Single-Molecule Analysis of Ribosome and Initiation Factor Dynamics during the Late Stages of Translation Initiation

Daniel David MacDougall

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

COLUMBIA UNIVERSITY

2012

Abstract

Single-Molecule Analysis of Ribosome and Initiation Factor Dynamics during the Late Stages of Translation Initiation

Daniel David MacDougall

Protein synthesis in all organisms is catalyzed by a highly-conserved ribonucleoprotein macromolecular machine known as the ribosome. Prior to each round of protein synthesis in the cell, a functional ribosomal complex is assembled from its component parts at the start site of a messenger RNA (mRNA) template during the process of translation initiation. In bacteria, rapid and high-fidelity translation initiation is promoted by three canonical initiation factors: IF1, IF2, and IF3. In this thesis, I report the use of single-molecule fluorescence methods to study the role of the initiation factors and ribosome-factor interactions in regulating molecular events that occur during late stages of the translation initiation pathway.

In Chapter 1, I provide a structural and biochemical framework for understanding one of the key events of the initiation pathway: docking of the large (50S) ribosomal subunit with the small subunit 30S initiation complex (30S IC). The 50S subunit joining reaction is catalyzed by GTP-bound IF2 and results in formation of a 70S initiation complex (70S IC) that contains an initiator transfer RNA (tRNA) and is primed for formation of the first peptide bond. During 50S subunit joining, IF2-GTP establishes interactions with RNA and protein components of the 50S subunit's GTPase-associated center (GAC), which play an important role in subunit recruitment as well as the subsequent activation of GTP hydrolysis by IF2.

In Chapter 2, I describe the development of a single-molecule fluorescence resonance energy transfer (smFRET) signal to monitor the interactions between IF2 and the ribosome's GAC during real-time 50S subunit joining reactions. Specifically, the role of the L11 region, comprising ribosomal protein L11 and its associated ribosomal RNA (rRNA) helices, was investigated. The L11 region is a prominent structural component of the GAC that is believed to undergo large-scale conformational changes during protein synthesis; however, the nature and timescale of these conformational dynamics, and their role in regulating the biochemical activities of IF2 during initiation, are not known. I demonstrate that my smFRET-based 50S subunit joining assay is sensitive to conformational rearrangements between IF2 and L11 within the 70S IC and can thus be used as a tool for characterizing GAC dynamics and elucidating their function during initiation. Furthermore, my smFRET approach is shown to provide information on the rate of 50S subunit joining as well as the rate of IF2 dissociation from the 70S IC. Notably, IF2-dependent GTP hydrolysis was found to influence the extent of 70S IC conformational dynamics as well as the dissociation rate of IF2.

The role of IF3 in regulating 50S-subunit joining dynamics is discussed in Chapter 3. IF3 plays an important role in ensuring the fidelity of translation initiation by preventing the formation of initiation complexes containing a non-initiator tRNA and/or a non-canonical mRNA start codon. Inclusion of IF3 within the 30S IC in the smFRET experiments was found to render the IF2-catalyzed 50S subunit joining reaction highly reversible. Direct observation of repetitive docking and undocking of the 50S subunit with the 30S IC indicates that IF3 may modulate translation initiation efficiency by influencing the stability of the 70S IC. The individual 50S subunit docking events were

found to result in the formation of very different classes of 70S IC, characterized by different stabilities and unique patterns of IF2-L11 interactions. I propose that these dynamics reflect an underlying conformational equilibrium of the IF3-bound 30S IC that is read out during 50S subunit joining, and that this equilibrium could be modulated in order to regulate the efficiency of translation initiation.

Following initiation-factor mediated assembly of the 70S IC, the first aminoacyl-tRNA is delivered to the ribosome in ternary complex with elongation factor Tu (EF-Tu) and GTP. Accommodation of aminoacyl-tRNA into the ribosome's peptidyl transferase center leads to formation of the first peptide bond, which signals the end of initiation and entry into the elongation phase of protein synthesis. The ternary complex binding site on the ribosome overlaps with that of IF2 at the GAC; a question of key mechanistic importance in understanding how the ribosome coordinates the transition from initiation to elongation thus concerns the relative timing of ternary complex binding with respect to IF2 dissociation from the 70S IC. In Chapter 4, I present preliminary results from two- and three-color fluorescence co-localization experiments aimed at characterizing the timing of these events at the single-molecule level. The data strongly suggest the occurrence of simultaneous occupancy of the ribosome by IF2 and ternary complex, implying that the ribosome is structurally capable of recruiting ternary complex prior to IF2 release from the 70S IC. The observation that the ribosome can accommodate more than one translation factor at a time may have important implications for understanding how it efficiently coordinates factor binding and release throughout protein synthesis, and opens the door to mechanistic studies of the ribosomal L7/L12 stalk, presumed to play a prominent role in these processes.

Table of Contents

List of Figures

List of Tables

List of abbreviations and acronyms

30S IC	30S initiation complex
70S IC	70S initiation complex
a/eIF	archaeal/eukaryotic initiation factor
A site	aminoacyl-tRNA binding site
aa-tRNA	aminoacyl-tRNA
BME	β-mercaptoethanol
BSA	bovine serum albumin
cDNA	complementary DNA
CAT	catalase
COT	1,3,5,7-cyclooctatetraene
cryo-EM	cryogenic electron microscopy
CTD	carboxy-terminal domain
Da	Dalton
DMSO	dimethyl sulfoxide
DNA	deoxyribonucleic acid
dNTPs	deoxynucleoside triphosphates
EDTA	ethylenediaminetetraacetic acid
EF	elongation factor
E_{FRET}	FRET efficiency
EMCCD	electron multiplying charge coupled device
E site	tRNA exit site
eTLC	electrophoretic thin layer chromatography
fMet-tRNA^{fMet}	initiator tRNA: *N*-formylmethionyl tRNA
FOV	field-of-view
FPLC	fast protein liquid chromatography
FRET	fluorescence resonance energy transfer
GAC	GTPase-associated center

