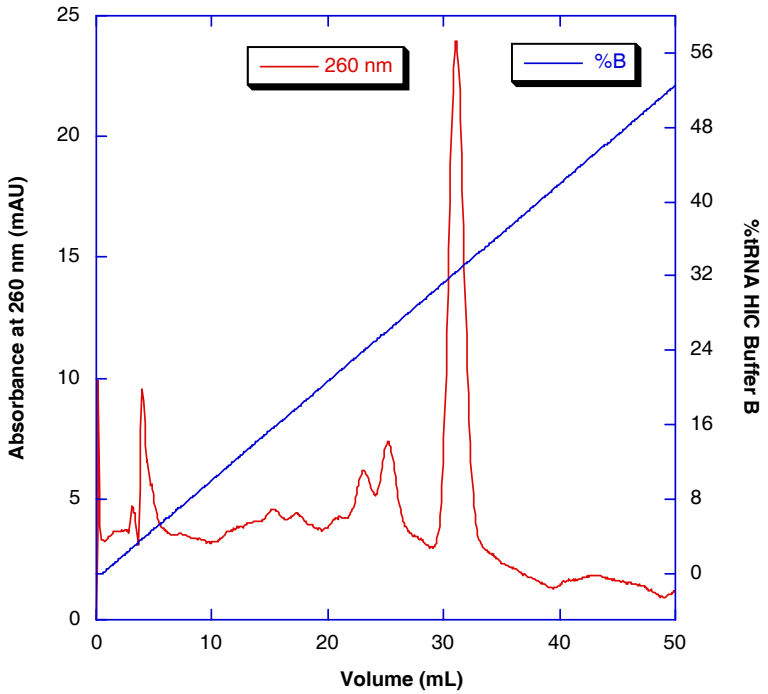

Figure 5.4 Assessment of Phe-$tRNA^{Phe}$ aminoacylation. HIC chromatogram of 0.1 nmol Phe-$tRNA^{Phe}$ eluted with a 0-100% Buffer B over 25 column volumes gradient. The major peak at ~31 mL is Phe-$tRNA^{Phe}$.

5.2 Preparation and purification of translation initiation factors

Translation initiation factors IF1, IF2, and IF3 were all overexpressed, purified, and biochemically tested. IF2γ was prepared by either Dr. Jiangning Wang or Mr. Daniel MacDougall. Please refer to Dr. Wang's dissertation [1] for detailed purification and biochemical testing protocols. The purification and biochemical testing of IF1 and IF3 are described in the following sections.

5.2.1 Preparation and purification of IF1

5.2.1.1 IF1 amino acid sequence

```
1         11        21        31       41         51        61
MAKEDNIEMQGTVLETLPNTMFRVELENGHVVTAHISGKMRKNYIRILTGDKVTVELTPYDLSKGRIVFRS
71
R
```

5.2.1.2 IF1 purification

Wild-type *E. coli* IF1 was originally cloned into a pProEX-HTb plasmid, transformed into DH5α/pRS3559 cells, and is identified in the laboratory strain collection as RG21. I transformed the plasmid into BL21-DE3 cells for optimal overexpression. Information on all cell strains can be found in Appendix D. A typical IF1 protein yield is 0.5 mg L^{-1}.

Overnight culture

Innoculate two 5 mL starter cultures of Terrific Broth (TB) (Difco) media with single colonies from an agar plate freshly streaked with the glycerol stock (MME Strain #2) (see Appendix D). α-Carboxybenzylpenicillin (carbenicillin) should be present in the starter cultures at 100 μg/mL. Grow with shaking at 250 rpm and 37°C. Minimize growth period to 12-14 hours.

Cell growth and protein overexpression

Innoculate two 1 L cultures of TB media with the overnight cultures and 100 μg/mL carbenicillin. Grow with shaking at 250 rpm and 37°C. Check the optical density (OD) at 600 nm every 30 minutes using a 1 mL sample from each culture and disposable 1.5 mL, 1 cm pathlength cuvettes. Once the OD_{600} reaches 0.6-0.8, save a 1 mL fraction of each culture for analysis and

induce protein overexpression by the addition of isopropyl β-D-1-thiogalactopyranoside (IPTG) to a final concentration of 1 mM to each 1 L culture. Continue to grow the cultures with shaking at 250 rpm and 37°C for 3-4 additional hours.

Cell harvesting of His_6-tagged IF1

Terminate cell growth by centrifuging the cultures in 0.5 L centrifuge bottles in a JA-10 rotor at 7,500 rpm for 15 minutes at 4°C. Store the cell pellet at -20°C, and properly dispose of the supernatant after treatment with bleach.

Purification

Purification should continue without interruption once the procedure is started; the procedure requires four days. This purification procedure is modified from a protocol developed in the research group of Joseph Puglisi (Stanford). The buffers should be prepared the day of use and chilled to 4°C before being pH adjusted to 7.5. All purification steps should be done at 4°C.

Required buffers:

Lysis and Equilibration Buffer
10 mM Tris-HCl, $pH_{4°C}$=7.5
60 mM NH_4Cl
10 mM $MgCl_2$
5 mM BME
0.1 mM phenylmethyl sulfonyl fluoride (PMSF)
10 mM imidazole

Wash Buffer
20 mM Tris-HCl, $pH_{4°C}$=7.5
10 mM $MgCl_2$
5 mM BME
0.1 mM PMSF
0.5 mM Ethylenediaminetetraacetic acid (EDTA)
30 mM imidazole

Elution Buffer
20 mM Tris-HCl, $pH_{4°C}$=7.5
10 mM $MgCl_2$
5 mM BME
0.1 mM PMSF
0.5 mM Ethylenediaminetetraacetic acid (EDTA)
250 mM imidazole

1X TEV Protease Buffer
20 mM Tris-HCl, $pH_{4°C}$=7.5
200 mM NaCl
0.1% Triton-X
2 mM BME

2X Storage Buffer (for protein storage, add equal volume of 100% glycerol to bring concentration to 1x)
20 mM Tris-OAc, $pH_{r.t.}$= 7.0
100 mM KCl
10 mM BME

Labeling Buffer
100 mM Tris-OAc, $pH_{r.t.}$ = 7.0
50 mM KCl

IF1 HIC Buffer A
1M $(NH_4)SO_4$
100 mM Na_2HPO_4, $pH_{r.t.}$=7.0

IF1 HIC Buffer B
100 mM Na_2HPO_4, $pH_{r.t.}$=7.0

Cell resuspension

Thaw the cell pellets on ice and then resuspend with 20-30 mL Lysis and Equilibration Buffer. Save 10 μL of the resuspended cell pellet for gel analysis. Pass suspension through French Press 3-4 times with the pressure gauge set to an internal gauge pressure of 1200 psi. Lysate should be released dropwise and collected in a side-arm flask on ice. Clear the lysate by centrifugation at 20,000 x g in a JA-17 rotor for 30 minutes at 4°C. Save 10 μL samples of the lysate, supernatant, and pellet for gel analysis. Carefully decant the supernatant into a disposable 50 mL Falcon tube and discard the cell pellet. If the lysate is not cleared after 30 minutes of centrifugation, as is often the case with IF1, add 0.5-1 mL of 5% poly(ethyleneimine) to precipitate genomic DNA, and then centrifuge the sample for another 30 minutes at 20,000 x g.

Preparation of 5% poly(ethyleneimine)

Take the tare weight of a 50 mL Falcon tube. Add ~5 mL of poly(ethyleneimine) (Sigma; P3143 – 100 mL; 50% (w/v) in water). Spin down in Sorvall centrifuge for 5 min. Measure the net weight,

and using the known density (1.08 g/mL), calculate the volume. Add enough water to make a final solution that is 5% v/v. Mix the solution on a shaker or rocker for at least an hour to fully mix.

Batch binding of His_6-tagged IF1 to Ni^{2+}-Nitrilotriacetic acid (NTA) resin

Equilibrate 2-3 mL of Ni^{2+}-NTA resin (Qiagen) in a disposable column with 20-30 mL Lysis and Equilibration buffer. Plug the bottom of the column and pour 10 mL of the cleared lysate into the column with the equilibrated Ni-NTA column material. Cap the top of the column and gently invert to mix the resin with the lysate. Then, transfer the suspension to the Falcon tube containing the remaining lysate. Secure the tube to a rotator and mix continuously for 30 minutes to ensure binding of all His_6-tagged IF1 to the Ni^{2+}-NTA resin beads.

First Ni^{2+}-NTA column

Pour the Ni^{2+}-NTA resin and lysate mixture back into the disposable column and collect the flow through. Wash the resin with 25 mL of Lysis and Equilibration buffer and collect the flow through. Then wash the resin with 25 mL of Wash buffer and collect the flow through. Elute the His_6-tagged IF1 protein with 25-30 mL of Elution buffer. Collect the eluate in 2-3 mL fractions. Analyze each fraction as well as the cell pellet, lysate supernatant, column binding flow through, equilibration flow through, wash flow through, and all the eluate fractions with a Tris-tricine gel or 10-20% pre cast Tris-tricine gradient gel (NuSep).

Tris-Tricine gel preparation

2x Tricine sample buffer
2 mL 4x Tris-HCl/SDS pH 6.8
2.4 mL Glycerol (26% final)
0.8 g SDS (8% final)
0.31 g DTT (0.2 M final)
2 mg Coomassie blue (0.02% final)
Add nanopure water to 10 mL and mix
Store 1 mL aliquots at -20°C

4x Tris-HCl/SDS pH 6.8 (0.5 M Tris-HCl containing 0.4% SDS)
6.05 g Tris-base
40 mL nanopure water
Adjust pH to 6.8, then add
0.4 g SDS
Add nanopure water to a final volume of 100 mL
Store at 4°C for up to 1 month

Separating gel solution
49 mL 40% acrylamide/bisacrylamide
50 mL 4x Tris-HCl/SDS, pH 8.45
21.1 mL Glycerol
Nanopure water to 200 mL

Stacking gel solution
9.72 mL 40% acrylamide/bisacrylamide
24.8 mL 4x Tris-HCl/SDS, pH 8.45
Nanopure water to 100 mL

Separating gel
8 mL separating gel solution
80 mL APS
8 μL TEMED

Stacking gel
3 mL stacking gel solution
30 μL APS
3 μL TEMED

*When preparing gel, add water-saturated butanol to the top of the separating gel and leave until polymerization is complete. Then discard and rinse with nanopure water.

10x Anode buffer
121.1 g Tris base (0.2 M final)
500 mL nanopure water
Adjust the pH to 8.9 with HCl

1x Cathode buffer
12.11 g Tris base (0.1 M final)
17.92 g Tricine (0.1 M final)
1 g SDS (0.1% final)
Dilute to 1 L with nanopure water
No pH adjustment

Isopropanol fixing solution
250 mL isopropanol
100 mL glacial acetic acid
Add nanopure water to 1 L

Rapid coomassie staining solution
100 mL glacial acetic acid
0.06 g Coomassie Brilliant Blue G-250
Add nanopure water to 1 L

Heat each sample for 5 min at 95°C. Run gel at constant voltage, 150 V, for 40 minutes. Fix the gel for 15 minutes, and then stain with Coomassie stain for 3 hr, and destain with water or a 10% acetic acid/10% methanol mixture, using a KimWipe to help soak up extra stain.

Cleavage of His_6-tag by TEV protease

Pool the protein-containing eluate fractions and carefully transfer this solution into 3,500 molecular weight cut off (mwco) dialysis tubing that has been rinsed with nanopure water. Secure the dialysis tubing with two clips on each end – magnetic sinkers on one end and plastic floaters on the other. Place the tubing in 1.5 L of 1x TEV protease buffer. Stir overnight at 4°C. Save a 10 μL fraction of the protein as a (-)TEV sample for gel analysis. Add 1 mL of 1 mg/mL TEV protease to the dialysis tubing and invert to mix. Allow the cleavage reaction to proceed overnight with stirring at 4°C. After 24 hours, check for cleavage completion by running a Tris-tricine gel of +/- TEV fractions. Cleavage is complete when >90% of the protein is cleaved as determined by gel analysis.

Removal of TEV protease and His_6-tags

Transfer IF1 from the dialysis tubing to a 50 mL Falcon tube. Equilibrate 2-3 mL of regenerated Ni^{2+}-NTA with 20-30 mL of Wash buffer, and then batch bind the protein to the resin for 30 minutes as performed in the first First Ni^{2+}-NTA column step. Transfer the resin back to the column and collect the flow through. Add 1 column volume (2-3 mL) of Wash buffer to the resin after all protein has eluted. Collect this in the same tube as the flow through. All His_6-tags should remain bound to the resin, and the His_6-tag-free IF1 should have been eluted and collected.

Ensure that protein is present either through Tris-tricine gel analysis or by measuring the protein concentration using the Bradford method [5]. Spin concentrate and buffer exchange IF1 into 1x Storage Buffer using a 5,000 mwco spin concentrator (Amicon). In the meantime, equilibrate the Superdex 75 gel filtration column with 300 mL of 2x IF1 Storage buffer.

Gel filtration purification of IF1

Gel filtration chromatography is used as a final purification step for IF1 to remove any remaining contaminants from the protein sample. Inject the entire 1 mL IF1 sample onto a Superdex 75 column pre-equilibrated with 2x IF1 Storage buffer. Elute the protein over 1.5 column volumes

with 2x IF1 Storage buffer. The protein elutes at ~85 mL (Figure 4.2A). Spin concentrate the purified protein to ~0.5 mL and add an equal volume of 100% glycerol for storage at -20°C. Verify the identity of the final product by Tris-tricine gel analysis and/or MALDI-TOF mass spectrometry at the Columbia University Medical Center's Protein Core Facility (see Appendix A). The concentration of the final IF1-in-glycerol stock should be determined by the Bradford method, using a carefully prepared calibration curve [5, 6].

5.2.1.3 IF1 mutagenesis

In order to site-specifically label IF1 with maleimide-conjugated Cy3 or Cy5, it was necessary to introduce single Cys residues into IF1 since wild-type IF1 from *E. coli* lacks Cys residues. Multiple sequence alignment combined with structural analysis and previous IF1 mutagenesis studies identified three positions in the IF1 amino acid sequence where replacement with Cys could be made: Q10C, L14C, and T33C [7-9]. See Table 5.1 for primers designed for mutagenesis with the Stratagene QuikChange II kit.

Table 5.1 Primers for IF1 mutagenesis	
IF1.Q10C.A	5'-GCCAAAGAAGACAATATTGAAATGTGCGGTACCGTTCTTGAAACGTTGCC-3'
IF1.Q10C.B	5'-CGGTTTCTTCTGTTATAACTTTACACGCCATGGCAAGAACTTTGCAACGG-3'
IF1.L14C.A	5'-CAATATTGAAATGCAAGGTACCGTTTGTGAAACGTTGCCTAATACCATGTTC-3'
IF1.L14C.B	5'-GTTATAACTTTACGTTCCATGGCAAACACTTTGCAACGGATTATGGTACAAG-3'
IF1.T33C.A	5'-TAGAAAACGGTCACGTGGTTTGTGCACACATCTCCGG-3'
IF1.T33C.B	5'-ATCTTTTGCCAGTGCACCAAACACGTGTGTAGAGGCC-3'

Preparation of the primers

1. Determine the concentration of the primers spectrophotometrically. Assume that 1 OD_{260} = 33 μg/mL ssDNA.
2. Prepare a diluted primer solution (12 ng/μL) from the stocks using sterile nanopure water.

Preparation of the DNA template

Template DNA can be purified using a large-scale or a mini-prep kit (Qiagen). Determine the concentration of the DNA template spectrophotometrically, assuming that 1 OD_{260} = 50 μg/mL dsDNA. Prepare a 20 ng/μL solution.

Preparation of a dNTP/10x buffer solution

First make the dNTP stock solution.

Mix:

30 μL sterile nanopure water
5 μL dATP (100 mM)
5 μL dTTP (100 mM)
5 μL dGTP (100 mM)
5 μL dCTP (100 mM)
50 μL 40 mM dNTPs (10 mM of each dNTP).

Add 250 μL 10x PfuUltra HF reaction buffer to the tube with the dNTP stock. Aliquot the dNTP/buffer solution in portions of 30 μL into PCR-tubes and store at -20°C. Note: the dNTP solution should not go through multiple freeze/thaw cycles. For a 50 μL PCR reaction you need 1 μL of the dNTP stock and 5 μL of the 10x buffer. Consequently, 6 μL of the premade dNTP/10x buffer mixture is added per reaction.

Preparation of the PCR reactions

There were a number of difficulties encountered early on in IF1 mutagenesis that were later attributed to 2° structure of the primers or template DNA. These were overcome through the use of 3% final (v/v) dimethyl sulfoxide (DMSO) (Sigma; anhydrous ≥99.9%) added to each reaction mixture (Table 5.2). Add the following components in the order listed, using thin-walled PCR tubes. Before adding the Pfu Turbo, spin each reaction down briefly to collect the samples at the bottom of each tube.

Table 5.2 IF1 mutagenesis reaction mixture for use with Stratagene Quik Change II kit

Water	20.25 μL
dNTPs	1 μL
Reaction buffer (from kit)	5 μL
Primer A (12 ng/μL)	10 μL
Primer B (12 ng/μL)	10 μL
Template (20 ng/μL)	3.75 μL
DMSO	1.5 μL
Pfu Turbo	1 μL

Running the PCR reactions

Table 5.3 Thermocycler program for IF1 mutagenesis		
1. 95°C	2.5 min	Denaturation
2. 95°C	30 sec	Denaturation
3. 50°C	1 min	Annealing
4. 68°C	23 min.	Polymerization
Repeat #2-4 25x		
5. 4°C	∞	Cooling

Gel analysis of the PCR amplification and *DpnI* digestion

1. Take 10 μL of each reaction and save for running an agarose gel after the digestion is complete.
2. Add 1 μL of the *DpnI* digestion enzyme (Stratagene or NEB) directly to the PCR reaction samples, mix gently and thoroughly, spin down the reaction mixtures and immediately transfer the tubes to a 37°C incubator. Incubate for 3 hours.
3. Run a 1% agarose gel with both the 10 μL reaction mixtures before digestion and 10 μL of each reaction mixture after digestion. A strong band should appear for both sets of reaction mixtures at the correct size for the plasmid in its linear form. A weaker band for the template may appear for the reaction mixtures before digestion but should disappear after digestion.

Heat shock transformation of a plasmid into Stratagene XL1-Blue supercompetent cells:

Place five Falcon tubes (15 mL) and one microfuge tube of XL1-Blue supercompetent cells on ice. Add 1 μL of DNA to each growth tube, then wash down with 50 μL of cells. Let sit on ice for 30 min. Transfer 5 mL of LB broth to a 15 mL Falcon tube. Equilibrate to 42°C in water bath. Heat shock cells for 45 seconds at 42°C. Place on ice for 2 min. Add 500 μL LB to each growth tube. Place tubes in shaking incubator for 1.5 hours at 37°C. Plate 100 μL onto LB + carbenicillin plates. Incubate overnight at 37°C. For transformation by electroporation, follow the directions in the Stratagene QuikChange manual.

Pick six colonies for overnight cultures (5 mL LB and 5 μL carbenicillin), followed by plasmid purification by plasmid mini-prep kit (Qiagen). Elute final sample with nanopure water. Check plasmid concentration and send ~500 ng of each sample for DNA sequencing to Genewiz (Plainfield, NJ). *Note: the first sequencing trial produced no reliable signal (all N's), but the second trial, using Genewiz's GC rich/hairpin structure protocol, produced a reliable signal. The

CAA → TGC mutation (Q10C) appeared in the correct location (see notebook, p. 45). IF1-L14C and IF1-T33C were never successfully generated. Transformed the plasmid corresponding to the positive hit (sample #1) into BL21(DE3) cells by electroporation and purified protein as described in section 5.2.1.2.

5.2.1.4 IF1 labeling with maleimide-conjugated Cy3 or Cy5

Preparation of Cy3- or Cy5-maleimide dye aliquots

Dissolve 1 mg Cy3 or Cy5 with 100 μL anhydrous DMSO. (Note: it is critical that the DMSO remain anhydrous. Purge bottle with nitrogen during DMSO removal using a nitrogen-filled balloon inserted into the SureSeal with a syringe and needle.) Prepare 10 μL aliquots. Rinse original tube with 20 μL DMSO and distribute this evenly into all the aliquots. Parafilm the top of each tube and poke holes into the parafilm with a needle. Place each tube in the speed vacuum lyophilizer and cover the lid of the centrivap with Al foil to minimize the dyes' exposure to light. Lyophilize for ~1 hour, or until dry. Store the aliquots under Al foil at 4°C in a container with desiccant.

This labeling protocol is based on the Thiol-Reactive Probes protocol (Invitrogen) and Amersham CyDye Maleimides protocol (GE Healthcare). Buffer exchange protein into Labeling Buffer (100 mM Tris-OAc, pH 7.0; 50 mM KCl) using a 10,000 MWCO spin concentrator. Add 10-fold molar excess of Tris(2-carboxyethyl)phosphine (TCEP) to reduce disulfide bonds and incubate at 37°C for 1 hour under Al foil. Agitate by pipeting up and down every 10 min. Then add 20-fold molar excess of dye and incubate in the dark at 4°C overnight, 2 hours at r.t., or 30 min. at 37°C. Quench reaction by adding 5 mM BME (final conc.). Remove unreacted dye by gel filtration chromatography (Superdex 75, GE Healthcare) with 2x IF3 Storage Buffer (Figure 4.2A). Remove unlabeled IF1 by hydrophobic interaction chromatography (HIC), using a 0 → 100% IF1 HIC Buffer B gradient over 20 CV (Figure 4.2B). Spin concentrate IF1(fluor) and buffer exchange into

2x Storage Buffer. Add equal volume of 100% glycerol for storage at -20°C. This HIC purification yields 100% labeled IF1, which is ideal for smFRET experiments.

5.2.2 Preparation and purification of IF3

5.2.2.1 *E. coli* IF3 amino acid sequence

```
          10         20         30         40         50         60
GAMAKGGKRVQTA RPNRINGEIR AQEVRLTGLE GEQLGIVSLR EALEKAEEAG VDLVEISPNA

        70         80         90        100        110        120
EPPVCRIMDY GKFLYEKSKS SKEQKKKQKV IQVKEIKFRP GTDEGDYQVK LRSLIRFLEE

       130        140        150        160        170        180
GDKAKITLRF RGREMAHQQI GMEVLNRVKD DLQELAVVES FPTKIEGRQM IMVLAPKKKQ
```

*Note: The Gly and Ala are positions -2 and -1, respectively, and Met is #1, Ala is #1, and Lys is #2. These extra residues were introduced due to the cloning strategy employed. The Met is position 1 and Lys is position 2 in wild-type *E. coli* IF3.

5.2.2.2 IF3 purification

Purification of His_6-tagged IF3 is very similar to the procedure followed for His_6-tagged IF1 purification with the exception that IF3 is cleaned up by cation exchange chromatography, not gel filtration chromatography. See Fei *et al.* for additional procedural details [4]. A typical protein yield for IF3 purification is 1 mg L^{-1}.

IF3 Buffers

*Note: Chill buffers prior to use. Do not add BME, EDTA, or Imidazole until the day of use. Adjust the pH of the final solution to 7.5.

Lysis and Equilibration Buffer
10 mM Tris-HCl, $pH_{4\ C}$=7.5
60 mM NH_4Cl
10 mM $MgCl_2$
5 mM BME
0.1 mM PMSF
10 mM Imidazole

Wash Buffer
20 mM Tris-HCl, $pH_{4\ C}$=7.5
10 mM $MgCl_2$
5 mM BME
0.1 mM PMSF
0.5 mM EDTA
30 mM Imidazole

Elution Buffer
20 mM Tris-HCl, $pH_{4\ C}$=7.5
10 mM $MgCl_2$
5 mM BME
0.1 mM PMSF
0.5 mM EDTA
250 mM Imidazole

5X Tev Protease Buffer
100 mL Tris-HCl, $pH_{4\ C}$=7.5
1 M NaCl
0.5% Triton-X
10 mM BME

IF3 Storage Buffer
10 mM Tris-Acetate
50 mM KCl
10 mM Mg-Acetate
6 mM BME
50% glycerol

2X TMNDN 70 (A)
80 mM Tris-HCl, $pH_{4°C}$ = 7.5
60 mM NaCl
10 mM $MgCl_2$
80 mM NH_4Cl
2 mM BME

1X TMNDN 70 (B)
40 mM Tris-HCl
750 mM NaCl
40 mM NH_4Cl
5 mM $MgCl_2$
2 mM BME

IF3 HIC Buffer A
1 M $(NH_4)_2SO_4$
100 mM Na_2HPO_4
pH to 7.0 with HCl

IF3 HIC Buffer B
100 mM Na_2HPO_4
pH to 7.0 with HCl

Labeling Buffer
100 mM Tris-OAc, $pH_{r.t.}$=7.0
50 mM KCl

IF3 Cell Growth and Harvesting

- Grow fresh 5 mL overnight cultures of BL21(DE3)-pProEx-IF3 (MME Strain #1) in TB Growth Media containing 100 μg/mL carbenicillin at 37°C and 250 rpm shaking. *Note: the cell strain is very important. IF3 does not overexpress well in DH5α cells, while overexpression is robust in BL21 cells.

- Use the 5 mL culture to inoculate 1 L of TB Growth Media containing 100 μg/mL carbenicillin at 37°C and 250 rpm shaking.

- Induce overexpression by adding 1 mM (final) IPTG when cells reach an OD_{600} of 0.5-0.8. (Prior to inducing, store a 1 mL aliquot of –IPTG cells)

- Grow for 3-4 hours post induction at 37°C and 250 rpm shaking.

- Harvest cells with 500 mL bottles by centrifuging at 5500 rpm in a JL-10 rotor for 15 minutes. (Prior to spinning, store a 1 ml aliquot of +IPTG cells)

- Resuspend cells in 20 mL Lysis and Equilibration buffer per 1 L of culture. Either store suspension at -20°C or proceed to lysis step.

- Run a Tris-Tricine gel to check +/- IPTG cell samples for induction.

IF3 Purification

All steps from here on are done at 4°C or on ice.

- Thaw IF3 cell suspension on ice.

- Lyze cells by French Press. Four passes through press with dropwise elution. 1200 gauge.

- Centrifuge lysate in polypropylene centrifuge tubes for 30 minutes at 20,000xg in a JL-17 rotor. After this, there should be a clear supernatant and substantial pellet in the tube. Carefully decant supernatant into a Falcon tube. Save a 20 μl sample of both the pellet and supernatant to check for IF3 localization.

Ni-NTA Column Equilibration

- Pour 1-2 mL Qiagen Ni-NTA resin into a disposable column

- Equilibrate the media with 50 ml Lysis and Equilibration Buffer

First Ni column

- Batch-bind His_6-tagged IF3 to the Ni-NTA resin by adding the cleared lysate to the resin and resuspending the resin by pipetting or agitated inversions. Transfer the resin-lysate mixture to a Falcon tube and place it on either the rotator or rocker for 30 minutes.

- Transfer mixture back to the disposable column and collect all the flow thru.

- Wash column with 25-50 ml Lysis and Equilibration Buffer. Collect flow thru.

Wash

- Wash column with 25 ml of Wash Buffer – collect flow thru into 2 tubes.

Elution

- Elute the IF3 with ~25-30 ml Elution Buffer. Collect the eluate in 2-3 ml fractions in small Falcon tubes. Analyze each fraction as well as the cell pellet, lysate supernatant, column binding flow through, equilibration wash flow through, both wash flow throughs, and all the eluate fractions by a Tris-Tricine gel.

Dialysis and Concentration

- Pool relevant fractions and pipet into 3,500 mwco dialysis tubing (Spectrum Laboratories) that has been rinsed with ddH_2O. Place in 1.5 L of 1x Tev Protease Buffer containing fresh BME. Stir overnight at 4°C.

- Transfer to fresh buffer (1.5 L of 1x Tev Protease Buffer). Save a 20 μl fraction as a (–)Tev sample for gel analysis.

- Add 1 mL Tev Protease to the dialysis bag. Allow the cleavage reaction to proceed overnight with stirring at 4°C.

- Check for cleavage completion (*Note: Allow 24 hours to pass before terminating TEV cleavage) by running a tris-tricine gel of +/- Tev samples. Cleavage is complete when >90% of the protein is cleaved as determined by gel analysis (see Figure 5.6)

Removal of Tev Protease and His_6-Tags

- Transfer IF3 from the dialysis tubing to a 50 mL Falcon tube.

- Equilibrate 1-2 mL Ni-NTA resin with wash buffer

- Pour the IF3 over the resin and collect one tube. Add 1 column volume of wash buffer to the resin after all protein has eluted. Collect this in the same tube as the protein.

- Check the protein concentration using the Bradford method.

Final Purification Step: Cation Exchange Column (HiTrap SP HP)

- Instead of spin concentrating IF3, use the Superloop on the FPLC to inject the entire sample (30-40 mL).

- Wash the superloop three times with 20 ml ddH_2O. Flow rate = 2 ml/min. Pressure alarm = 2 MPa.

- MME method: if3frac

  Details:
  Column: HiTrap SP HP
  1 CV = 5 CV
  Flow rate: 0.7 mL/min
  Pressure limit: 0.30 MPa
  Monitor absorbance at: 280, 260, 230 nm

Column equilibration: 5 CV
Wash out unbound sample with 3 CV
Linear gradient: 0 → 100% TMNDN 70 Buffer B over 20 CV
Gradient delay: 2.0 mL
Clean after elution with 5 CV

- Buffers: 1x TMNDN 70 A and B (see "IF3 Buffers")
- Rinse and fill line B with buffer B so that when the gradient switches from 100% A to A & B, the line contains B.
  - Position: LOAD
  - Column Position: 1Bypass
  - Outlet Valve: Waste
  - Gradient: 100% B
  - Pressure Alarm: 0.3 MPa
  - Flow: 0.5 ml/min for ~10 ml
- Collect fractions. IF3 elutes at ~120 mL (see Figure 5.5).

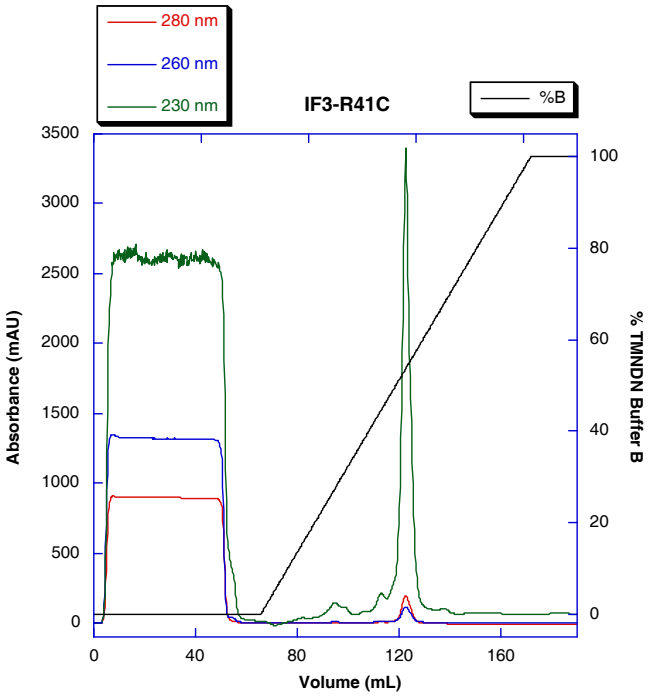

Figure 5.5 Cation exchange column purification of IF3. Representative cation exchange (HiTrap SP HP) chromatogram of IF3-R41C final purification. Gradient: 0-100% TMNDN Buffer B over 20 CV. Pure IF3 elutes at ~125 mL.

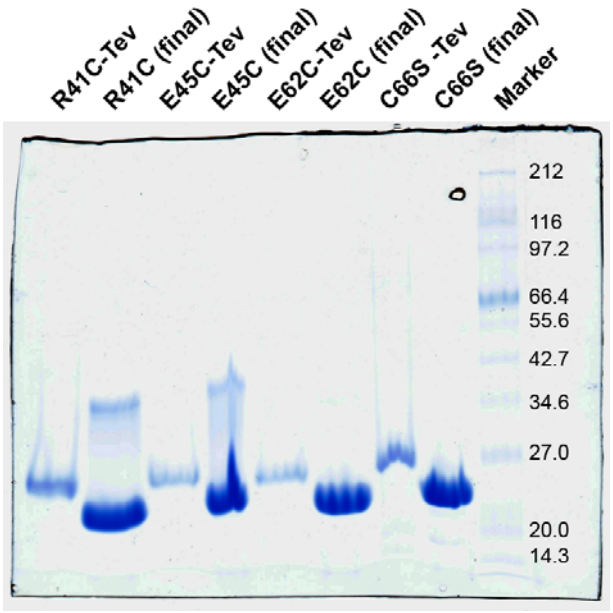

Figure 5.6 Purification of IF3 mutants. Tris-tricine gel analysis of four IF3 point mutants (R41C, E45C, E62C, and C66S) before removal of the His_6-tag by TEV cleavage ("-Tev"), and after the second Ni-NTA column and cation exchange column ("final"). The units on the protein ladder ("marker") are kDa.