GDP	guanosine diphosphate
GDPNP	guanosine-5’-[(β,γ)-imido]triphosphate
GMPPCP	guanosine-5’-[(β,γ)-methyleno]triphosphate
GOD	glucose oxidase
GTP	guanosine triphosphate
GTPase	guanosine triphosphatase
h	16S rRNA helix
H	23S rRNA helix
HIC	hydrophobic interaction chromatography
HMM	hidden Markov model
IPTG	β-D-1-thiogalactopyranoside
k_a	association rate constant
k_d	dissociation rate constant
mRNA	messenger RNA
NBA	3-nitrobenzyl alcohol
NMR	nuclear magnetic resonance
NTA	nitrilotriacetic acid
NTD	amino-terminal domain
nts	nucleotides
P site	peptidyl-tRNA binding site
PAGE	polyacrylamide gel electrophoresis
PCR	polymerase chain reaction
PDB	Protein Data Bank
PEG	polyethylene glycol
P_i	inorganic phosphate
PMSF	phenylmethyl sulfonyl fluoride
RF	release factor
RNA	ribonucleic acid
ROI	region of interest

r-protein	ribosomal protein
RRF	ribosome recycling factor
rRNA	ribosomal RNA
S (30S, 50S, 70S)	Svedberg unit
SD	Shine-Dalgarno
smFRET	single-molecule fluorescence resonance energy transfer
SRL	sarcin-ricin stemloop
T3	ternary complex
T4gp32	gene product 32 from T4 bacteriophage
TCEP	*tris*(2-carboxyethyl)phosphine
TDP	transition density plot
TEV	tobacco etch virus
TLC	thin layer chromatography
TIR	translation initiation region or total internal reflection
TIRFM	total internal reflection fluorescence microscopy
tRNA	transfer RNA
Tris	*tris*(hydroxymethyl)aminomethane

Acknowledgments

This thesis would not have been possible without the support and encouragement of my family, friends, and coworkers. I am grateful to my advisor Professor Ruben L. Gonzalez, Jr. for giving me the opportunity to work in his lab and learn first-hand about the exciting field of single-molecule biophysics. His mentorship has been extremely valuable for my growth as a scientist, and I am thankful for the unwavering support he has shown toward me throughout my grad school career.

In addition, I am privileged to have been able to work alongside so many wonderful people in the Gonzalez lab. My coworkers have been a constant source of inspiration for me; their intelligence and drive has motivated me and their input and advice have been crucial for the advancement of my research project as well as my overall scientific development. They made the lows of grad school bearable and were a large part of the reason why working in the lab could, not infrequently, be both rewarding and fun. These dear colleagues include Margaret Elvekrog, Jingyi Fei, Jiangning Wang, Mike Englander, Dileep Pulukkunat, Wei Ning, Sam Sternberg, Noam Prywes, Corey Perez, Colin Kinz-Thompson, Bridget Huang, Kelvin Caban, Bo Chen, Jan-Willem van de Meent, Somdeb Mitra, Rathi Srinivas, Peggy Bermel, Madeleine Jensen, and Victor Naumov, and I would like to thank them all for their scientific and personal support over the years. In particular, the training that I received early on from Jingyi Fei and Jianging Wang was invaluable for getting my research project off the ground and moving in the right direction.

Finally, I would like to thank my mom and dad, my brother Peter, and my sister Audrey for their constant love and support, which has made all the difference for me, and for which I am truly grateful.

Chapter 1
Introduction

1.1 Protein synthesis and the translational machinery

The flow of genetic information in the cell from DNA to protein comprises two major steps. In the first, gene sequences are transcribed from DNA into temporary messenger RNA (mRNA) copies by RNA polymerase, and in the second, the nucleotide sequence of the mRNA is translated into a sequence of amino acids by the ribosome. The ribosome is a highly conserved macromolecular machine composed of numerous ribosomal RNA (rRNA) and protein molecules, which is responsible for catalyzing the process of protein synthesis, or translation, in all organisms. During translation, the ribosome incorporates amino acids into a growing polypeptide chain as specified by the sequence of triplet-nucleotide codons on the mRNA. Individual amino acids are delivered to the ribosome by transfer RNA (tRNA) molecules, which are selected based on Watson-Crick base pairing between the tRNA anticodon and the mRNA codon. Thus, during each round of protein synthesis, the ribosome translocates in the 5’ to 3’ direction along the mRNA template and repetitively catalyzes selection of aminoacyl-tRNA (aa-tRNA) substrates and peptide bond formation in order to faithfully synthesize the mRNA-encoded gene product [1].

The translation cycle can be divided into four major stages: initiation, elongation, termination, and ribosome recycling (Figure 1.1). During each stage, essential protein translation factors interact with the ribosome and its aa-tRNA substrates in order to

catalyze different steps of the translation process. Translation factors thus constitute an important part of the translational machinery essential for achieving rapid and tightly controlled protein synthesis. In the initiation stage of protein synthesis, initiation factors IF1, IF2, and IF3 promote assembly of a functional ribosomal complex at the correct start site on the mRNA. During the elongation stage, elongation factor Tu (EF-Tu) catalyzes delivery of aa-tRNA substrates to the ribosome, while elongation factor G (EF-G) catalyzes translocation of the mRNA-tRNA complex through the ribosome in steps of precisely one codon. When synthesis of the mRNA-encoded polypeptide has been completed, the termination stage ensues, which involves release of the nascent polypeptide from the ribosome catalyzed by a class I release factor (RF1 or RF2) followed by recycling of RF1/2 by the class II release factor RF3. Finally, during the ribosome recycling stage of protein synthesis, the joint action of ribosome recycling factor (RRF) and EF-G promotes splitting of the ribosome into subunits; the subunits can then be reassembled at a new mRNA start site in order to begin the process anew [2].