5.2.2.3 IF3 mutagenesis

Table 5.4 Primers for IF3 mutagenesis	
IF3.C66S.A	5'-GCCGAGCCGCCGGTTAGTCGTATAATGGA-3'
IF3.C66S.B	5'-TCCATTATACGACTAACCGGCGGCTCGGC-3'
IF3.R41C.A	5'-GCTTGGTATTGTGAGTCTGTGCGAAGCTCTGGAGAAAGCAG-3'
IF3.R41C.B	5'-CTGCTTTCTCCAGAGCTTCGCACAGACTCACAATACCAAGC-3'
IF3.E45C.A	5'-GAGTCTGAGAGAAGCTCTGTGCAAAGCAGAAGAAGCCGGAG-3'
IF3.E45C.B	5'-CTCCGGCTTCTTCTGCTTTGCACAGAGCTTCTCTCAGACTC-3'
IF3.E62C.A	5'-GATCAGCCCTAACGCCTGCCCGCCGGTTTGTCGTA-3'
IF3.E62C.B	5'-TACGACAAACCGGCGGGCAGGCGTTAGGGCTGATC-3'
IF3.C66A.A	5'-CTAACGCCGAGCCGCCGGTTGCTCGTATAATGGA-3'
IF3.C66A.B	5'-TCCATTATACGAGCAACCGGCGGCTCGGCGTTAG-3'
IF3.Q23C.A	5'-CAATGGCGAAATTCGCGCCTGCGAAGTTCGCTTAACAGGTC-3'
IF3.Q23C.B	5'-GACCTGTTAAGCGAACTTCGCAGGCGCGAATTTCGCCATTG-3'
IF3.L30C.A	5'-GGAAGTTCGCTTAACAGGTTGCGAAGGCGAGCAGCTTGGTA-3'
IF3.L30C.B	5'-TACCAAGCTGCTCGCCTTCGCAACCTGTTAAGCGAACTTCC-3'
IF3.S39C.A	5'-AGCAGCTTGGTATTGTGTGTCTGAGAGAAGCTCTG-3'
IF3.S39C.B	5'-CAGAGCTTCTCTCAGACACACAATACCAAGCTGCT-3'
IF3.S58C.A	5'-GACTTAGTCGAGATCTGCCCTAACGCCGAGC-3'
IF3.S58C.B	5'-GCTCGGCGTTAGGGCAGATCTCGACTAAGTC-3'

5.2.2.4 IF3 labeling with maleimide-conjugated Cy3 or Cy5

Buffer exchange IF3 into Labeling Buffer and add 10-fold molar excess of TCEP, followed by 20-fold molar excess of Cy3- or Cy5-maleimide. The dye should be dissolved in enough anhydrous DMSO so that the final reaction contains 5% DMSO. Allow the reaction to proceed for 1 hr at r.t.

followed by overnight at 4°C. Quench the reaction by addition of BME. Bring the volume up to 1.5 mL by addition of an appropriate amount of labeling buffer. Remove unreacted dye by extensive dialysis into IF3 HIC Buffer A using a 10,000 MWCO 0.5-3 mL capacity slide-a-lyzer cassette (Pierce Biotechnology).

5.2.2.5 Purification of IF3(Cy3-Cy5) with hydrophobic interaction chromatography

After removing unreacted free dye by extensive dialysis, centrifuge the IF3(Cy3-Cy5) sample at 14,000 x g for 5 min to pellet any aggregated protein. Then inject the sample onto the HIC column and elute the labeled protein species with a 0-100% HIC buffer B gradient over 16 CV. This labeling strategy produces a number of labeled products, including singly and doubly labeled IF3 (see Figure 2.3A). The identity of the doubly labeled, IF3(Cy3-Cy5) species can be verified by MALDI-TOF mass spectrometry (see section 5.2.2.6 and Figure A1). See Figure 2.2B for a representative HIC chromatogram.

5.2.2.6 Trypsin digestion and MALDI-TOF Mass Spectrometry analysis of IF3(Cy3-Cy5)

The presence of two IF3(Cy3-Cy5) peaks on the HIC chromatogram (Fig. 2.2B) suggested that each peak may represent a uniquely labeled sample (*i.e.* (1) IF3(C66S/S39C-Cy3/K98C-Cy5) or (2) IF3(C66S/S39C-Cy5/K98C-Cy3)). The identities of the two IF3(Cy3-Cy5) samples (peaks 1 and 2) were verified by trypsin digestion of IF3(Cy3-Cy5), followed by matrix assisted laser desorption ionization time-of-flight (MALDI-TOF) mass spectrometry. The linker region of IF3 has previously been shown to be highly susceptible to trypsin digestion, making this an effective means to separate the two, similarly-sized globular domains for m.s. analysis [10]. Trypsin digestion and MALDI-TOF mass spectrometry analysis were performed by Dr. Mary Ann Gawinowicz at the Columbia University Medical Center's Protein Core Facility. The mass spectra are found in Appendix A.

Sample preparation: IF3(Cy3-Cy5) samples from HIC peaks 1 and 2 (Fig. 2.2B) were buffer exchanged into IF3 Storage Buffer (without glycerol) (section 5.2.2.2) and concentrated with a 10,000 MWCO Microcon spin column (Millipore).

Trypsin digestion of IF3(Cy3-Cy5): Trypsin digestion was carried out by diluting 1 μL (approximately 2 μg) of the protein solution with 10 μL 25 mM Tris, pH 8.5. Then 0.08 mg of trypsin (modified sequencing grade, Roche Applied Science) was added to this solution in 8 mL 25 mM Tris, pH 8.5 and heated overnight at 32ºC. The digestion solution was desalted using a Millipore C18 ZipTip and the eluted peptides were dried completely in a Speed-Vac concentrator.

Sample preparation for MALDI-TOF analysis: The dried peptides were re-dissolved in 3 μL matrix solution prepared as follows: A 10 mg/mL solution of α-cyano-4-hydroxycinnamic acid was prepared in 50% acetonitrile/0.1% TFA and two internal standards added, angiotensin and ACTH (7-38) peptide, so that the final concentrations were 193 fmol/mL and 340 fmol/mL, respectively. The peptide/matrix solution was spotted onto a MALDI target plate and allowed to air-dry.

MALDI-TOF analysis: The digest was analyzed on an Applied Biosystems Voyager DE-PRO in the linear mode with the following settings: Accelerating voltage 21000; Grid voltage 95%; Guide wire voltage 0.050%; Delay time 200 nsec; Laser power 1800-2000. The resulting average masses were smoothed using 19-point Gaussian smoothing, then manually calibrated using the two internal standards.

This m.s. work confirmed that peaks 1 and 2 are both mixtures of (1) IF3(C66S/S39C-Cy3/K98C-Cy5) and (2) IF3(C66S/S39C-Cy5/K98C-Cy3) (Figure A1). All smFRET data was collected using IF3(Cy3-Cy5) from peak 1, though use of IF3(Cy3-Cy5) from peak 2 gave very similar results (Figure 2.3B).

5.3 Toeprinting biochemical activity assay

The activities of ribosomes, fMet-tRNA^{fMet}, and the IFs are tested using a well-established primer-extension inhibition, or "toeprinting" assay [11, 12]. Initiation reactions are carried out on an mRNA that has been preannealed with a 5'[^{32}P]-labeled DNA primer. Subsequent reverse transcription of the primer-annealed, initiated mRNA is strongly blocked when the reverse transcriptase encounters an mRNA-bound ribosome, thereby producing a 5'[^{32}P]-labeled cDNA of defined length, or "toeprint." Analysis of the cDNA products on a 9% sequencing PAGE gel reports the position of the ribosome on the mRNA with single-nucleotide resolution. The assay for each IF is distinct, and the ones for IF1 and IF3 are described in the following sections and also in Fei *et al.* [4].

5.3.1 Toeprinting primer labeling

Thaw a sample of frozen γ-^{32}P-ATP (PerkinElmer Life and Analytical Sciences, 1 mCi/mL) at r.t. for at least 15 minutes.

Then mix:
70 pmol gp32.toe.2 primer (TATTGCCATTCAGTTTAG) (Integrated DNA Technologies)
42 pmol γ-^{32}P-ATP (6000 Ci mmol^{-1}, PerkinElmer)
14 U T4 polynucleotide kinase (New England Biolabs)
final volume of 1x T4 polynucleotide kinase buffer (New England Biolabs)

Incubate for 30 min. at 37°C, then 10 min. at 75°C to inactivate the T4 polynucleotide kinase. Remove unincorporated γ-^{32}P-ATP by gel filtration through a G-25 Sephadex spin column (GE Healthcare).

5.3.2 Primer annealing to mRNA

Mix:
4 μL ^{32}P-labeled primer
100 pmol RG13 mRNA (a.k.a mRNA #4; See Table 2.1 for sequence)
Bring to 40 μL with 25 mM Tris-OAc $pH_{r.t.}$=7.0

Heat at 90°C in a dry block heater for 1.5 min, then cool to r.t. by moving the heat block to benchtop. Store at -20°C.

5.3.3 Preparation of 9% denaturing PAGE sequencing gel

9% sequencing gel solution
100 mL 10x TBE
225 mL 40% 19:1 acrylamide:bis-acrylamide (Calbiochem)
420.5 g urea
Add nanopure water to 1 L. Then filter through a 0.45 μm (minimum) filter. Store under Al foil.

10% Ammonium persulfate, 5 mL
0.5 g Ammonium persulfate
Add water to 5 mL
Aliquot and store at -20°C

1% Dye solution, 10 mL
0.1 g xylene cyanol
0.1 g bromophenol blue
10 mL 1x TAE

Formamide loading buffer, 10 mL
900 μL Dye solution
9.1 mL formamide
Store at 4°C for up to three months

10x TBE, 1 L
108 g Tris base
55 g boric acid
40 mL 0.5 M EDTA, pH 8.0

50x TAE, 1 L
242 g Tris Base
57.1 mL glacial acetic acid
37.2 g $Na_2EDTA \bullet 2H_2O$
bring volume to 1 L with nanopure water and adjust the pH to 8.5

Prepare the gel plates (20 x 45 cm x 4.8 mm, and 20 x 43 cm x 4.8 mm) by washing with soap and water the night before, then soaking in a base bath (~1 M KOH) overnight. Before use, rinse with water and ethanol and place in a fume hood for drying. Place the plates horizontally on top of two pipet tip boxes with bench paper underneath. Silanize each plate by pipeting 2 x 500 μL Sigmacote (Sigma) onto the interface side of each plate. Wipe down with a KimWipe. Add 0.4 mm spacers, then top plate. Clamp together with six binder clips. Then, mix 60 mL 9% sequencing gel solution, 45 μL tetramethylethylenediamine (TEMED), and 200 μL ammonium persulfate (APS) in a 250 mL beaker. Swirl to mix. Use 60 mL syringe to deliver to plates, being careful not to introduce bubbles. Lastly, add comb and allow to polymerize for ~1 hour.

Gel running and imaging

Assemble gel box (Owl) and secure the gel plates using binder clamps. Be sure to insert cushions into the gaps at the top of the plates to prevent the top running buffer from leaking. Add metal plate to back of gel and secure with binder clamps. Prepare 1 L of 1x TBE buffer for top and bottom chambers. Mark the locations of each well with a marker and then remove the comb. Pre-run gel for 1 hour at 55 W, then load 5 μL samples with flat gel loading pipet tips and run gel at constant power of 55 W for ~1.5 hours or until 2nd dye front has reached the 2/3 point on the gel. Transfer the gel to Whatman filter paper (3 MM Chr; 0.34 mm) and place in gel dryer. Add dry ice and ethanol to condenser trap and open the dryer to vacuum. Dry at 55°C for 2 hours under vacuum before storing on a storage phosphor screen (GE Healthcare) overnight. Scan the phosphorimage screen with a STORM PhosphorImager and then erase the screen under fluorescent light. Analyze the data in ImageQuant, following the specific criteria specified in sections 5.3.4 (IF1 activity) or 5.3.5 (IF3 activity).

5.3.4 IF1 toeprinting activity assay

The IF1 assay tests the ability of IF1 to enhance the formation of a correctly initiated 70S IC in the presence of IF2 and IF3 [12]. Each initiation reaction is prepared as follows:

Mix the following:
12 pmol 30S subunits
12 pmol 50S subunits
Incubate at 37°C for 15 minutes.

Then add:
12 pmol IF3
48 pmol IF2
48 pmol IF1
35 nmol GTP
Incubate at 37°C for 10 minutes, then r.t.

Next add:
2.4 μL Primer annealing reaction
Incubate at 37°C for 10 minutes, then add:

35 pmol fMet-tRNA^{fMet}
35 pmol tRNA^{Phe}
Incubate at 37°C for 10 minutes, then on ice for 10 minutes.

Place initiation reactions on ice until ready for use in primer-extension reactions.

Toeprinting Mix
Prepare 10x dNTP mix by mixing together equal volumes (20 μL each) of each NTP (Promega).

Prepare sequencing polymix:
100 μL 5x polymix
5 μL 1 M $Mg(OAc)_2$
0.21 μL 14.3 M BME

Mix:
50 nmol ATP
20 nmol dATP
20 nmol dGTP
20 nmol dCTP
20 nmol dTTP
sequencing polymix to a final concentration of 1x

Add 15 μL of the Toeprinting Mix and 10 units of AMV Reverse Transcriptase (RT) (Promega) to each initiation reaction and incubate at 37°C for 15 minutes. Phenol extract the reaction mixture once, and then chloroform extract the mixture once.

Add:
2.5 μL 3 M NaOAc, pH 5.2
82.5 μL 100% r.t. ethanol

Incubate at r.t. for 10 minutes, followed by centrifugation at 14,000 rpm for 10 minutes. Wash the pellet with 70% ethanol. Centrifuge for 10 min. at 14,000 rpm. Resuspend pellet in 5 μL Denaturing PAGE loading buffer, heat at 95°C for 3 min., and load onto a 9% denaturing, sequencing PAGE gel (see section 5.4.3). For the IF1 toeprinting assay, the intensity of the +15 toeprint band should show a ~3-fold enhancement in the presence versus the absence of IF1, indicating more efficient 70S IC formation. Calculate the relative toeprint intensity as:

$$\text{Relative toeprint} = \frac{+15\ \text{Toeprint}}{\text{Full length} + +15\ \text{Toeprint}}$$

See Figure 4.3 for a representative example of a toeprinting gel to test IF1's activity.

5.3.5 IF3 toeprinting activity assay

The IF3 toeprinting activity assay measures IF3's ability to promote selection of $tRNA^{fMet}$ over $tRNA^{Phe}$ on 30S-mRNA complexes, in the absence of IF1 and IF2 [4, 12, 13]. Reactions were

performed as described in Fet *et al.* [4], Hartz *et al.* [12], and Maar *et al*. [13]. Briefly, 2 pmol 30S subunits, 0.05 μM ^{32}P-primer-annealed T7 gp32 mRNA, 20 pmol tRNA^{fMet}, and 200 pmol tRNA^{Phe} (both tRNAs were not aminoacylated) were incubated in Tris-polymix buffer (5 mM Mg^{2+}) for 10 minutes at 37°C. Varying concentrations of IF3 were added to the reaction mixtures and they were then incubated for 10 minutes at 37°C, followed by 10 minutes on ice. Reverse transcription was initiated by adding 31.3 nmol ATP, 15.6 nmol each of dGTP, dATP, dCTP, and dTTP, and 15 U AMV reverse transcriptase (Promega or NEB) and incubating for 15 minutes at 37°C. The cDNA was purified by phenol and chloroform extractions, and then precipitated with sodium acetate and ethanol. The cDNA products were resuspended in Gel Loading Buffer (see section 5.4.3) and run on a 9% Denaturing PAGE gel for 1.5 hours (see section 5.4.3). The gel was dried and exposed to a PhosphorImager screen (GE Life Sciences). The screen was scanned with a Storm PhosphorImager (Molecular Devices) and analysis was performed with Image Quant (Molecular Dynamics).

The total phosphor intensity at the +15 (tRNA^{fMet}) and +18 (tRNA^{Phe}) bands was quantified using equally sized regions and the ratio of the intensities within these regions (18/(15+18)) was used as an indication of the shift from tRNA^{Phe} (UUC) to tRNA^{fMet} (AUG) selection as the concentration of IF3 was increased. Specifically, the increase in tRNA^{fMet} selectivity upon increasing IF3 concentration was defined as follows:

$$\%\text{Selectivity of tRNA}^{fMet}\text{ over tRNA}^{Phe} = \left(1 - \left(\left(\frac{UUC}{AUG}\right)_{\text{lane x}} \div \left(\frac{UUC}{AUG}\right)_{(-)IF3}\right)\right) \times 100$$

In addition to testing wild-type IF3, unlabeled IF3(C66S/S39C/K98C), and IF3(Cy3-Cy5), the singly-labeled variants IF3(C66S/S39C)-Cy3 and IF3(C66S/K98C)-Cy5 were tested for their ability to select tRNA^{fMet} over tRNA^{Phe}. Gradual loss in activity was seen upon introduction of one, and then two, fluorophores. Full activity was achieved under increased IF3 concentrations. See Figure 2.5B and C for examples of typical IF3 toeprinting gels.