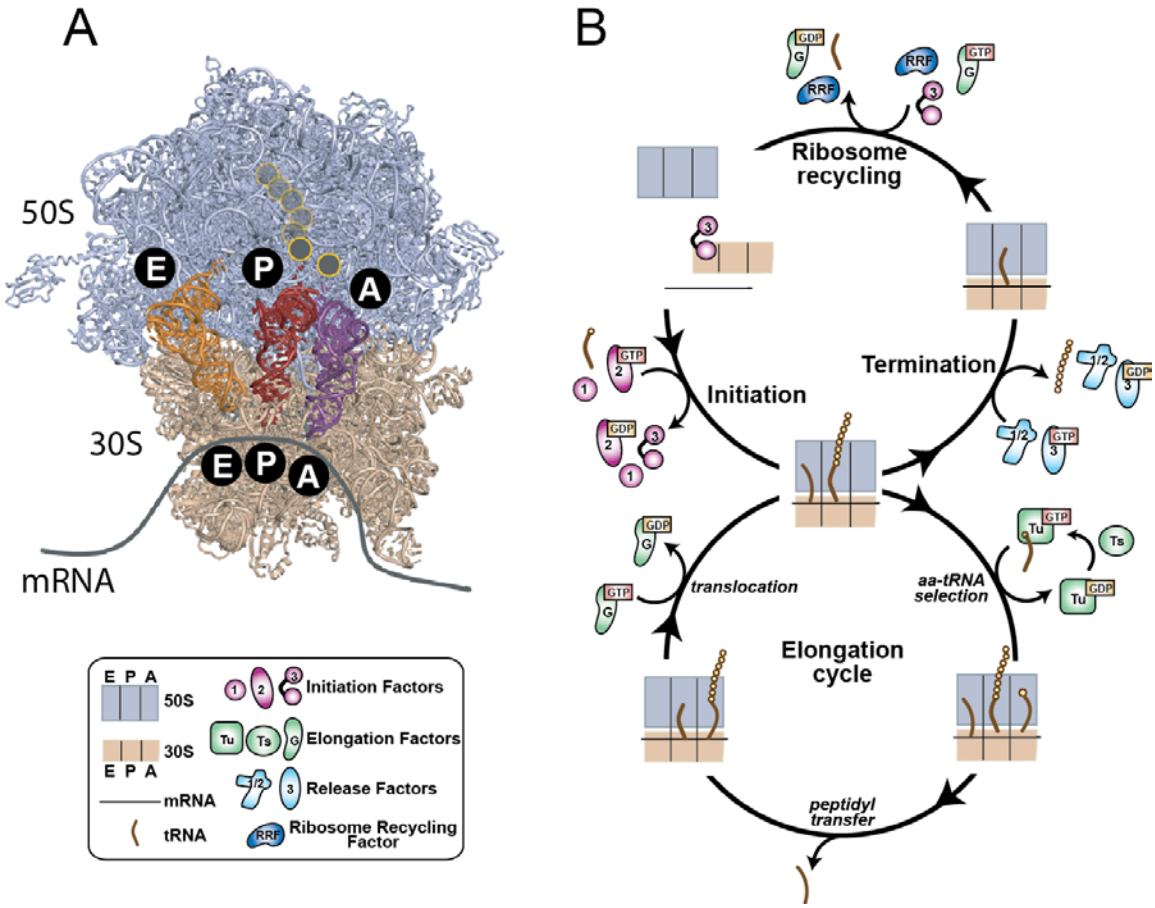

Figure 1.1: The ribosome and the protein synthesis cycle.
(A) X-ray crystallographic structure of the 70S ribosome from *Thermus thermophilus* (PDB ID: 2J00 and 2J01) depicting the three tRNA binding sites. (B) Cartoon schematic of the protein synthesis cycle. The stages of initiation, protein chain elongation, termination, and ribosome recycling are depicted, as well as the translation factors involved at each step. Components of the translational machinery are identified in the legend to the lower left. Figure reproduced from [3].

The multitude of biochemical and mechanical tasks that the ribosome must accomplish during the protein synthesis cycle is reflected in the complexity of its molecular architecture. The ribosome is composed of two subunits, which, in prokaryotes, are termed the large, 50S and small, 30S subunits, according to their respective sedimentation coefficients. The 30S subunit comprises one rRNA molecule

(16S, ~1500 nts) and about 20 ribosomal proteins (r-proteins), while the 50S subunit comprises two rRNA molecules (23S, ~2900 nts and 5S, ~120 nts) and over 30 r-proteins [4]. The two subunits associate to form the intact 70S ribosome, which is a roughly spherical, 2.5 MDa ribonucleoprotein complex. Subunit association is mediated through the formation of twelve RNA-RNA, RNA-protein, and protein-protein contacts at the subunit interface called intersubunit bridges (see Figure 1.6 and Table 1.1) [5-7]. The two-subunit architecture of the ribosome, which is conserved in all organisms, is intimately tied to ribosome function during all stages of protein synthesis: a relative rotation of the subunits plays a critical role in the ribosome's ability to translocate along the mRNA template during peptide chain elongation [8, 9], recycling of ribosomes following completion of protein synthesis entails a splitting of the subunits [10], and the start of a new round of protein synthesis involves reassembly of the intact ribosome from its subunits onto a new mRNA start site [11].

X-ray crystal structures of ribosomal subunits and the intact ribosome have revealed atomic-level details of its functional centers, as well as its interactions with mRNA, tRNA, and translation factor ligands [12, 13]. In addition, cryogenic electron microscopy (cryo-EM) reconstructions of ribosomal complexes have brought our structural understanding of translation into clearer focus [14]. Within the intact ribosome, the mRNA threads through a channel located between the head and body domains of the 30S subunit [15]. Three tRNA binding sites span the intersubunit space, termed the A (aminoacyl-tRNA binding), P (peptidyl-tRNA binding), and E (tRNA exit) sites, which are traversed sequentially as tRNAs make their way through the ribosome during the

elongation cycle (Figure 1.1). Two important functional centers of the ribosome—the decoding center within the small ribosomal subunit and the peptidyl transferase center of the large ribosomal subunit—were found to be composed largely of rRNA, indicating that the ribosome acts as a ribozyme in catalyzing the essential reactions of protein synthesis. In the decoding center, the universally conserved 16S rRNA nucleotides G530, A1492, and A1493 inspect the geometry of the codon-anticodon minihelix formed during aa-tRNA delivery as part of the mechanism for achieving high accuracy aa-tRNA selection [16]. In the peptidyl transferase center, interactions between the so-called A-loop of 23S rRNA (residues U2552-C2556) with the 3’ CCA end of aa-tRNA at the A site, and between the P-loop of 23S rRNA (G2250-C2254) with the 3’ CCA end of peptidyl tRNA at the P site, orient the two tRNAs to promote the chemistry of peptide bond formation, which involves nucleophilic attack of the peptidyl-tRNA ester linkage by the α-amine of the amino acid on the A-site tRNA [17].