5.4 TIRFM-based fMet-(Cy3)tRNA^{fMet} dissociation assay

This assay probes IF3's ability to destabilize tRNAs, specifically fMet-(Cy3)tRNA^{fMet}, in the presence of IF1 [14]. Complexes were prepared with 1.8 μM biotinylated-mRNA, 0.6 μM 30S subunits, 0.9 μM IF1, 0.9 μM IF3 (when included), and 0.6 μM fMet-(Cy3)tRNA^{fMet}. The complexes were incubated at 37°C for ten minutes, aliquoted, flash frozen, and stored at -80°C until use. To quantify the extent of fMet-(Cy3)tRNA^{fMet} dissociation, complexes were diluted to 300 pM, keeping IF1 and IF3 (when included) in solution at 0.9 μM. After incubating the complexes on the slide surface for 5 minutes, unbound molecules were washed out with a buffer containing 0.9 μM IF1, 0.9 μM IF3, enzymatic oxygen scavengers, and triplet state quenchers in low-salt Tris-polymix buffer (see section 5.5.1). The first of fourteen movies was collected ten minutes after sample immobilization. Movies were collected at a frame rate of 10 frames per sec for 100 frames. The data were analyzed by summing the maximum intensity of each frame and then setting a lower threshold of 3500 arbitrary units. Regions larger than 4 pixels (2 x 2 binning) were composed of multiple dyes and discarded in all datasets. The two highest and the two lowest total region counts from all fourteen movies were discarded as outliers and the average and standard deviation of the remaining ten movies was taken. Raw counts were converted to "% bound" by normalizing the (+)IF3 data to the (-)IF3 data and assuming "100% bound" for the (-)IF3 complexes. See Figure 2.4. The activity of IF3(C66S/S39C/K98C) and IF3(Cy5-Cy5) (here doubly labeled with Cy5 due to the use of fMet-(Cy3)tRNA^{fMet}) is within error of wild-type IF3. Wild-type IF3 generally shows ~15% of the total number of fMet-(Cy3)tRNA^{fMet}-bound 30S ICs bound in the absence of IF3 (e.g. 45 molecules vs. 300 molecules).

5.5 smFRET imaging using TIRF microscopy

5.5.1 Oxygen scavenging system and triplet state quenchers

The oxygen scavenging system for single-molecule fluorescence work was optimized for the IF3(Cy3-Cy5) signal based on the following references: [15-17]. The system used previously (1%

β-D-glucose, 55 U/mL glucose oxidase, 365 U/mL catalase) was inadequate for the IF3(Cy3-Cy5) smFRET signal, allowing for only 1-2 seconds of observation time before fluorophore photobleaching. I modified my concentrations to those reported by Selvin and Ha: 165 U/mL glucose oxidase, and 2170 U/mL catalase.

80x stocks
13,200 U/mL glucose oxidase
173,600 U/mL catalase

Our laboratory stocks of glucose oxidase and catalase are 5,000 U/mL and 200,000 U/mL, respectively. These are prepared as follows:

Glucose oxidase (GOD) is Type VII-S from *Aspergillus niger* (Sigma; G-7016). 10,000 U is 0.055 g. Catalase (CAT) is from bovine liver (Sigma; C-40). CAT is typically 9,360,000 U/g solid and is purchased in 100 mg (936,000 U) bottles. 98% 1,3,5,7-cyclooctatetraene (COT) (Aldrich) is a liquid at r.t. with molarity of 8.7 M. 99.5% 3-nitrobenzyl alcohol (NBA) (Fluka) is barely solid at r.t. but it becomes a liquid at 37°C. Therefore, incubate NBA at 37°C immediately before preparing the 1000x stock. Its molarity is 8.4 M.

GOD/CAT Storage buffer:
50 mM Tris-OAc, $pH_{r.t.}$=7.0
50 mM KCl
500 mM BME
50% glycerol

500 μL 5 U/μL GOD
0.014 g 181,500 U/g GOD
487 μL GODCAT Storage buffer

500 μL 200 U/μL CAT
0.011 g 9,360,000 U/g CAT
489 μL GODCAT Storage buffer

*Note: Do not vortex GOD or CAT stocks. Simply pipet up and down, being careful not to introduce bubbles into the solution. Let the solutions sit at 4°C overnight. Once fully dissolved, centrifuge the solutions for 5 min. at 14,000 rpm and r.t. Transfer the top 75-80% of the supernatants into new tubes.

To prepare 80x stocks:

Mix: 7.92 μL 5 U/μL glucose oxidase
2.61 μL 200 U/μL catalase

Dilute the 80x stock to 1x in a buffer containing 1% β-D-glucose for smFRET work.

1000x COT/NBA stock solution (1 M each)
115 μL 8.7 M COT
119 μL 8.4 M NBA
766 μL Ethanol

Dilute the 1000x stock to 1x for smFRET experiments. The imaging buffer should contain GOD, CAT, COT, and NBA.

5.5.2 5x Low salt Tris-polymix

Mix:
50 mM Tris-OAc ($pH_{25°C}$ = 7.0)
100 mM KCl
5 mM NH_4OAc
0.5 mM $Ca(OAc)_2$
0.5 mM EDTA
5 mM putrescine-HCl
1 mM spermidine
5% β-D-glucose

*Note: filter salts before addition of polyamines. Store 1 mL aliquots at -20°C.

5.5.3 Passivation of microfluidic flowcells

This protocol was developed by Prof. Ruben Gonzalez at Stanford University and is based on a protocol developed by T. Ha and co-workers in the S. Chu research group.

Chemicals
1 M KOH (EMD)
30% Hydrogen peroxide (J.T. Baker)
100% (200 proof) Ethanol (Decon)
Spectroscopic grade acetone (J.T. Baker)
Vectabond silanization reagent (Vector Laboratories; SP-1800; 7 mL)
PEG reagents (PEG-Succinimidyl Valerate; 5,000 m.w.) (Laysan Bio., Inc.)
Biotin-PEG-SVA-5000-100mg
mPEG-SVA-5000-1g
100 mM Potassium borate, pH 8.4 (prepare 100 mM $KHBO_4$ and then pH to 8.4 with Boric acid)

Equipment
10 quartz slides from G. Finkenbeiner
12 coverslips (24 x 30 mm) (VWR; No. 1½, 48404-466)
1 1 L plastic graduated cylinder
1 250 mL plastic graduated cylinder
2 5 mL glass graduated pipets

1 Propane torch
1 scissor-type forceps for holding quartz slides
1 tweezers for holding coverslips
4 drill bits (Starlite Industries)
2 polypropylene slide staining rack (Electron Microscopy Sciences; 70321-10)
3 porcelain coverslip staining racks (Thomas Scientific 8542E40)
4 large (500 mL) plastic containers for dipping slides (Nalgene)
4 small (250 mL) plastic containers for dipping coverslips (Nalgene)

Drill holes

Using acrylic jig, drill 10 holes in each quartz slide. Drill all holes slowly, with the drill tip submerged in water on the surface. Approximately 30 holes (3 slides) can be drilled with each diamond tip drill bit. Remove slide after drilling and rinse thoroughly with nanopure water to remove quartz dust residue. Store in water overnight.

Wash all glassware, porcelainware and plasticware

Bleach all glassware, porcelainware, and plasticware to be used with a 100 mM KOH / 5% H_2O_2 solution.

Mix:
100 mL 1M KOH
166 mL H_2O_2
bring volume to 1 L with nanopure water

Fill all containers and lids with bleach solution and allow to sit for $\geq$30 min. Then, rinse thoroughly with nanopure water and air dry overnight.

Degrease

Place slides in plastic holder in a large plastic container. Fill container with 100% ethanol and sonicate for 15 min. Rinse completely with nanopure water. Do the same for coverslips using porcelain rack and small container.

Activate

Place slides in plastic holder in a large plastic container. Fill container with 1M KOH and sonicate for 15 min. Treat coverslips identically using a porcelain holder and a small plastic container. After

sonication, rinse the slides and coverslips 5x with nanopure water. Fill containers with nanopure water and sonicate slides and coverslips for 5 min.

Flame

Use a KimWipe to soak up excess water from edges of the slides and coverslips. Then flame the degreased and activated slides and coverslips with a propane torch. Four slow passes for each side of the slide with the slide ~1 inch from metal opening on the propane torch and three quick passes for each side of the coverslip to avoid warping. Place slides in plastic holder and coverslips in dry porcelain rack to cool.

Silanize

Place slides in plastic holder in a large plastic container and fill the container with 250 mL acetone. Place the coverslips in a porcelain holder in a small plastic container and fill the container with 140 mL acetone. Using a glass pipet, add 4.5 mL Vectabond to the 250 mL acetone in the large plastic container (final Vectabond concentration of 1.8%). Slowly stir the slides in the acetone/Vectabond solution (do not mix by dipping slide holder in and out of the acetone/Vectabond solution). Allow the Vectabond reaction to proceed for 2 min. Quench the reaction by transferring the rack to a fresh container with 50% acetone/50% nanopure water. Slowly stir the slides in the acetone/water solution for 2 min. Decant acetone/water solution and replace with nanopure water. Rinse 3x with nanopure water. Slowly stir the slides in the nanopure water for 2 min. Allow the slides to sit in nanopure water for 10 min. and then air dry.

Repeat the above silanization reaction with the coverslips. Using a glass pipet, add 2.5 mL fresh Vectabond to the 140 mL acetone in the small plastic container with the coverslips (final Vectabond concentration of 1.8%). Slowly stir the slides in the acetone/Vectabond solution (do not mix by dipping slide holder in and out of the acetone/Vectabond solution). Allow the Vectabond reaction to proceed for 2 min. Quench the reaction by transferring the rack to a fresh container with 50% acetone/50% nanopure water. Slowly stir the coverslips in the acetone/water

solution for 2 min. Decant the acetone/water solution and replace with nanopure water. Rinse 3x with nanopure water. Slowly stir the coverslips in the nanopure water. Allow the coverslips to sit in nanopure water for 10 min. and then air dry.

First PEG derivatization

Take PEG reagents out of the freezer and warm rapidly to room temperature. Place slides flat with only the upper and lower edges touching the flat surface. Use a large Eppendorf rack with tape along the top and bottom. Make sure the "blown-out" side of the drilled holes is facing up—this is the side you will PEG. Weigh out 60 mg of mPEG-SVA and 10 mg Biotin-PEG-SVA-5000 for 10 slides. Dissolve biotin-PEG in 1 mL ice-cold nanopure water. Vortex thoroughly and add 60 μL to the mPEG-SVA. Add 600 μL of ice-cold 100 mM potassium borate, pH 8.4 to this mixture. Vortex ~10 sec, centrifuge for ~30 sec to pellet any insoluble material and rapidly pipet 60 μL of this mixture onto the center of each quartz slide. Place a coverslip on top of each quartz slide by carefully laying it down edgewise so as to not trap any air bubbles in the space between the quartz slide and the coverslip. Pipet 60 μL of this mixture onto one of the extra coverslips and place the other coverslip on top (this gives you two extra coverslips in case of breakage). Incubate 1 hr at r.t. in the dark (a drawer works well). Before separating the coverslips from the quartz slides, label each slide and coverslip at the lower right hand corner with a marker so that you can tell which sides of the slide and coverslip were treated with PEG. Carefully remove the coverslip from the slides. Leave the quartz slides PEG side facing up toward you on the taped Eppendorf rack and place the coverslips on a dry porcelain coverslip staining rack.

Second PEG derivatization

Dissolve a second aliquot of mPEG-SVA in 660 μL of ice-cold 100 mM potassium borate, pH 8.4 and pipet another 60 μL onto the center of each quartz slide. Replace the coverslip on top of each quartz slide by carefully laying it down edgewise so as to not trap any air bubbles in the space between the quartz slide and the coverslip. Make sure that both the quartz slides and

coverslips are reassembled such that the sides treated with PEG during the first PEG derivatization are treated a second time. Incubate 1 hr at r.t. in the dark. Carefully remove the coverslip from the slides. If separation is difficult, briefly dip the slide/coverslip in water before separating. Rinse slides and coverslips continuously with nanopure water for 5-10 min.

Dry

Air dry the slides in a large plastic container and coverslips in a small plastic container overnight before assembling flowcell.

Flowcell assembly

See Figure 1.17A for depiction of an assembled flowcell. Cut thin (~2 mm) strips of double sided tape (3M; ½" x 250") using a razor blade, and trim the strips so they are ~1 mm long. Adhere six strips of tape to a PEG derivatized slide as depicted in Figure 1.17A. Place a glass cover slip on top, PEG slide toward the slide. Using a PCR tube, firmly but gently press the cover slip onto the tape to seal each flowcell chamber. Then seal the sides of the flowcell with epoxy (Devcon). Each of the five flowcells on the slide should now be leak-proof. Samples can be delivered to each flowcell through the drilled holes using a P200 pipet, and collected from the other hole in the flowcell.

5.5.4 Sample delivery to microfluidic flowcells

Buffers and reagents

TP50 buffer
10 mM Tris-OAc, $pH_{r.t.}$=7.0
50 mM KCl

DNA duplex (100 μM)
DNA oligo sequences:

DNA A: CGT TTA CAC GTG GGG TCC CAA GCA CGC GGC TAC TAG ATC ACG GCT CAG CT
DNA B: AGC TGA GCC GTG ATC TAG TAG CCG CGT GCT TGG GAC CCC ACG TGT AAA CG

Preparation instructions

Make 1 mM solutions of DNA A and DNA B. (e.g. 128 nmol DNA + 128 μL nanopure water)

120 μL 1 mM DNA A
120 μL 1 mM DNA B
12 μL 1 M Tris-OAc ($pH_{r.t.}$=7.0)

Nanopure water to 1.2 mL

Then, heat at 95°C for 2 min. in a dry heat block and then move to a benchtop to cool.
Once the solution reaches 70°C, add 30 μL of 2 M KCl.
Cool to r.t. and store at -20°C.

Block
10 μM Ultrapure BSA (Ambion, cat. #AM2616, 50 mg/mL)
10 μM DNA duplex
Dilute with TP50 buffer

Strep & Block
1 μM Streptavidin (Invitrogen, 5 mg, S888)
10 μM BSA
10 μM DNA duplex
Dilute with TP50 buffer

For sample delivery to microfluidic flowcell, add the following using a P200 pipet:
200 μL TP50 buffer
25 μL Block (incubate for 5 min)
25 μL Strep & Block (incubate for 5 min)
200 μL TP50 buffer
200 μL Low-salt polymix buffer
25 μL sample (incubate for 5 min)
100 μL Low-salt polymix buffer containing GOD/CAT, and COT/NBA

30S initiation complex assembly

The following components are added in the order and at the concentrations listed:

Biotin-mRNA	0.6 μM
fMet-tRNA^{fMet}	0.9 μM
IF1	0.9 μM
IF2	0.9 μM
IF3(Cy3-Cy5)	0.6 μM
GTP	1 mM
Low-salt polymix	1x
30S subunits	0.6 μM

The mixture is then incubated at 37°C for 10 minutes. 0.5 μL aliquots are prepared, flash frozen with liquid N_2, and stored at -80°C until use.

5.5.5 Data acquisition

Turn on green and red lasers 30 minutes before use. Add 2 drops of nanopure water to the microscope objective and secure the microfluidic flowcell to the microscope stage using the stage clips. Center the flowcell of interest over the objective. Add one drop of oil to the prism base,

place over the flowcell and tighten. Focus the microscope using the eyepiece first, then using MetaMorph.

Data collection

A good strategy for determining the presence and location of Cy5 (acceptor) fluorophores is to directly excite Cy5 with a red (643 nm) laser for the first frame of each movie, followed by continuous green (532 nm) laser direct excitation of Cy3 and excitation of Cy5 through FRET for the remainder of the movie. A journal should be written in MetaMorph (Molecular Devices) that switches illumination from red to green at the second frame of the movie. Following acquisition of each movie, stack the single, red laser-illuminated Cy5 frame on top of the green laser-illuminated frames by following these steps:

To focus the FOV:

→ Acquire

→ Acquire

→ Show Live

Standard camera settings

Acquire → Acquire … → Special

Setting	Value
Digitizer	10 MHz (EM Gain)
Gain	Gain 2 (2x)
EM Gain	3500
Camera Shutter	Always Open
Clear Mode	Clear PRE Sequence
Clear Count	2
Frame to Average	1
Trigger Mode	Normal (Timed)
Live Trigger Mode	Normal (Timed)
Show Focus Indicator Box	Unchecked

Acquire → Acquire … Main window

Setting	Value
Exposure Time	100 ms
Binning	2
Live Bin	2
Use Active Region	Created Region

To collect a movie

→ Acquire

→ Stream Acquisition

-- Adjust the number of frames to the appropriate number (usually 600 to 1200)

→ Acquire

To add a frame to a movie

→ Stack

→ Add Plane

-- Adjust the source and resultant stack accordingly

5.6 smFRET data processing and analysis

*Note: This data analysis protocol was used for the analysis of IF3(Cy3-Cy5) data.