The functional core of the ribosome is flanked by two highly mobile regions known as the L1 stalk and the GTPase-associated center (GAC). The L1 stalk, which forms part of the ribosomal E site, is composed of 23S rRNA helices 76 to 78 (H76-78, where capital “H” refers to 23S rRNA helices within the 50S subunit) and r-protein L1. Movements of the L1 stalk are thought to play a role in promoting translocation of deacylated tRNA from the P to the E site following peptide bond formation as well as subsequent release of the E-site tRNA from the ribosome [18, 19]. The GAC is located on the other side of the 50S subunit, near the A site. It serves as the ribosomal binding site for translation factors, many of which, such as IF2, EF-Tu, EF-G, and RF3,

hydrolyze GTP as part of their catalytic cycle. Binding to the ribosome is required to stimulate the GTPase activity of the translation factors, which exhibit very low levels of GTP hydrolysis on their own [20]. The most prominent feature of the GAC is the L7/L12 protein stalk, which consists of, depending on the species, between four and six copies of r-protein L12 (L7 is the same as L12, except with an acetylated N-terminus). The individual copies of L12 form dimers that associate with r-protein L10, which, in turn, binds to the surface of the ribosome via contacts with 23S rRNA [21]. The L7/L12 stalk promotes translation factor binding to the ribosome and has been demonstrated to play a role in the stimulation of GTP hydrolysis by EF-G and EF-Tu [22, 23]. The region at the base of the L7/L12 stalk comprising helices H42-44 of 23S rRNA (nucleotides 1030-1124, *E. coli* numbering) serves as the binding site for the L10-(L7/L12) complex as well as r-protein L11 (Figure 1.2) [21]. L11 binds specifically to an rRNA platform composed of H43/44 at the tip of the stalk base through interactions with its C-terminal domain (CTD). Another prominent structural component of the GAC is the sarcin-ricin stemloop of 23S rRNA (SRL, nucleotides 2654-2665) located within helix H95, so-called because it is the target for cleavage or chemical modification by the cytotoxins α-sarcin and ricin, respectively [24, 25]. The role of the GAC in factor recruitment and GTPase activation is well-established, though the precise mechanistic contribution of the individual GAC components to these processes is still being delineated.

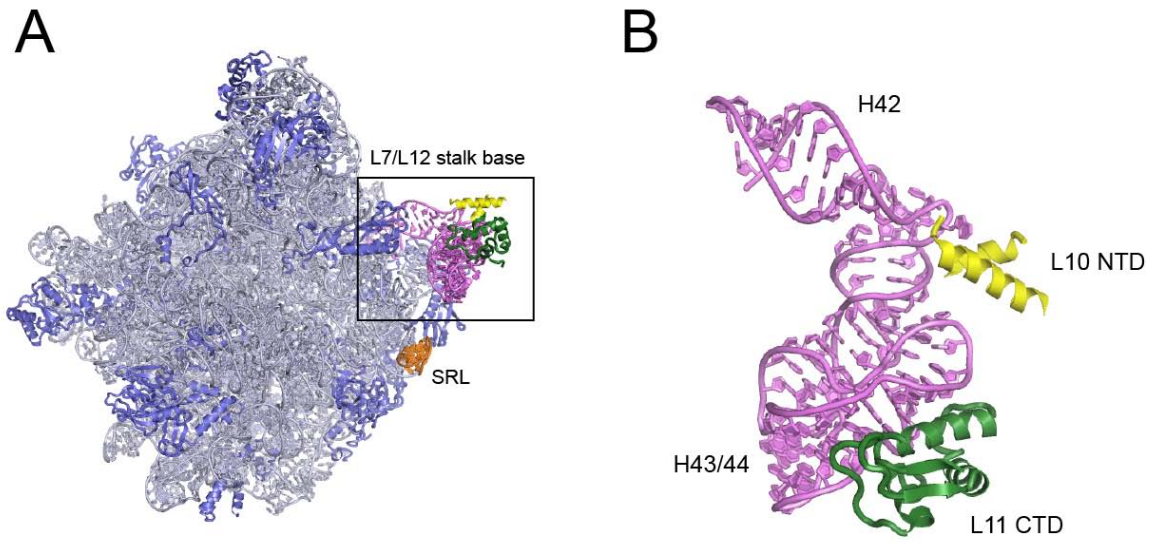

Figure 1.2: Structural depiction of the ribosomal GTPase-associated center.
(A) X-ray crystallographic structure of the *Haloarcula marismortui* 50S ribosomal subunit (PDB ID: 1S72), which contains one of the more complete high-resolution depictions of ribosomal GAC components in the context of the intact 50S subunit. The sarcin-ricin stemloop (SRL: H95, nucleotides A2654 to A2665 of 23S rRNA) is colored orange; the stalk base comprising H42, 43, and 44 of 23S rRNA (nucleotides G1034 to C1121) is colored light purple. N- and C-terminal fragments, respectively, of r-proteins L10 (yellow) and L11 (green) bind to the stalk base rRNA, thereby forming a bifurcated stalk that projects away from the ribosome's core into solution. The L11 NTD, as well as the L10 CTD and associated L7/L12 stalk, were not able to be modeled due to missing electron density. (B) The boxed area corresponding to the stalk base is enlarged and rotated to better visualize the L10 and L11 binding sites.

The structural description of the ribosome and its functional centers provided by X-ray crystallography and cryo-EM sets the stage for detailed mechanistic studies of ribosome dynamics. Protein synthesis is an inherently dynamic process in which movements of ribosomal domains regulate mechanical events such as mRNA-tRNA translocation and factor binding/release, and conformational changes of the translational machinery limit the rates of biochemical steps such as GTP hydrolysis and peptide bond formation [26]. A thorough characterization of the ribosomes's structural dynamics, and its interactions with translation factor and tRNA ligands, will thus be critical for gaining a

better understanding of the molecular mechanisms underlying rapid and accurate protein synthesis. This thesis presents efforts aimed at characterizing ribosome dynamics and ribosome-factor interactions during the late stages of translation initiation.

1.2 The prokaryotic translation initiation pathway

Translation initiation is a multi-step process in which 30S and 50S ribosomal subunits, mRNA, and the initiator *N*-formylmethionyl-tRNA (fMet-tRNA^{fMet}) are brought together to form a functional 70S initiation complex (70S IC). Initiation is considered to be the rate-limiting step of translation, taking several seconds *in vivo* [27]. It represents an important hub for post-transcriptional regulation of gene expression. The efficiency of initiation on a particular mRNA transcript can be modulated by both *cis*- and *trans*-acting elements to control whether the encoded protein is expressed, and if so, at what levels [28]. Furthermore, initiation establishes the reading frame on the mRNA and is thus of critical importance in ensuring the fidelity of protein synthesis. In bacteria, three essential initiation factors, IF1, IF2, and IF3 guide the translation initiation process, helping to achieve the requisite levels of speed and accuracy.

The prokaryotic translation initiation pathway is schematized in Figure 1.3 and reviewed in [11]. In the first major step of the pathway, IF1, IF2, IF3, mRNA, and fMet-tRNA^{fMet} associate with the 30S ribosomal subunit to form a 30S initiation complex (30S IC). The precise order of ligand binding events leading to 30S IC formation, however, remains unclear. The mRNA is initially bound in a standby site, and the initiator tRNA in a codon-independent manner. A rate-limiting conformational change takes place that

promotes formation of the codon-anticodon interaction in the 30S P site and a concomitant stabilization of the 30S IC. Selection of the correct initiator tRNA and mRNA start codon (typically AUG) is promoted by the synergistic action of IF1, IF2, and IF3 at the level of 30S IC assembly.

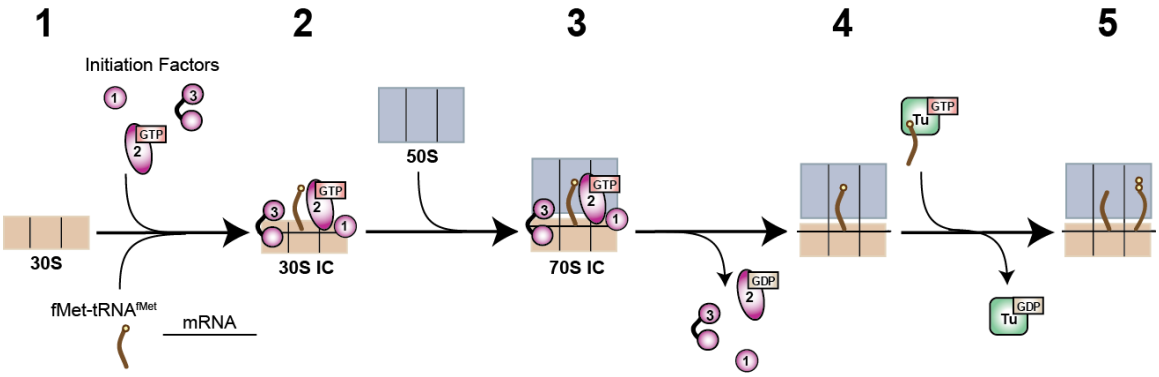

Figure 1.3: Overview of the prokaryotic translation initiation pathway.
Cartoon schematic of major molecular events during translation initiation. As noted in the text, the precise order and timing of ligand binding and dissociation events is still being delineated. The extent of reversibility for each step is also unclear. Step 1→2: Assembly of the 30S initiation complex (30S IC) containing mRNA, initiator fMet-tRNA^{fMet}, IF1, IF2-GTP, and IF3. Step 2→3: Joining of the 50S subunit to the 30SIC to form the 70S initiation complex (70S IC). 50S subunit joining stimulates IF2-dependent hydrolysis of GTP to GDP and P_i. Step 3→4: Dissociation of initiation factors from the ribosome. Step 4 → 5: Delivery of aa-tRNA to the ribosome and formation of the initiation dipeptide. The first elongator aa-tRNA is brought to the ribosome in ternary complex with EF-Tu and GTP. Following GTP hydrolysis and dissociation of EF-Tu, aa-tRNA is accommodated into the A site and peptide bond formation is catalyzed at the peptidyl transferase center.

The next major event in the initiation pathway is joining of the 50S subunit to the 30S IC, which is catalyzed by IF2 in its GTP-bound form. The subunit docking event results in formation of a 70S IC in which interactions between IF2 and the 50S GAC stimulate rapid GTP hydrolysis by IF2. Inorganic phosphate (P_i) is released, and IF2 undergoes a conformational change from its GTP- to its GDP-bound form, followed by

its dissociation from the ribosome. The order and timing of IF1, IF2, and IF3 release from the ribosome has not been well-characterized, and in some cases, represents a point of dispute in the field. For example, different research groups are at odds regarding whether IF3 dissociation precedes or occurs subsequent to 50S subunit joining [29, 30]. Another critical event during 70S IC formation is adjustment of fMet-tRNA^{fMet} within the ribosomal P site, which places it in a configuration conducive to formation of the first peptide bond. The 70S IC can then bind and incorporate the first elongator aa-tRNA, which is delivered to the ribosome in complex with GTP-bound EF-Tu (the so-called EF-Tu:GTP:aa-tRNA ternary complex). Accommodation of aa-tRNA into the ribosomal A site proceeds through a process involving ribosome-dependent GTP hydrolysis by EF-Tu [31], and results in formation of the initiation dipeptide catalyzed by the ribosomal peptidyl transferase center. This event signals completion of initiation and entry into the elongation phase of translation.