Open movie file (Stack file .stk) in MetaMorph. Scroll to the beginning of the movie and follow these steps to split the image into Cy3 and Cy5 channels:

→ Display

→ Split view

→ Align

Region size:

Height: 256 pixels Width: 127 pixels

	H Shift	V Shift
W1:	A	B
W2:	X	Y

→ ‘A’, ‘B’, ‘X’, and ‘Y’ are values determined during DualView alignment. The Height and Width parameters may also vary, depending on the DualView alignment.

→ Split

Source image: All planes

Destination: Separate images

→ Apply

→ Close

Cy3 (green) channel is labeled W1, and Cy5 (red) channel is labeled W2.

Scroll to the beginning of the W2 movie. On the left, locate the “Threshold” button, then click “Auto Threshold for Light Objects.” Depending on whether or not the auto-threshold seems adequate by eye, the threshold can be adjusted manually using “Threshold Image.” Then adjust the “Low” value accordingly. To determine if the threshold is adequate, compare the thresholded image with the second frame of the Cy3 movie. There should be approximately the same number of spots in both the Cy3 and Cy5 FOVs. If not, adjust the threshold manually.

→ Regions

→ Create Regions Around Objects

→ Regions

→ Transfer Regions

→ Source Image: W2

Destination Image: W1

→ OK

Click on the W1 window.

→ Apps

→ Graph Intensities

→ Measure From: → Stack

→ OK

→ Begin

Right click on graph

→ Show Graph Data

Copy data to Excel

Excel can only merge 116 data sets at a time, so it is necessary to break data sets containing more than 116 molecules into separate spreadsheets (not worksheets)

Crtl + C = copy

Sheet 1 = Cy3 data

Repeat procedure for Cy5 data.

Sheet 2 = Cy5 data

Merge the two sets of data with the macro:

Crtl + Q

Graph all fluorescence trajectories using the macro:

Crtl + G

Save Excel spreadsheets and data in same folder:

Example: One1.xls

5.6.1 Selection of smFRET trajectories

5.6.1.1 Selection of smFRET trajectories using Microsoft Excel

Each trajectory needs to be visually inspected and either discarded or saved for further analysis. The selection criteria are as follows:

1. Anticorrelation of Cy3 and Cy5 intensities, as determined by visual inspection.
2. Single-step photobleaching, as determined by visual inspection.

The data from those traces that meet these criteria should be copied and pasted into a new spreadsheet and saved appropriately.

After saving Excel file, convert to .dat file to open in Matlab.

Delete the top row of all data (the row indicating the region identifier). Also, format the data to have two decimal places:

→ Format cells

→ Number

→ Decimal Places: 2

To save:

→ File

→ file.dat (Be sure to add this extension)

→ Save as type: Text (Tab delimited)

In the warning windows which appear upon saving, click 'OK' and then 'Yes'. When exiting Excel, do not save changes to the file.

5.6.1.2 Selection of smFRET trajectories using Matlab

Obtain a copy of FDAP v1.7, which contains all the Matlab scripts necessary for data processing, as well as a manual written by Pallav Kosuri, a former Gonzalez group rotation student. All scripts described in the following pages can be found in Appendix E. Copy the script "LoadTraces" and paste it into the folder containing all .dat data.

Open Matlab and change the current directory to the folder containing the .dat files and the LoadTraces script. The following example assumes each file was saved as 'red1.dat' 'red2.dat' etc.

Load traces into Matlab:

```
>> X=LoadTraces('red',2);
```

where X is a variable, 'red' is the name of the experiment, and 2 is the number of 'red' files.

Now set the current directory to the FDAP v1.7 folder.

Separate the Cy3 and Cy5 trajectories:

```
>> [cy31,cy51]=separateCy(X);
```

Plot the traces:

```
>> plotTraces(cy31,cy51);
```

Now, individually inspect each trajectory and note which trajectories should be omitted from further analysis steps. These trajectories will make up a vector in the J-filter step. Also take note of those trajectories that do not show photobleaching before the end of the trajectory. These trajectories will be included in the baseline correction step of data processing.

```
>> J = [A,B,C,D, … ];
```

Where A,B,C,D, … are the labels (the number above each graph) of each bad trajectory.

Apply the J-filter to purify out any unwanted FRET trajectories:

```
>> [cy32,cy52]=J_Filter(cy31,cy51,J);
```

Baseline correction:

```
>> [cy33,cy53]=correctBaseline_end(cy32,cy52,[X,Y,Z]);
```

Where X,Y,Z are the trajectories which do not show photobleaching. Note that the correctBaseline_end script can be modified according to the experimentally determined bleedthrough coefficient. For all work done in this thesis, the bleedthrough coefficient used was 0.0700.

Plot FRET trajectories:

```
>> plotFRETtraces(cy33,cy53);
```

Plot time evolution of population FRET histogram:

```
>> plotTimeFRET(cy33(2:401,:),cy53(2:401,:),24,2,10);
```

This is based on: plotTimeFRET(cy3x,cy5x,FRETbins,Tbinsize,cutoffT), where FRETbins are the number of bins in the FRET dimension (default: 24), Tbinsize is the number of datapoints in the time averaged over (default: 9) and cutoffT is the cutoff in the time dimension (in seconds). The first frame is omitted from the plot by using: cy33(2:401,:), where 401 is the total number of frames and 2 is the first frame to appear in the histogram.

Plot a 1D population FRET histogram:

```
>> FH=plotFRET(cy33(2:401,:),cy53(2:401,:),30);
```

This is based on: FH=plotFRET(cy3x,cy5x,bins), where bins is the number of bins in the histogram. The histogram data is stored in the variable FH as a two-row matrix with x-values in the first row and y-values in the second row.

To transpose the FH data from two rows into two columns:

>> FH=FH';

Save the processed FRET data using:

>> saveTraces(cy33,cy53,'<filename.dat>');

*It is important to include .dat in the filename when saving the data set.

5.6.2 Fitting a population FRET histogram with Gaussian distributions in Origin

In order to fit a 1D population FRET histogram with Gaussian distributions, copy and paste the contents of the FH variable (see above) into two columns in OriginPro 8 (Origin Lab). Highlight these columns.

→ Plot
    → Columns/Bars
        → Column
→ Analysis
    → Fitting
        → Fit Multi-peaks

5.6.3 smFRET trajectory idealization

See the vbFRET manual (http://vbfret.sourceforge.net/) for detailed directions on the use of the program. To open vbFRET, set the current directory to the 'vbFRET gui' folder. In the command window, type:

>> vbFRET;

In the vbFRET window, open a .dat file through File → Load Data. Then select one or many .dat files in the loadData window.

<u>Photobleaching removal</u>

Traces → Remove Photobleaching

Photobleaching Identification Method → 1D FRET

Truncate data when FRET exceeds 1 or 0 by more than → 0.1

Truncate an extra → 2 time steps

Minimum trace length → 10

Click 'Remove Photobleaching' button

Analysis Settings

Number of FRET states possible:

Min: 1 Max: 5

Fitting attempts per trace: 10

→ Analyze Data

Following trace fitting, save the data:

File → Save Data

→ Save Idealized Traces

→ Save as concatenated text file (.dat)

This will be the 'PATH' file that can be imported into Matlab for dwell time analysis (section 5.6.4). To determine the number of states sampled in each trajectory, visually inspect each idealized trajectory. Ignore transitions that occur between 0.05 FRET or less.

5.6.4 Dwell time analysis

The following protocol was followed to determine the 30S subunit binding and dissociation rate constants for IF1(Cy5). It is adapted from protocols developed by Drs. Jingyi Fei and Jiangning Wang [1, 18].

Set the current directory in Matlab to the "lifetime" folder within the hammy scripts folder of the FDAP 1.7 folder.

To plot an idealized 1D FRET histogram:

```
>> X=linspace(-0.2, 1.2, 30);
```

Drag PATH data (saved in vbFRET) into the command window, and open using:

```
>> H=hist(PATH DATA(:,2),X);
>> H=H';
```

Then, open and fit the data in Origin.

Now set the current directory in Matlab to the "lifetime analysis_W&L" folder. These scripts were written by Dr. Jiangning Wang.

To extract the dwell times spent in each zero or non-zero FRET state:

```
>> dwellData = getRawDwell(pathData);
```

First, remove the first and last dwells with:

```
>> dwellData_s1 = purifyRawDwell(dwellData);
```

Then, remove single data point transitions with:

```
>> dwellData_s2 = purifyOnOffDwell(dwellData_s1);
```

Separate the zero from non-zero FRET data points using a cutoff threshold of 0.2. If there are transitions between two non-zero FRET states, these will be combined.

```
>> [ts,N]=getDecay(dwellData_s2,[-0.25,0.2,0.2,1.2],time cutoff);
>> [ts,N]=getDecay(dwellData_s2,[0.2,1.2,-0.25,0.2],time cutoff);
```

For a 'time cutoff', I use 200 sec.

The data are automatically saved in a folder with the name decaycurvedata.dat. To obtain the lifetimes of each state, open "decaycurvedata.dat" in Excel, and copy the data to Origin. Fit the data with an exponential decay curve to extract the lifetimes.

5.7 References

1. Wang, J., *Regulation of IF2 Binding Kinetics and 30S IC Conformational Dynamics during Translation Initiation*, 2010, Columbia University.

2. Englander, M.T., *The ribosome discriminates the structure of the amino acid at its peptidyl-transferase center*, 2011, Columbia University: New York.

3. Dubnoff, J.S. and U. Maitra, *Isolation and properties of polypeptide chain initiation factor FII from Escherichia coli: evidence for a dual function.* Proc Natl Acad Sci U S A, 1971. **68**(2): p. 318-23.

4. Fei, J., J. Wang, S.H. Sternberg, D.D. MacDougall, M.M. Elvekrog, D.K. Pulukkunat, M.T. Englander, and R.L. Gonzalez, Jr., *A highly purified, fluorescently labeled in vitro translation system for single-molecule studies of protein synthesis.* Methods Enzymol, 2010. **472**: p. 221-59.

5. Bradford, M.M., *A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding.* Anal Biochem, 1976. **72**: p. 248-54.

6. Simonian, M.H. and J.A. Smith, *Spectrophotometric and Colorimetric Determination of Protein Concentration*, in *Current Protocols in Molecular Biology*2001, John Wiley & Sons, Inc.

7. Croitoru, V., M. Bucheli-Witschel, P. Hagg, F. Abdulkarim, and L.A. Isaksson, *Generation and characterization of functional mutants in the translation initiation factor IF1 of Escherichia coli.* Eur J Biochem, 2004. **271**(3): p. 534-44.

8. Spurio, R., M. Paci, R.T. Pawlik, A. La Teana, B.V. DiGiacco, C.L. Pon, and C.O. Gualerzi, *Site-directed mutagenesis and NMR spectroscopic approaches to the elucidation of the structure-function relationships in translation initiation factors IF1 and IF3.* Biochimie, 1991. **73**(7-8): p. 1001-6.

9. Gualerzi, C.O., R. Spurio, A. La Teana, R. Calogero, B. Celano, and C.L. Pon, *Site-directed mutagenesis of Escherichia coli translation initiation factor IF1. Identification of the amino acid involved in its ribosomal binding and recycling.* Protein Eng, 1989. **3**(2): p. 133-8.

10. Fortier, P.L., J.M. Schmitter, C. Garcia, and F. Dardel, *The N-terminal half of initiation factor IF3 is folded as a stable independent domain.* Biochimie, 1994. **76**(5): p. 376-83.

11. Hartz, D., D.S. McPheeters, R. Traut, and L. Gold, *Extension inhibition analysis of translation initiation complexes.* Methods Enzymol, 1988. **164**: p. 419-25.

12. Hartz, D., D.S. McPheeters, and L. Gold, *Selection of the initiator tRNA by Escherichia coli initiation factors.* Genes Dev, 1989. **3**(12A): p. 1899-912.

13. Maar, D., D. Liveris, J.K. Sussman, S. Ringquist, I. Moll, N. Heredia, A. Kil, U. Blasi, I. Schwartz, and R.W. Simons, *A single mutation in the IF3 N-terminal domain perturbs the fidelity of translation initiation at three levels.* J Mol Biol, 2008. **383**(5): p. 937-44.

14. Antoun, A., M.Y. Pavlov, M. Lovmar, and M. Ehrenberg, *How initiation factors maximize the accuracy of tRNA selection in initiation of bacterial protein synthesis.* Mol Cell, 2006. **23**(2): p. 183-93.

15. Dave, R., D.S. Terry, J.B. Munro, and S.C. Blanchard, *Mitigating unwanted photophysical processes for improved single-molecule fluorescence imaging.* Biophys J, 2009. **96**(6): p. 2371-81.

16. Aitken, C.E., R.A. Marshall, and J.D. Puglisi, *An oxygen scavenging system for improvement of dye stability in single-molecule fluorescence experiments.* Biophys J, 2008. **94**(5): p. 1826-35.

17. Joo, C. and T. Ha, *Single-Molecule FRET with Total Internal Reflection Microscopy*, in *Single-Molecule Techniques*, P. Selvin and T. Ha, Editors. 2008, Cold Spring Harbor Laboratory Cold Spring Harbor, New York. p. 3-36.

18. Fei, J., *Coupling of Ribosome and tRNA Dynamics during Protein Synthesis*, 2010, Columbia University: New York. p. 211.

Appendix A

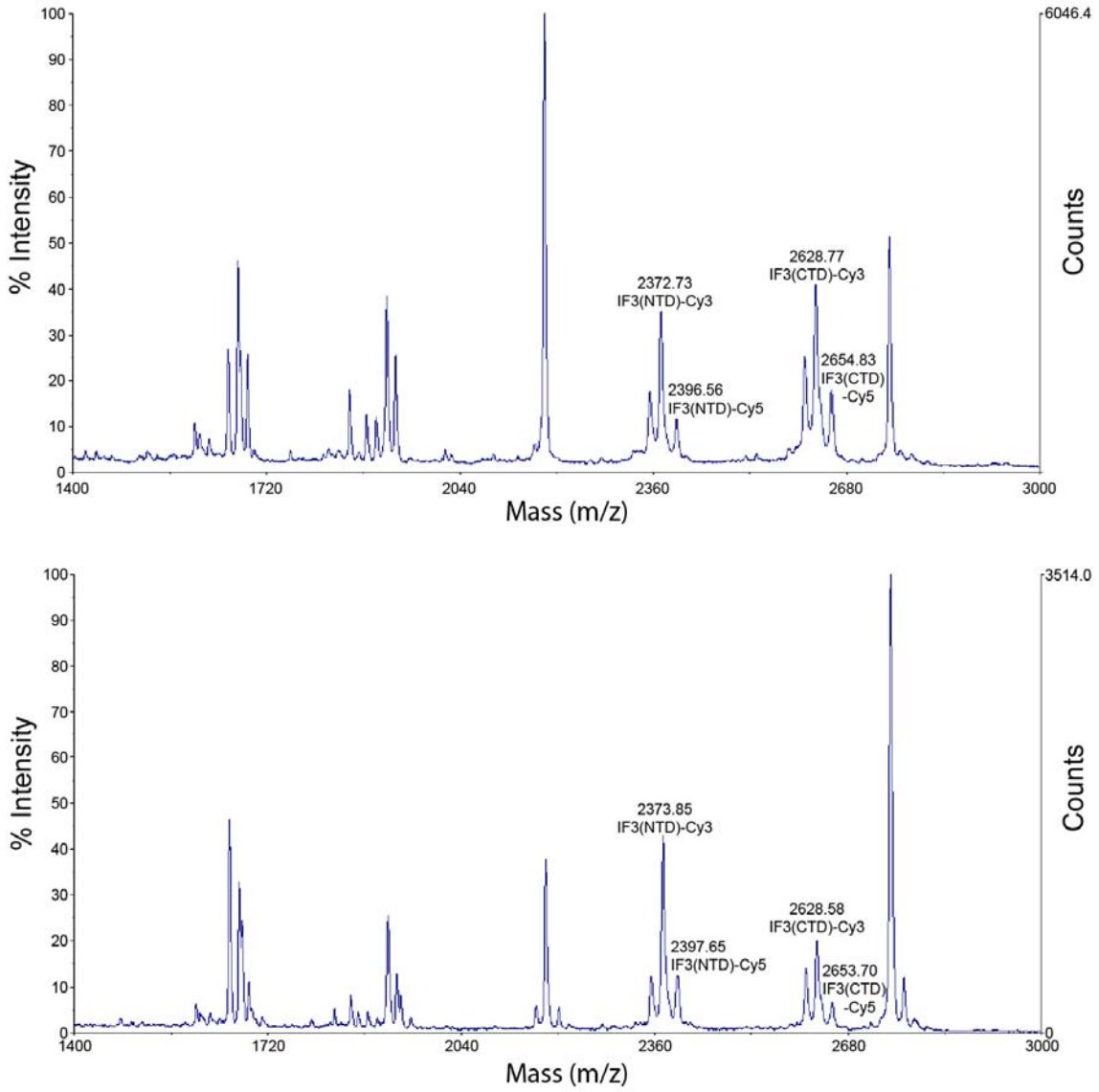

Figure A1. MALDI-TOF mass spectrometry analysis of trypsin digested IF3(Cy3-Cy5). Samples from HIC purification **(top)** peak 1, and **(bottom)** peak 2 (See Figure 2.2). See section 2.3.9 for a description of the experiment, and 5.2.2.6 for detailed materials and methods.

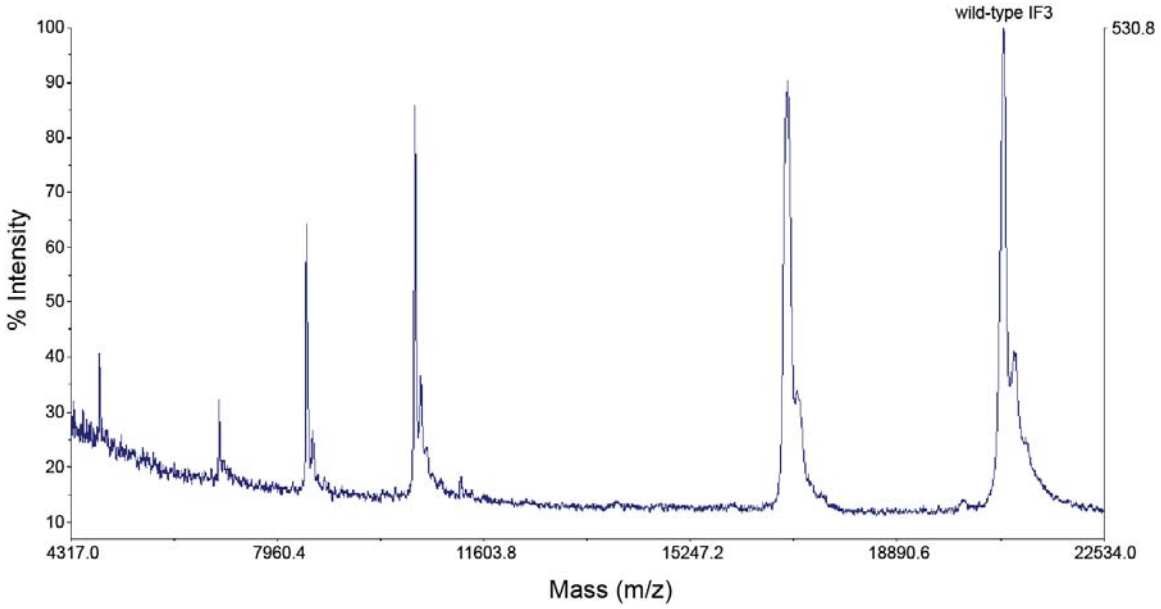

Figure A2. MALDI-TOF mass spectrometry analysis of wild-type *E. coli* IF3. The peak at 20,764.61 m/z is wild-type IF3 (expected m.w. of Gly + Ala + IF3 = 20,763.1 Da). The peaks at 8,476.78 and 16,952.56 are the singly- and doubly-ionized forms of myoglobin, the internal standard.

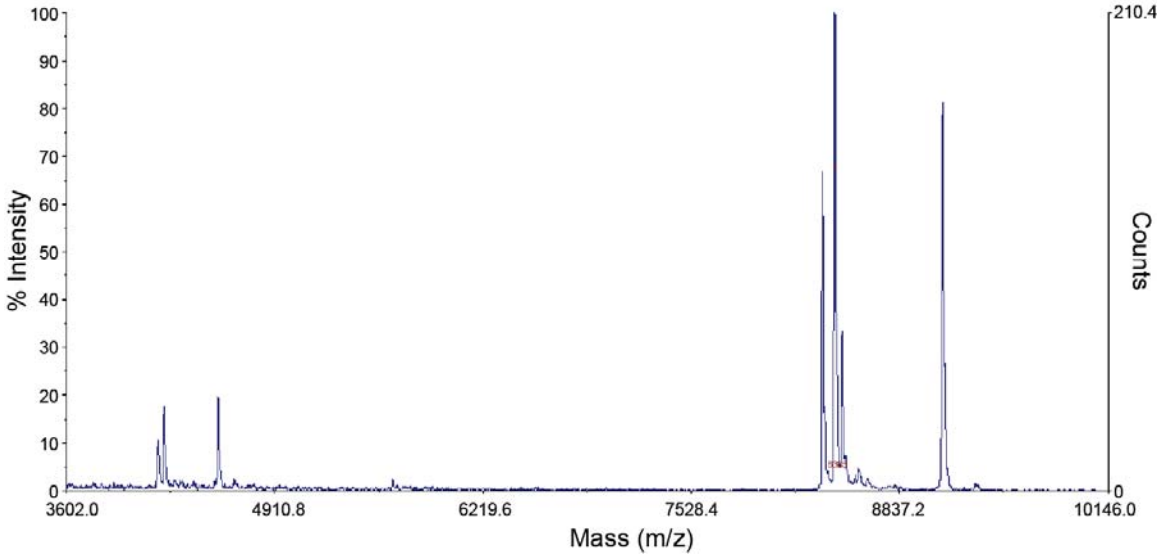

Figure A3. MALDI-TOF mass spectrometry analysis of IF1-Q10C(Cy3). The peak at 9107.46 m/z corresponds to IF1-Q10C(Cy3). The peak at 8353.87 is unlabeled IF1-Q10C, and 8476.78 is the doubly-ionized form of myoglobin, the internal standard.

Appendix B

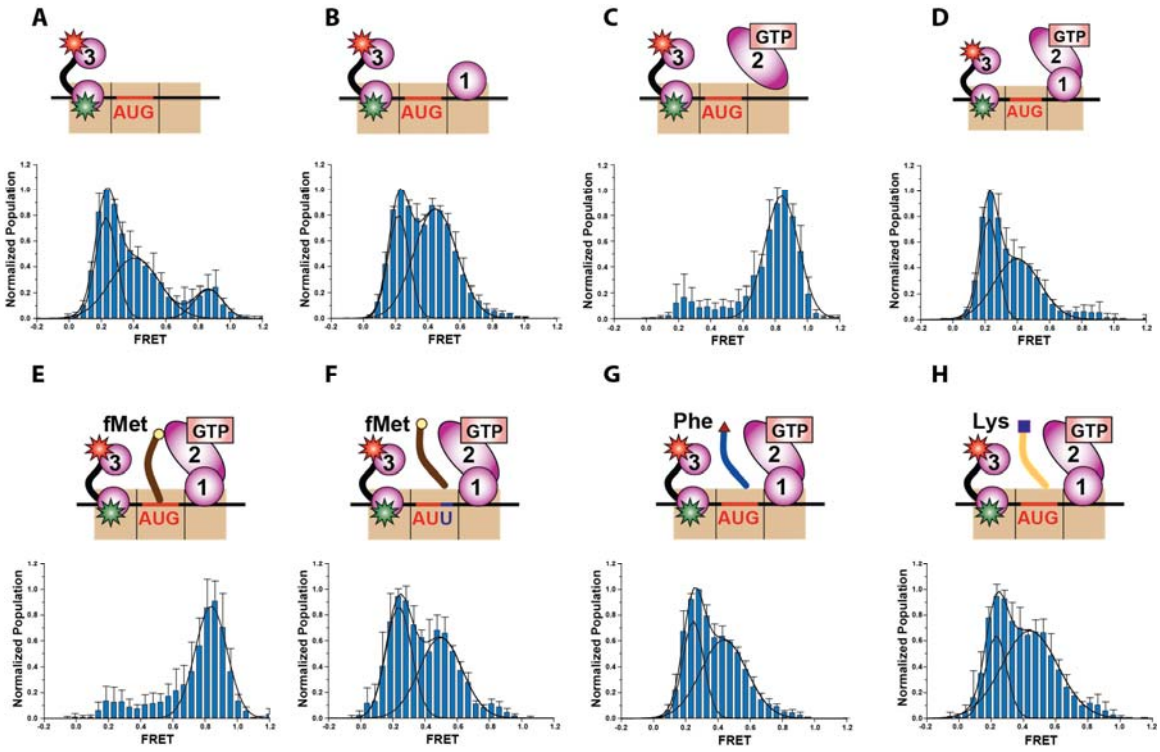

Figure B1 smFRET results are reproducible on a day-to-day basis. Three independent data sets were collected for the eight 30S initiation complexes investigated. Each blue bin represents the average normalized population at that FRET value from three data sets, and the error bars represent the standard deviation from these measurements. Bin size = 0.047 FRET. The histograms were fit with Gaussian distributions (black lines). The cartoons above each population histogram represent the contents of each initiation complex.

Appendix C

Interactions between IF3 and the 30S subunit

Ribosomal proteins or rRNA residues	Method	REF	Notes
S1, S11, S12, S13, S19, S21	Crosslinking: *N,N'-p*-phenylenedimaleimide or dimethylsuberimidate	[1]	Was not possible to distinguish between proteins directly cross-linked to IF3 and those indirectly cross-linked through the mediation of one or more other proteins. S1, S11, S12, S13, S19, S21 are *near* the binding site of IF3.
S12	Crosslinking: DMS	[2]	
S7	Tartaryl diazide	[3]	
S4, S7, S12, S13, S14, S19	Inhibition of IF3 binding by pre-incubation of 30S subunits with antibodies against S4, S7, S12, S13, S14, S19	[4]	
S7, S11, S12, S18, S21	Direct Near-UV (>285 nm) irradiation (crosslinking)	[5]	The r-proteins were identified by immunochemical techniques
S1, S2, S3, S11, S12, S13, S18, S21	Photochemical crosslinking: FMN *p*-nitrobenzylmaleimide (PNBM)	[6]	See note #1
S11, S13, S19	Crosslinking [3H]NEM IF3 to 30S proteins by reaction with dimethylsuberimidate	[7]	Suggests that IF3 bridges the head and platform of the 30S, making contact with both sites on the side of the 30S facing the 50S subunit
S7, S11, S13, S19	Direct cross-linking with 2-iminothiolane	[8]	"this study shows unambiguously the direct cross-linking of S7, S11, and S19 to IF3 and adds S13 as a direct cross-link"
16S rRNA: major site in the central domain: 819-859 minor site near the 3' end: 1506-1529	Cross-linking by transdiamminedichloroplatinum(II)	[9]	"we *unambiguously* identified two distinct RNA regions cross-linked to the factor"
G791	Mutagenesis (G791A), resulted in subunit association defect and 10-fold loss in IF3 affinity for 30S	[10]	G791 is universally conserved. Nucleotides in the region of residue 790 in 16S rRNA are involved in subunit association. It is possible that IF3 controls subunit association by altering the accessibility of the 790 loop in free 30S subunits
A792	Mutagenesis (A792C), lost the ability to bind IF3	[11]	Note #2
G700, G703, G791	Protection from attack by kethoxal	[12]	The binding of IF3 causes virtually complete protection of G700 and G703, and partial protection of G791, from kethoxal attack…these results suggest that IF3 makes direct contact

			with the region of 16S rRNA that includes positions 700 to 703, but that protection of the 790 region is either indirect or due to close proximity (but not direct contact) with the factor. **Note** #3
G1487	Significant enhancement of the reactivity of G1487 toward kethoxal	[12]	
G1530	Mutagenesis (G1530A). Loss of IF3 binding.	[13]	
A1531	Mutagenesis (A1531G). Loss of IF3 binding.	[13]	

#1. "It is, however, surprising that a total of seven proteins (S1, S2, S3, S11, S12, S18, and S21) are found crosslinked to IF3 in our studies with PNBM-IF3 and the previous work with PDM, since the chemical selectivity of the maleimido group should insure that in both of these studies crosslinking to IF3 will occur principally or exclusively via its unique cysteine residue (Cys66)"…this result could be used to argue that IF3 can bind to 30S subunits in more than one relative orientation, due either to multiple 30S conformations (reflecting perhaps 30S heterogeneity), multiple IF3 conformations, or the presence of more than one IF3 binding sites per 30S subunit (binding to such sites would have to be mutually exclusive, since IF3 binds to 30S subunits with 1:1 stoichiometry."

#2. In Laughrea and Tam…"together with the results of Tapprich et al. (1989) and Santer et al. (1990), our data suggests that hairpin 24 promotes IF3 binding through allosteric effects rather than direct interaction."

#3. The 700 and 790 regions of 16S rRNA have been implicated in subunit association by chemical protection. Thus, binding of IF3 to these sites could compete for their interactions with the 50S subunit.

#3 (cont). "IF3 has been localized to the central domain of 16S rRNA by crosslinking (Ehresmann et al.), site-directed mutagenesis (Tapprich et al., 1989), and protection from chemical and enzymatic probes (Muralikrishna & Wickstrom)

IF3 Residues involved in interactions with the 30S subunit

Residue	Perturbation	Results	REF
Tyr109	Labeled *in vitro* with 125I by lactoperoxidase or chloramine T	Loss of binding	[14]
Tyr76	Labeled *in vitro* with 125I by lactoperoxidase or chloramine T	Does not result in loss of biological activity	[14]
Tyr71	Labeled *in vitro* with 125I by lactoperoxidase or chloramine T	Does not impair the 30S binding of IF3 but prevents the functional interaction of the factor with the 30S subunit (test: capacity to promote the dissociation of ternary complexes of poly(U)-N-AcPhe-tRNA-30S ribosomal subunits	[14]
Lys112	Site-specific chemical modification	Essential for ribosome binding	[15]

	with pyridoxal phosphate (PLP) at lysine residues. “In many instances, this reaction proved to be selective for those Lys residues displaying affinity for negatively charged phosphate groups.” The presence of Lys112 in the ribosomal-binding site of IF3 is further supported by the fact that 2 residues away from it, there is a tyrosine residue (Tyr109) which displays essentially similar properties.	(Lys shows affinity for negatively charged phosphate groups) **Note**: Lys 2, 5, 99, 112, & 166 were modified. Only Lys112 is important (major cause for loss of IF3 binding). “concerning the other lysines, it seems unlikely that a role in binding to 30S ribosomes is played by Lys99, since this residue is hardly protected by the 30S ribosomal subunits and its extensive modification does not interfere with IF3 activity”	
Tyr107	Y107F, Y107L	Required for 30S binding, fMet-tRNA-mRNA proofreading	[16]
Lys110	K110R, K110L	Required for 30S binding, fMet-tRNA-mRNA proofreading (Lys110 is more important than Tyr107 for binding)	[16]
A42T G71D D106N L111F R116C M172I P176L	Unable to normally discriminate against several non-canonical initiation codons such as AUU and ACG		[17]
G71	G71D – ribosome binding affinity only marginally affected	NTD mutant probably directly affects the proofreading activity of IF3, although a final proof requires more direct evidence	[17]
D106	D106N – increased IF3 affinity for the 30S	Y107L and K110L (other studies) are also located in this region, thereby demonstrating that this solvent-exposed side of the helix is critical for ribosome-binding.	[17]
R116	R116C	(cont from D106) For both K110 and R116, the most likely explanation is an electrostatic effect, as both these basic residues could promote contacts with the acidic rRNA. **Note** #1	[17]
Glu134Lys	Defect in the fidelity function of IF3		[18]
P176	P176L shows a decreased affinity for the 30S		[17]
Asp106 Gln108 Lys110 Arg112 Leu114 Ile115	Helix H3 (CTD) residues involved in binding to the 30S (NMR titration)		[19]
Leu145 Asn146 Lys149 Leu155	Helix H4 (CTD) residues involved in binding to the 30S (NMR titration)		[19]
Arg99 Gly101 Thr102 Asp103 Gly105	Loop L6 residues involved in binding to the 30S (NMR titration)		
Phe130	Loop L7 residues involved in binding		

His137	to the 30S (NMR titration)		
Thr163 Lys164 Glu166	Loop L8 residues involved in binding to the 30S (NMR titration)		
Thr127	Beta strand B7 residues involved in binding to the 30S (NMR titration)		
Gln169 Ile171 Val173 Leu174 Ala175	Beta strand B9 residues involved in binding to the 30S (NMR titration)		
Glu8 Ala10	N-terminal tail of NTD residues involved in binding to the 30S (NMR titration)		
Asn16 Gly17	Beta strand B1 (NTD) residues involved in binding to the 30S (NMR titration)		
Asp52 Ile56	Beta strand B4 (NTD) residues involved in binding to the 30S (NMR titration)		
Cys65	Beta strand B5 (NTD) residues involved in binding to the 30S (NMR titration)		
Glu18 Gln22	Loop (L1) connecting B1 and B2 (NTD) residues involved in binding to the 30S (NMR titration)		
Lys72 Phe73	C-terminal helix		
Ser78	Start of linker		

#1. D106N…the dramatic change in the ribosome-binding properties of the D106N variant is striking. It is not the result of a structural rearrangement, as an NMR analysis of the structure of the corresponding protein shows that it is unaffected by the mutation. This aspartate residue is one of the eight strictly conserved residues in IF3 (N16, I19, A46, D52, L53, D106, G132 and P176) out of which only two are solvent-exposed. This suggests that D106 plays a key role in specific RNA recognition.