1.3 Structure and function of the initiation factors

IF1, IF2, and IF3 are encoded by the *infA*, *infB*, and *infC* genes, respectively, all of which are essential in *E. coli* [11]. The three bacterial initiation factors have structural or functional counterparts in archaea and eukaryotes: a/eIF1A and a/eIF5B are homologs of IF1 and IF2, respectively [32-34], and a/eIF1 is a functional analog of IF3 that binds to the same site on the 30S subunit platform and has similar activities in proofreading of the codon-anticodon interaction at the P site [35]. Although the eukaryotic translation initiation pathway exhibits marked differences from the bacterial pathway (e.g. the

mechanisms of 5' cap loading and 5' to 3' mRNA scanning, and the larger collection of initiation factors involved), the presence of eukaryotic IFs with sequence, structure, and function reminiscent of IF1, IF2, and IF3 suggests that investigations of the bacterial system could shed light on highly conserved features of initiation applicable to higher organisms. In the following sections, I discuss structural and functional characteristics of the bacterial initiation factors. The localization of the initiation factors in the context of the 30S IC is depicted in Figure 1.4.

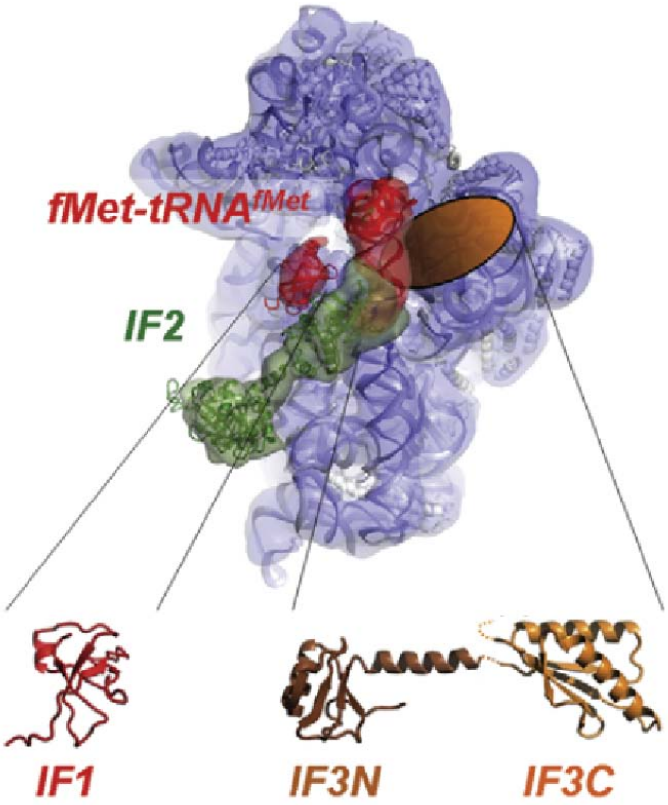

Figure 1.4: Localization of initiation factors and fMet-tRNA^{fMet} on the 30S IC.
Top: Cryo-EM reconstruction of the 30S IC from *Thermus thermophilus* containing IF1, IF2, mRNA and fMet-tRNA^{fMet}. Atomic models of the individual components have been fitted into the electron density. The 30S subunit is colored blue; IF1 and fMet-tRNA^{fMet} are red; IF2 is green. The putative binding site for IF3 is indicated by a brown oval at the platform region of the 30S interface. Bottom: Magnified view of the atomic structures of IF1 (red), IF3N (brown), and IF3C (orange). Figure reproduced from [36].

1.3.1 Initiation Factor 1 (IF1)

IF1, the smallest of the initiation factors, is an 8.2 kDa protein in *E. coli*. Its solution structure has been determined by nuclear magnetic resonance (NMR) spectroscopy, showing a compact five-stranded beta barrel in which beta strands β3 and β4 are connected by a loop that contains a short 3_{10} helix [37]. A crystal structure of the *T. thermophilus* 30S subunit in complex with IF1 details its binding site in a cleft between r-protein S12, the 530 loop of 16S rRNA, and helix h44 (where lower-case "h" is used to denote 16S rRNA helices within the 30S subunit), overlapping with the A site [38]. The major role attributed to IF1 during translation initiation is the stimulation of IF2 and IF3 functions. IF1 binding to the 30S subunit increases the binding affinity of IF2, and it has been shown to enhance IF3's anti-subunit association activity (see section 1.3.3) [39, 40]. Interestingly, mammalian mitochondria do not possess a stand-alone IF1 homolog, but instead, genetic and biochemical data suggest that a conserved 37 amino acid insertion in mitochondrial IF2 substitutes for the IF1 function [41]. These studies lend support to the notion that the primary role of IF1 is the enhancement of IF2 binding and activity.

1.3.2 Initiation Factor 2 (IF2)

IF2 is the largest of the initiation factors, and it plays a central role during formation of both the 30S IC and 70S IC [11, 42]. Three isoforms of IF2 have been identified in *E. coli* (IF2-α, IF2-β, and IF2-γ), which are translated from three different but in-frame start sites on the same *infB* mRNA transcript. The isoforms thus contain the

same C-terminal amino acid sequence but differ in the length of their N-termini: IF2-β and IF2-γ lack the first 157 and 164 amino acids encoded by *infB*, respectively (Figure 1.5B) [43]. The presence of all three isoforms is required for optimal growth in *E. coli* [44], though the functional difference between them has yet to be established, and each isoform is independently capable of promoting *in vitro* translation initiation [45].