References

1. Heimark, R.L., L. Kahan, K. Johnston, J.W. Hershey, and R.R. Traut, *Cross-linking of initiation factor IF3 to proteins of the Escherichia coli 30 S ribosomal subunit.* J Mol Biol, 1976. **105**(2): p. 219-30.

2. Hawley, D.A., L.I. Slobin, and A.J. Wahba, *The mechanism of action of initiation factor 3 in protein synthesis. II. Association of the 30S ribosomal protein S12 with IF-3.* Biochem Biophys Res Commun, 1974. **61**(2): p. 544-50.

3. van Duin, J., C.G. Kurland, J. Dondon, and M. Grunberg-Manago, *Near neighbors of IF3 bound to 30S ribosomal subunits.* FEBS Lett, 1975. **59**(2): p. 287-90.

4. Gualerzi, C. and C.L. Pon, *Nature of the ribosomal binding site for initiation factor 3 (IF-3).* Biochem Biophys Res Commun, 1973. **52**(3): p. 792-9.

5. MacKeen, L.A., L. Kahan, A.J. Wahba, and I. Schwartz, *Photochemical cross-linking of initiation factor-3 to Escherichia coli 30 S ribosomal subunits.* J Biol Chem, 1980. **255**(21): p. 10526-31.

6. Cooperman, B.S., A. Expert-Bezancon, L. Kahan, J. Dondon, and M. Grunberg-Manago, *IF-3 crosslinking to Escherichia coli ribosomal 30 S subunits by three different light-dependent procedures: identification of 30 S proteins crosslinked to IF-3--utilization of a new two-stage crosslinking reagent, p-nitrobenzylmaleimide.* Arch Biochem Biophys, 1981. **208**(2): p. 554-62.

7. Pon, C.L., R.T. Pawlik, and C. Gualerzi, *The topographical localization of IF3 on Escherichia coli 30 S ribosomal subunits as a clue to its way of functioning.* FEBS Lett, 1982. **137**(2): p. 163-7.

8. Boileau, G., P. Butler, J.W. Hershey, and R.R. Traut, *Direct cross-links between initiation factors 1, 2, and 3 and ribosomal proteins promoted by 2-iminothiolane.* Biochemistry, 1983. **22**(13): p. 3162-70.

9. Ehresmann, C., H. Moine, M. Mougel, J. Dondon, M. Grunberg-Manago, J.P. Ebel, and B. Ehresmann, *Cross-linking of initiation factor IF3 to Escherichia coli 30S ribosomal subunit by trans-diamminedichloroplatinum(II): characterization of two cross-linking sites in 16S rRNA; a possible way of functioning for IF3.* Nucleic Acids Res, 1986. **14**(12): p. 4803-21.

10. Tapprich, W.E., D.J. Goss, and A.E. Dahlberg, *Mutation at position 791 in Escherichia coli 16S ribosomal RNA affects processes involved in the initiation of protein synthesis.* Proc Natl Acad Sci U S A, 1989. **86**(13): p. 4927-31.

11. Santer, M., E. Bennett-Guerrero, S. Byahatti, S. Czarnecki, D. O'Connell, M. Meyer, J. Khoury, X. Cheng, I. Schwartz, and J. McLaughlin, *Base changes at position 792 of Escherichia coli 16S rRNA affect assembly of 70S ribosomes.* Proc Natl Acad Sci U S A, 1990. **87**(10): p. 3700-4.

12. Moazed, D., R.R. Samaha, C. Gualerzi, and H.F. Noller, *Specific protection of 16 S rRNA by translational initiation factors.* J Mol Biol, 1995. **248**(2): p. 207-10.

13. Firpo, M.A., M.B. Connelly, D.J. Goss, and A.E. Dahlberg, *Mutations at two invariant nucleotides in the 3'-minor domain of Escherichia coli 16 S rRNA affecting translational initiation and initiation factor 3 function.* J Biol Chem, 1996. **271**(9): p. 4693-8.

14. Bruhns, J. and C. Gualerzi, *Structure--function relationship in Escherichia coli initiation factors: role of tyrosine residues in ribosomal binding and functional activity of IF-3.* Biochemistry, 1980. **19**(8): p. 1670-6.

15. Ohsawa, H. and C. Gualerzi, *Structure-function relationship in Escherichia coli initiation factors. Identification of a lysine residue in the ribosomal binding site of initiation factor by site-specific chemical modification with pyridoxal phosphate.* J Biol Chem, 1981. **256**(10): p. 4905-12.

16. De Bellis, D., D. Liveris, D. Goss, S. Ringquist, and I. Schwartz, *Structure-function analysis of Escherichia coli translation initiation factor IF3: tyrosine 107 and lysine 110 are required for ribosome binding.* Biochemistry, 1992. **31**(48): p. 11984-90.

17. Sacerdot, C., E. de Cock, K. Engst, M. Graffe, F. Dardel, and M. Springer, *Mutations that alter initiation codon discrimination by Escherichia coli initiation factor IF3.* J Mol Biol, 1999. **288**(5): p. 803-10.

18. Sacerdot, C., C. Chiaruttini, K. Engst, M. Graffe, M. Milet, N. Mathy, J. Dondon, and M. Springer, *The role of the AUU initiation codon in the negative feedback regulation of the gene for translation initiation factor IF3 in Escherichia coli.* Mol Microbiol, 1996. **21**(2): p. 331-46.

19. Sette, M., R. Spurio, P. van Tilborg, C.O. Gualerzi, and R. Boelens, *Identification of the ribosome binding sites of translation initiation factor IF3 by multidimensional heteronuclear NMR spectroscopy.* RNA, 1999. **5**(1): p. 82-92.

Appendix D

Bacterial strain collection

No.*	Strain	Plasmid	Resistance	Remarks
1	BL21(DE3)-IF3	pProEX-HTb	Amp	Dated: 20 Dec 07
2	BL21(DE3)-IF1	pProEX-HTb	Amp	Dated: 21 Jul 08
3	16S A h6 mt SQ	pKK3535	Amp/Cam	† hp1 insertion.
4	16S A h6 mt DH5α	pKK3535	Amp	† hp1 insertion. Sequence confirmed.
5	16S B h33a SQ380	pKK3535	Amp	†
6	16S B h33a DH5α	pKK3535	Amp	†
7	16S C h44 mt SQ380	pKK3535	Amp	† hp5 insertion.
8	16S C h44 mt DH5α	pKK3535	Amp	† hp5 insertion. Sequence confirmed.
9	16S D h39 mt SQ380	pKK3535	Amp/Cam	† hp1 insertion.
10	16S D h39 mt DH5α	pKK3535	Amp	† hp1 insertion. Sequence confirmed.
11	16S F h10 mt SQ380	pKK3535	Amp	† hp5 insertion.
12	16S F h10 mt DH5α	pKK3535	Amp	† hp5 insertion. Sequence confirmed.
13	BL21(DE3)-IF3-C66S	pProEX-HTb	Amp	
14	BL21(DE3)-IF3-R41C	pProEX-HTb	Amp	
15	BL21(DE3)-IF3-E45C	pProEX-HTb	Amp	
16	BL21(DE3)-IF3-E62C	pProEX-HTb	Amp	
17	BL21(DE3)-IF3-C66S-K98C	pProEX-HTb	Amp	Dated: 01 Sept 08
18	BL21(DE3)-IF3-C66S-M143C	pProEX-HTb	Amp	Dated: 01 Sept 08
19	BL21(DE3)-IF3-M143C-C66A	pProEX-HTb	Amp	
20	BL21(DE3)-IF3-K98C-C66A	pProEX-HTb	Amp	
21	BL21(DE3)-IF1-Q10C	pProEX-HTb	Amp	Dated: 09 Apr 07
22	DH5α-IF1(RG21 strain)	pRS3559	Amp	Dated 30 May 07; See Ruben's Strains
23	BL21(DE3)-IF3-Q23C	pProEX-HTb	Amp	
24	BL21(DE3)-IF3-L30C	pProEX-HTb	Amp	
25	BL21(DE3)-IF3-S39C	pProEX-HTb	Amp	
26	BL21(DE3)-IF3-S58C	pProEX-HTb	Amp	
27	BL21(DE3)-IF3-C66S-S39C	pProEX-HTb	Amp	
28	BL21(DE3)-IF3-C66S-S58C	pProEX-HTb	Amp	
29	BL21(DE3)-IF3-C66S-K98C-S39C	pProEX-HTb	Amp	
30	BL21(DE3)-IF3-C66S-K98C-L30C	pProEX-HTb	Amp	
31	BL21(DE3)-IF3-C66S-K98C-S58C	pProEX-HTb	Amp	
32	Avi-tag His-tag IF3-C66S-K98C in BL21	pProEX-HTb	Amp	Prepared by Dr. Armando del Rio**
33	Avi-tag His-tag IF3-C66S-S39C in BL21	pProEX-HTb	Amp	Prepared by Dr. Armando del Rio**
34	Avi-tag His-tag IF3-C66S-S39C-K98C in BL21	pProEX-HTb	Amp	Prepared by Dr. Armando del Rio**
35	KLF 1085 wt SQZ10 containing p278	pKF207	Amp	From K. Fredrick (D. Qin); See Ref. [1]
36	KLF 1106 SQZ10 containing p278 A790G	pKF207	Amp	From K. Fredrick (D. Qin); See Ref. [1]
37	KLF 1104 SQZ10 containing p278 G1338A	pKF207	Amp	From K. Fredrick (D. Qin); See Ref. [1]
38	BL21(DE3)-IF3-C66S-S39C-K98C-Y75N	pProEX-HTb	Amp	Prepared by Anthony Testa
39	BL21(DE3)-RRF	pProEX-HTb	Amp	Plasmid from S. Sternberg plasmid #76
40	B105 C1G A72C	pUC13		MetY mutant from Jingyi; See Refs. [2, 3]
41	B105 G32A C40T	pUC13		MetY mutant from Jingyi; See Refs. [2, 3]
42	B105 A11C T25G	pUC13		MetY mutant from Jingyi; See

				Refs. [2, 3]

*Location in marked freezer box

†16S rRNA extension mutants from the Puglisi lab. See Ref. [4]. I have only sequenced the plasmids from the DH5α strains, not the Squires strains.

**A His tag is on the N terminus and an Avi tag on the C terminus. All the mutants are in pProEX HTb, inserted between EheI and HindIII restriction sites. There is a linker between the C terminal of IF3 and the Avi tag sequence (15 AA): A G T P G A P T G I G T D Y D. The avi tag is: M S G L N D I F E A Q K I E W H E D T. Then, there is a 26 AA between the last AA in IF3 and the K in the Avi sequence that binds the biotin.

References

1. Qin, D., N.M. Abdi, and K. Fredrick, *Characterization of 16S rRNA mutations that decrease the fidelity of translation initiation.* RNA, 2007. **13**(12): p. 2348-55.

2. Mandal, N. and U.L. RajBhandary, *Escherichia coli B lacks one of the two initiator tRNA species present in E. coli K-12.* J Bacteriol, 1992. **174**(23): p. 7827-30.

3. Fei, J., A.C. Richard, J.E. Bronson, and R.L. Gonzalez, Jr., *Transfer RNA-mediated regulation of ribosome dynamics during protein synthesis.* Nat Struct Mol Biol, 2011.

4. Dorywalska, M., S.C. Blanchard, R.L. Gonzalez, H.D. Kim, S. Chu, and J.D. Puglisi, *Site-specific labeling of the ribosome for single-molecule spectroscopy.* Nucleic Acids Res, 2005. **33**(1): p. 182-9.

Appendix E

FRET Data Analysis Package (FDAP) v1.7 Matlab scripts

The following FDAP v.1.7 scripts were written by Mr. Pallav Kosuri.

E1. loadTraces.m

```
    %Loads traces from multiple files with filenames formatted as
"<filename>#.dat",
    %where # is the index number. The traces are returned as columns in
a
    %matrix with the first row as a label formatted as: "<file#>*1000 +
<trace#>"

function traces = loadTraces(filename, no)

        %one file
    if nargin == 1
        newTraces=load([filename '.dat']);
        ms=size(newTraces);
            %create labels as [1 1 2 2 3 3 ...]
        labels=1:ms(2)/2;
        labels=[labels; labels];
        labels=labels(:)';
            %insert labels as the first row
        newTraces=[labels; newTraces];
        traces=newTraces;

        %multiple files
    elseif nargin == 2

            %first file
        i=1;
        reading_file=i
        newTraces=load([filename int2str(i) '.dat']);
        ms=size(newTraces);
            %create labels as [1 1 2 2 3 3 ...]
        labels=i*1000+(1:ms(2)/2);
        labels=[labels; labels];
        labels=labels(:)';
            %insert labels as the first row
        newTraces=[labels; newTraces];
        traces=newTraces;


            %if more than one file
        if no > 1
            for i = 2:no
                reading_file=i
                newTraces=load([filename int2str(i) '.dat']);
```

```
                ms=size(newTraces);
                    %create labels as [1 1 2 2 3 3 ...]
                labels=i*1000+(1:ms(2)/2);
                labels=[labels; labels];
                labels=labels(:)';
                    %insert labels as the first row
                newTraces=[labels; newTraces];

                traces=[traces newTraces];
            end
        end
    end

    ms=size(traces);
    number_of_traces=ms(2)/2
```

E2. separateCy.m

```
function [cy3, cy5] = separateCy(in)

    ms=size(in);
    for i = (1:ms(2)/2)
        cy3(:,i)=in(:,i*2-1);
        cy5(:,i)=in(:,i*2);
    end

    traces=ms(2)/2
```

E3. P_filter.m

```
    %takes the cy3 and cy5 traces as arguments and
    %returns the filtered traces - only the traces
    %with P-values lower than or equal to the cutoff are kept.
    %Also plots a histogram. The counting shows which
    %trace is currently analyzed.

function [cy3, cy5, P] = P_Filter(cy3, cy5)

    [C, P, mPl]=evalTraces(cy3, cy5);

    CUTOFF=0.5

    disp('Applying filter...')

    labels=cy3(1,:);

        %selective filter
    n=0;
   % j=1
```

```
    for i = 1:length(P)
        if P(i) > CUTOFF
       %     cy3d(:, j)=cy3(:, i);
        %    cy5d(:, j)=cy5(:, i);
            [cy3, cy5]=removeTrace(cy3, cy5, labels(i));
            n=n+1;
        %    j=j+1;
        end
    end

        %display results
    total=length(P)
    discarded=n
    kept=total-discarded

    plotP(P);
```

E4. plotTraces.m

```
function plotTraces(cy3, cy5, len1, len2)

    [cy3, cy5, labels]=removeLabels(cy3, cy5);

    ms=size(cy3);
    cols=ms(2);
    rows=ms(1);

        %no photobleach specified
    if nargin==2
        j=1;
        while j <= cols
            figure;
            for n = 1:9
                subplot(3,3,n);
                plot(cy3(:,j), 'g'), hold on, plot(cy5(:,j), 'r');
                axis([0 rows -200 4000]);
                title(labels(j));
                j=j+1;
                if j > cols, break, end
            end
        end

        %1 photobleach event specified
    elseif nargin==3
        tLength=0;

            %remove labels
        ms1=size(len1);
        if ms1(1) == 2
            ind=len1(1,:);
```

```
        for i = 1:length(ind)
            tLength(i)=len1(2,find(labels==ind(i)));
        end
    else
        tLength=len1;
    end

        %plot traces
    j=1;
    while j <= cols
        figure;
        for n = 1:9
            subplot(3,3,n);
            plot(cy3(:,j), 'g'), hold on, plot(cy5(:,j), 'r');
            title(labels(j));
            v=axis;
            plot([tLength(j) tLength(j)], [v(3) v(4)], '--')
            j=j+1;
            if j > cols, break, end
        end
    end

    %2 photobleach events specified
elseif nargin==4
    tLength1=0;
    tLength2=0;

        %remove labels
    ms1=size(len1);
    if ms1(1) == 2
        ind=len1(1,:);
        for i = 1:length(ind)
            tLength1(i)=len1(2,find(labels==ind(i)));
        end
    else
        tLength1=len1;
    end
    ms2=size(len2);
    if ms2(1) == 2
        ind=len2(1,:);
        for i = 1:length(ind)
            tLength2(i)=len2(2,find(labels==ind(i)));
        end
    else
        tLength2=len2;
    end

        %plot traces
    j=1;
    while j <= cols
        figure;
        for n = 1:9
```

```
                subplot(3,3,n);
                plot(cy3(:,j), 'g'), hold on, plot(cy5(:,j), 'r');
                title(labels(j));
                v=axis;
                plot([tLength1(j) tLength1(j)], [v(3) v(4)], '--b')
                plot([tLength2(j) tLength2(j)], [v(3) v(4)], '--k')
                j=j+1;
                if j > cols, break, end
            end
        end
    end

    [cy3, cy5]=addLabels(cy3, cy5, labels);
```