The domain organization of IF2 is depicted in Figure 1.5. The N-terminal region of the protein (domains I-III, following the numbering scheme of Mortensen, et al. [46]) is highly variable across species both in terms of length and amino acid composition. The C-terminal region, on the other hand (domains IV, V, VI-1, and VI-2), is more conserved. IF2 homologs from bacteria, archaea, and eukaryotes possess high levels of amino acid sequence identity in the C-terminal region, suggesting that they share a common tertiary structure [47].

There is currently no high-resolution structure available of full-length prokaryotic IF2, though the three-dimensional solution structures of isolated domains VI-1 and VI-2 from *Bacillus stearothermophilus* have been solved independently by NMR spectroscopy [48, 49]. Additionally, three X-ray structures of *Methanobacterium thermoautotrophicum* aIF5B (the archaeal homolog of IF2) have been solved, corresponding to the GTP-bound, GDP-bound, and nucleotide-free forms of the factor [47]. aIF5B lacks the extended N-terminus found in prokaryotic and eukaryotic IF2, and the X-ray structures thus correspond to the conserved C-terminal domains IV, V, VI-1, and VI-2 of *E. coli* IF2.

The crystal structures reveal a “chalice-shaped” enzyme in which domains IV, V, and VI-1 cluster together, and domain VI-1 is connected to domain VI-2 via an extended

~100 amino acid long α-helix (Figure 1.5A) [47]. Domain IV is the GTP binding domain (G domain), which contains the highest level of interspecies sequence conservation among the IF2 homologs. Additionally, it exhibits structural similarity to the G domains of $p21^{Ras}$ as well as EF-Tu and EF-G. It is an eight-stranded beta-sheet flanked by six α-helices and a 3_{10} helix, and it contains the four conserved sequence elements characteristic of the guanine-nucleotide binding site of GTPases (G1/P loop, G2, G3, and G4). The G1/P loop motif participates in phosphate binding, making contacts with the α- and β-phosphates of GTP, while the G3 and G4 loops form the walls of a hydrophobic pocket in which the guanine nucleotide is bound. The G domain contains two segments known as Switch 1 and Switch 2 which change conformation depending on whether GTP or GDP is bound at the active site. The Switch 2 region contains the G2 motif, which coordinates an essential Mg^{2+} ion required for catalysis of GTP hydrolysis.

Domain V of aIF5B is a β-barrel similar to the analogous domain II of EF-G and EF-Tu. It was proposed, based on the similarity in structure and organization of the G-domain/domain II module of EF-G and EF-Tu with the corresponding domains of IF2, that all three translation factors participate in a similar set of interactions with the ribosome [47]. Domain VI-1 has an α/β/α-sandwich fold, while the C-terminal domain VI-2 is a closed beta barrel, which, in the case of the bacterium *B. stearothermophilus*, consists of six anti-parallel β-strands [48].

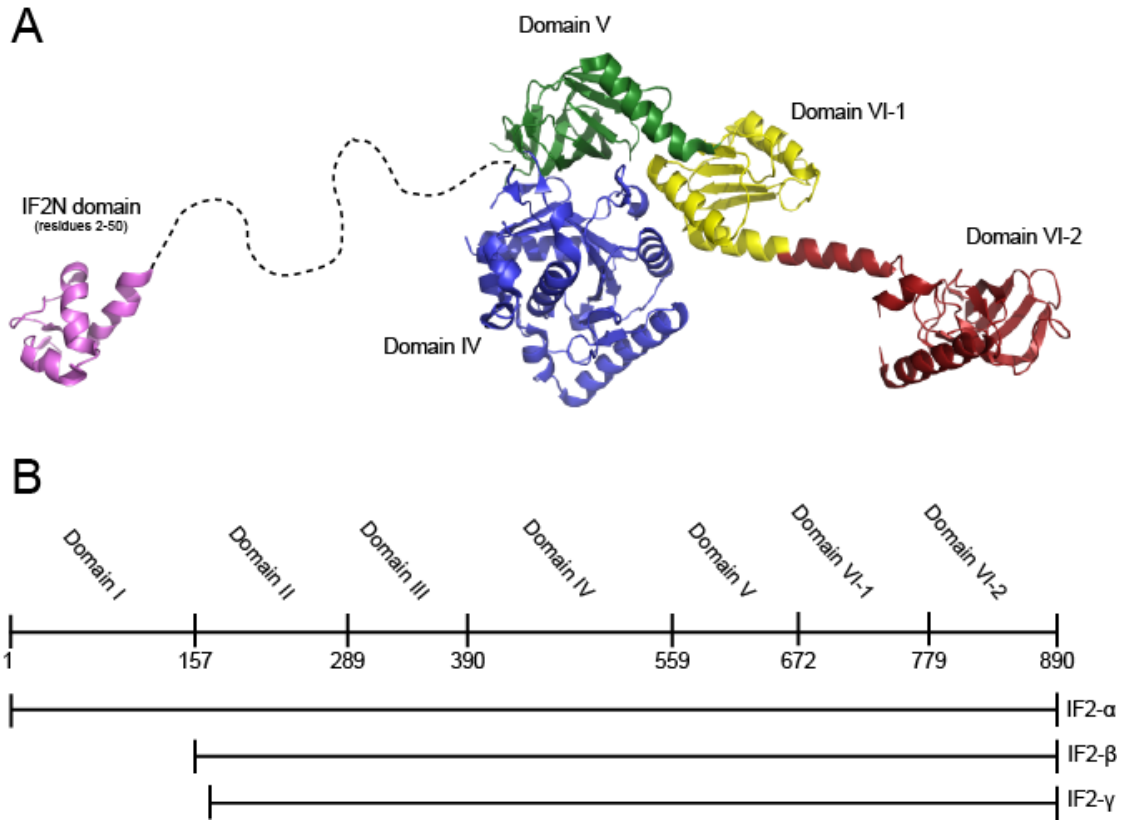

Figure 1.5: IF2 domain architecture.
(A) IF2 structural model. IF2 is divided into domains based on sequence and biochemical data. It consists of a highly conserved C-terminal portion (domains IV, V, VI-1, and VI-2) and an N-terminal portion (domains I, II, and III) that is highly variable in both length and sequence. Right: The X-ray crystal structure of aIF5B from *Methanobacterium thermoautotrophicum* (PDB ID: 1G7T). Based on sequence homology, the C-terminal portion of bacterial IF2 is expected to have a similar structure. Left: A small ~50 amino acid N-terminal subdomain, found in all bacterial IF2s, was solved by NMR spectroscopy (PDB ID: 1ND9). (B) IF2 domain numbering and comparison of the three IF2 isoforms found in *E. coli*. Compared to full-length IF2-α, IF2-β and IF2-γ lack 157 and 164 amino acids from their N-termini, respectively. All three isoforms can perform all biochemical functions attributed to IF2 *in vitro*. Figure adapted from [11].