E5. J_filter.m

```
    %takes the cy3 and cy5 traces as well as a vector with the
    %J-selections (bad traces) as arguments and returns the filtered
traces

function [cy3, cy5] = J_Filter(cy3, cy5, J)

    ms=size(cy3);

        %selective filter
    for i = 1:length(J)
        [cy3, cy5]=removeTrace(cy3, cy5, J(i));
    end

        %display results
    ms2=size(cy3);
    total=ms(2)
    discarded=ms(2)-ms2(2)
    kept=total-discarded
```

E6. S_Filter.m

```
    %filters out the traces with a Cy3 photobleach event before "minL"
    %datapoints.

function [cy3, cy5] = S_Filter(cy3, cy5, minL)

        %correct for the 4 datapoints kept after photobleaching
    minL=minL+4;

        %detect Cy3 photobleach times
    [cy3d, cy5d]=deleteCy3photobleach(cy3, cy5);
    [cy3d, cy5d, labels]=removeLabels(cy3d, cy5d);
    d3t=traceLength(cy3d);
```

```
    %selective filter
n=0;
for i = 1:length(d3t)
    if d3t(i) < minL
        [cy3, cy5]=removeTrace(cy3, cy5, labels(i));
        n=n+1;
    end
end

    %display results
total=length(d3t)
discarded=n
kept=total-discarded
```

E7. correctBaseline.m

```
function [cy3, cy5] = correctBaseline(cy3, cy5)

    BLEED_COEF = 0.07

    [cy3, cy5, labels]=removeLabels(cy3, cy5);

    X = linspace(600, 3000, 400);

    H3 = hist(cy3(:), X);
        %truncation correction
    H3(length(H3))=H3(length(H3)-1);
    figure, bar(X, H3), title('Cy3 Intensity Histogram');

        %Cy3 baseline correction
    Cy3_Baseline = X(find(H3==max(H3)))
    cy3 = cy3 - Cy3_Baseline;

        %Cy5 bleedthrough correction
    cy5 = cy5 - BLEED_COEF*cy3;

    H5 = hist(cy5(:), X);
        %truncation correction
    H5(length(H5))=H5(length(H5)-1);
    figure, bar(X, H5), title('Cy5 Intensity Histogram');

        %Cy5 baseline correction
    Cy5_Baseline = X(find(H5==max(H5)))
    cy5 = cy5 - Cy5_Baseline;

    [cy3, cy5]=addLabels(cy3, cy5, labels);
```

E8. correctBaseline_end.m

```
%the argument noBleach is a list of all the traces that do not show
%photobleaching

function [cy3, cy5] = correctBaseline_end(cy3, cy5, noBleach)

        %AVER specifies how many datapoints at the end of the trace
that
        %are averaged to get the baseline.
    BLEED_COEF = 0.07
    AVER = 10

    [cy3, cy5, labels]=removeLabels(cy3, cy5);

    ms = size(cy3);

        %individually correct all traces that show photobleaching
    for i = 1:ms(2)
        if find(noBleach==labels(i))
            cy3_baseline(i)=0;
            cy5_baseline(i)=0;
        else
            disp(labels(i))

            cy3_baseline(i) = mean(cy3((ms(1)-AVER):ms(1), i));
            cy3(:, i) = cy3(:, i) - cy3_baseline(i);

            cy5(:, i) = cy5(:, i) - BLEED_COEF*cy3(:, i);
            cy5_baseline(i) = mean(cy5((ms(1)-AVER):ms(1), i));
            cy5(:, i) = cy5(:, i) - cy5_baseline(i);
        end
    end

        %correct all non-photobleaching traces using an average
    cy3_average =
mean(cy3_baseline)*length(cy3_baseline)/(length(cy3_baseline)-
length(noBleach));
    cy5_average =
mean(cy5_baseline)*length(cy5_baseline)/(length(cy5_baseline)-
length(noBleach));
    no = find(cy3_baseline == 0);
    cy3(:, no) = cy3(:, no) - cy3_average;
    cy5(:, no) = cy5(:, no)-BLEED_COEF*cy3(:, no)- cy5_average;
    cy3_baseline
    [cy3, cy5]=addLabels(cy3, cy5, labels);

%    base_list = [labels; cy3_baseline; cy5_baseline];
```

E9. saveTraces.m

```
function saveTraces(cy3, cy5, filen)
```

```
    merged=mergeCy(cy3, cy5);
    save(filen, 'merged',  '-ASCII');
```

E10. plotFRET.m

```
    %calculates FRET and plots a histogram with 'n' bins.
    %returns the histogram data as a two-row matrix
    %with x-values in row 1 and y-values in row 2

function FH = plotFRET(cy3, cy5, n)

    [cy3, cy5]=deleteCy3photobleach(cy3, cy5);
    F=getFRET(cy3, cy5);

    X=linspace(-0.2,1.2,n);
    H=hist(F(:),X);
        %normalize
    H=H/max(H);
    figure, bar(X, H, 'k');
    axis([-0.2 1.2 0 max(H)*1.05]);
    xlabel('FRET');
    ylabel('normalized frequency');
    title('FRET histogram')

    FH=[X; H];
```

E11. plotTimeFRET.m

```
    %Generates and plots a 2D histogram of the FRET time evolution.
    %'FRETbins' and 'Tbins' are the number of bins in each dimension.
    %cutoffT is the cutoff time in datapoints. If no cutoff time is
given,
    %no cutoff is applied.

function [T, Y, H] = plotTimeFRET(cy3, cy5, FRETbins, Tbinsize,
cutoffT)

        %Exposure time = 50 ms
    FPS = 10;

        %Minimum/maximum count shown
    MINCOUNT = 1;
    MAXCOUNT = 16;
    RESOLUTION = 800;

        %calculate FRET
    F = getFRET(cy3, cy5);
    ms=size(F);
```

```
    %generate the histogram
T = (Tbinsize:Tbinsize:ms(1))'/FPS;
Y = linspace(-0.2, 1.2, FRETbins)';
H = getTimeFRET(F, FRETbins, Tbinsize);

    %apply cutoff in time dimension
if nargin == 5
    cutoffT = cutoffT*FPS;
    T = T(1:min(floor(cutoffT/Tbinsize), ms(1)));
    H = H(:, 1:min(floor(cutoffT/Tbinsize), ms(1)));
end

TI = linspace(min(T), max(T), RESOLUTION);
YI = linspace(-0.2, 1.2, RESOLUTION);
HI = interp2(T', Y, H, TI', YI, 'cubic');

    %plot figure
figure, pcolor(TI', YI, HI);
colormap([1 1 0.8; ones(15, 3); JET]);
hold on

MAXCOUNT = max(max(H))*13.2;

    %minimum intensity: MINCOUNT
    %minimum intensity: MAXCOUNT
caxis([MINCOUNT MAXCOUNT]);
axis([min(T) max(T) min(Y) max(Y)]);

    %plot contour lines
%contour(T, Y, H, (MINCOUNT:1:MAXCOUNT))

colorbar; shading interp; axis tight square;

    %add labels
xlabel('T (seconds)') ; ylabel('FRET') ; title('FRET Time Evolution
Histogram') ;

    %set tick marks between data points
%set(gca,'XTick',([2*T(1):2*(T(2)-T(1)):max(T)] + T(1)/2))
%set(gca,'XTickLabel',T)
```

Dwell time analysis: “lifetime analysis_W&L” scripts

These scripts were written by Dr. Jiangning Wang.

E12. getRawDwell.m

```
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%5
```

```
% group and stack the data into the form [ fileno(i)  FRET(i) FRET(i+1)
n ]
% remove the first and last dwell states in each trace.
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
function dwellData = getRawDwell(pathData)

    fileno = pathData(:, 1);
    FRET = pathData(:, 2);

        %transition no
    t = 1;
        %frame no in current transition
    n = 1;
    for (i=(1:(length(FRET)-1)))

            %new file
        if diff(fileno(i:i+1))
            %dwellData(t, :) = [fileno(i) FRET(i) NaN n];
            %t = t + 1;
            n = 1;
            continue;          % skip the last dwell state of each
trace without saving them, equals to remove the last dwell state
        end

            %transition
        if diff(FRET(i:i+1))
            dwellData(t, :) = [fileno(i) FRET(i) FRET(i + 1) n];
            t = t + 1;
            n = 1;
            %no transition
        else
            n = n + 1;
        end
    end

    rows = [2:t-1];

    for k=2:t-1
       if dwellData(k-1,1)<dwellData(k,1)  % remove the first dwell
state of each trace
           rows = removeEntry(rows, k);
       end
    end

    dwellData = dwellData(rows', :);
```

E13. purifyRawDwell.m

```
% purify traces by deleting or combining single datapoint transtions.
```

```
function dwellData = purifyRawDwell(dwellData)

    filecol = dwellData(:,1);
    sizedw = size(dwellData);
    filemax = dwellData(sizedw(1),1);      % max file number

    minorind = zeros(1,1);  % record the rows to be deleted

    for i=1:filemax
        datascope = find(filecol==i);
        if length(datascope)==1 & dwellData(datascope, 4)==1
            minorind = [minorind; datascope];
        elseif length(datascope)>=2
            singlefile = dwellData(datascope, :);   % singlefile = a
specific file
            for j=1:length(datascope)
                tempind = datascope(j);
                if dwellData(tempind,4) == 1        % single datapoint
transition
                    if j == 1                       % if the single
datapoint transition is at the beginning of the trace, add it to the
following state
                        minorind = [minorind; datascope(j)];
                        dwellData(tempind,4) = -1;
                        dwellData(datascope(j+1),4) =
dwellData(datascope(j+1),4)+1;
                    elseif j == length(datascope)   % if the single
datapoint transition is at the beginning of the trace, add it to the
previous state
                        minorind = [minorind; datascope(j)];
                        dwellData(tempind,4) = -1;
                        k = j-1;
                        while dwellData(datascope(k),4)<0
                            k = k-1;
                        end
                        dwellData(datascope(k),4) =
dwellData(datascope(k),4)+1;
                    else                            % if the single
point transition is in the middle, add it to the closer state of the
two sides.
                        minorind = [minorind; datascope(j)];
                        dwellData(tempind,4) = -1;
                        diffpre = abs(dwellData(datascope(j-1),2)-
dwellData(datascope(j),2));
                        diffpos = abs(dwellData(datascope(j),2)-
dwellData(datascope(j+1),2));
                        if diffpre <= diffpos
                            k = j-1;
                            while dwellData(datascope(k),4)<0
                                k = k-1;
                            end
```

```
                            dwellData(datascope(k),4) =
dwellData(datascope(k),4)+1;
                            dwellData(datascope(k),3) =
dwellData(datascope(j),3);
                        else
                            k = j+1;
                            while dwellData(datascope(k),4)<0 &
k<length(datascope)
                                k = k+1;
                            end
                            dwellData(datascope(k),4) =
dwellData(datascope(k),4)+1;
                            m = j-1;
                            while dwellData(datascope(m),4)<0
                                m = m-1;
                            end
                            dwellData(datascope(m),3) =
dwellData(datascope(k),2);
                        end
                    end
                end
            end
        end
    end

    rows = [1: sizedw(1)];
    minorlen = length(minorind);
    minorind = minorind(2:minorlen, 1);
    for k = minorind.'
        rows = removeEntry(rows, k);
    end

    dwellData = dwellData(rows', :);
```

E.14. purifyOnOffDwell.m

```
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% combine all the successive high and low FRET states(defined by the
threshold)
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
function dwellData = purifyOnOffDwell(dwellData, Filter)

    if nargin==2
        KEY = Filter;
    else
        KEY = 0.20;
    end

    dwellData = dwellData(:,[2,3,4]);
```

```
    sizedw = size(dwellData);
    selind = 1;

    for t = 1:sizedw(1)-1
        if dwellData(t,1)>KEY
            if dwellData(t,2)>KEY && dwellData(t+1,1)>KEY
                selind = [selind; t];
                dwellData(t+1,3) = dwellData(t,3)+dwellData(t+1,3);
            end
        else
            if dwellData(t,2)<=KEY && dwellData(t+1,1)<=KEY
                selind = [selind; t];
                dwellData(t+1,3) = dwellData(t,3)+dwellData(t+1,3);
            end
        end
    end

    lensel = length(selind);
    if selind(2)==1
        selind = selind(2:lensel);
    end

    mergeind = selind;
    lenmerg = length(mergeind);

    rowind = [1:sizedw(1)];
    for k = 1:lenmerg
        rowind = removeEntry(rowind, mergeind(k));
    end
    dwellData = dwellData(rowind, :);

    % There's basicly no use for the following code which removes the
row with 'NAN'

    newsize = size(dwellData);
    nanind = 0;
    for i=1:newsize(1)
        if isnan(dwellData(i,2))
                nanind = [nanind; i];
        end
    end

    newlen = length(nanind);
    nanind = nanind(2:newlen);
    newrowind = [1:newsize(1)];

    for k = 1:newlen-1
        newrowind = removeEntry(newrowind, nanind(k));
    end
```

```
    dwellData = dwellData(newrowind, :);
```

E.15. getDecay.m

```
function [ts, N] = getDecay(dwellData, bounds, cutoff_t)

    start_low = bounds(1);
    start_high = bounds(2);
    end_low = bounds(3);
    end_high = bounds(4);

    selection = find(dwellData(:,1) > start_low & dwellData(:,1) <
start_high & dwellData(:,2) > end_low & dwellData(:,2) < end_high);
    maxT = max(dwellData(selection,3));
    t = 1:min((maxT-1), cutoff_t/0.1);
    ts = t*0.1;
    for i = t
        N(i) = length(find(dwellData(selection, 3) > t(i)));
    end

    figure, plot(ts, N, '.')
    title(['Decay curve for state bounded by: start=['
num2str(start_low) ', ' num2str(start_high) '], end=[' num2str(end_low)
', ' num2str(end_high) ']'])
    ylabel('Population')
    xlabel('t (s)')
    DC=[ts;N].'
    save('decaycurvedata.dat', '-ascii', 'DC');
```

Dwell time analysis: FDAP v1.7 “lifetime” scripts

These script was written by Mr. Pallav Kosuri.

E.16. plotTDP.m

```
    %type:
    %caxis([3500 7000])
    %to set the color limits

function [X, Y, Z] = plotTDP(dwellData, res)

        %size of gaussians in TDP
    VAR = 0.00075

    RESOLUTION = 800;

    X = linspace(-0.2, 1.2, res)';
    Y = X';

        %remove NaN transitions
    n = 1;
```

```
    ms = size(dwellData);
    while n <= ms(1)
        if isnan(dwellData(n, 2))
            if n == 1
                dwellData = dwellData(2:ms(1), :);
            elseif n == ms(1)
                dwellData = dwellData(1:(n-1), :);
                break;
            else
                dwellData = [dwellData(1:(n-1), :);
dwellData((n+1):ms(1), :)];
            end
            ms = size(dwellData);
        else
            n = n + 1;
        end
    end

        %start and stop vectors
    start = dwellData(:, 1);
    stop = dwellData(:, 2);

    size(start)

        %build TDP function
    for j = (1:res)
        for i = (1:res)
            Z(j, i) = sum((1/(2*pi*VAR))*exp(-((X(i) - start).^2 +
(Y(j) - stop).^2)/(2*VAR)));
        end
    end

    %figure, pcolor(X, Y, Z), colormap([0 0 0; JET]), shading flat,
axis square tight

        %interpolate
    XI = linspace(-0.2, 1.2, RESOLUTION);
    ZI = interp2(X, Y, Z, XI', XI, 'cubic');
    figure, pcolor(XI', XI, ZI), colormap([1 1 1; JET]), shading flat,
axis square tight
```