At the level of 30S IC assembly, one of the main functions of IF2 is the recruitment and stabilization of fMet-tRNA^{fMet} on the 30S ribosomal subunit. IF2 has a weak, micromolar affinity for fMet-tRNA^{fMet} in solution, so it is unlikely that IF2 acts as a tRNA carrier that actively delivers fMet-tRNA^{fMet} to the ribosome. Instead, rapid

kinetics measurements suggest that binding of IF2-GTP to the 30S subunit precedes and subsequently promotes fMet-tRNA^{fMet} binding [50]. IF2 is, by itself, capable of discriminating against elongator tRNAs and preferentially selecting the correct initiator tRNA for incorporation into the 30S IC [51]. Specific recognition of the initiator tRNA by IF2 is based on interactions between domain VI-2 of IF2 and the 3' CCA acceptor stem and amino acid of fMet-tRNA^{fMet} [52]. The presence of the formyl group blocking the α-NH_2 group of the amino acid plays a key role in this recognition [53].

The cryo-EM reconstruction of the 30S IC from *T. thermophilus* adds a level of structural detail that helps to explain the mutual stabilization of IF2 and fMet-tRNA^{fMet} on the small subunit (Figure 1.4). The intermolecular contacts that bridge domain VI-2 of IF2 with the aminoacyl-acceptor stem region of fMet-tRNA^{fMet} generate an IF2–fMet-tRNA^{fMet} sub-complex that is bound to the 30S subunit through two anchor points [54, 55]. The first anchor point consists of domains IV and V of IF2 bound to the 30S subunit in the vicinity of helices h5 and h14 of the 16S rRNA. The second anchor point is established through binding of the tRNA's decoding stem at the P site near the neck region of the 30S subunit, within a pocket formed by 16S rRNA helices h24, h29, h30, h31, h34, and h44, and r-proteins S9 and S13.

Following 30S IC formation, IF2 is responsible for catalyzing 50S subunit joining. Rapid docking of the 50S subunit to the 30S IC has been shown by light scattering measurements to be completely dependent on the presence of both IF2 and fMet-tRNA^{fMet} (Figure 1.7) [56]. Additionally, subunit association has been shown to occur ~20-fold faster in the presence of GTP versus GDP [57]. In the context of a

completely assembled 30S IC, IF2 is likely present in the GTP-bound form, since its affinity for GTP, but not GDP, is enhanced in the presence of 30S subunits and fMet-tRNA^{fMet} [57]. Thus, correct assembly of a 30S IC containing fMet-tRNA^{fMet} and IF2-GTP will be quickly followed by joining of the 50S subunit. During the subunit docking event, formation of interactions between IF2's G-domain and the 50S subunit's GAC lead to rapid GTPase activation and GTP hydrolysis by IF2 [58]. The functional consequences of IF2-dependent GTP hydrolysis during initiation will be discussed below in Section 1.5.2. In addition to stimulating fMet-tRNA^{fMet} binding to the 30SIC and rapid 50S subunit joining, IF2 has been proposed to play a role in positioning of fMet-tRNA^{fMet} within the P site during the late stages of 70S IC formation, such that it can act as an efficient donor in the first peptidyl transfer reaction [59].

1.3.3 Initiation Factor 3 (IF3)

IF3 is a 20.4 kDa protein of 180 amino acids that consists of two globular domains of roughly equal size (IF3C and IF3N) connected by a long, flexible, lysine-rich linker [60]. The structures of isolated IF3C and IF3N have been solved by X-ray crystallography and NMR [61-63]. The IF3C domain folds into a structure in which two parallel α-helices pack against a mixed four-stranded β-sheet. The IF3N domain exhibits a similar globular α/β fold in which a single α-helix packs against a five-stranded antiparallel β-sheet. The flexible linker connecting IF3C and IF3N is highly conserved in both length and hydrophilic character, and is essential for the function of IF3 *in vivo* [64].

Data from chemical probing and hydroxyl radical footprinting of the 16S rRNA [65, 66], as well as a low-resolution cryo-EM reconstruction of the 30S-IF3 complex [67], place the binding site for IF3C at the platform on the interface side of the 30S ribosomal subunit. The binding site for IF3N, on the other hand, is less certain. Some of the data suggest that it is located at the neck region of the small subunit near the P site [66], while other data would place it near the E site, neighboring r-proteins S7 and S11 [65]. In theory, discrepancies regarding the specifics of IF3 localization on the 30S IC could be resolved if IF3 or its isolated domains could populate different binding sites on the 30S subunit. In support of this possibility, single-molecule fluorescence resonance energy transfer (smFRET) data from our laboratory suggests that the IF3-bound 30S subunit can adopt at least three distinct conformations, corresponding to different interdomain distances between IF3C and IF3N [68]. These data are indicative of conformational rearrangements of the IF3-30S complex, which could correspond to conformational changes of the 30S subunit itself and/or transitions of one or both IF3 domains between different 30S binding sites.

Multiple functions have been ascribed to IF3 at different stages of the translation initiation pathway. Binding of IF3 to the 30S subunit inhibits association of the 50S subunit, and this anti-association function may play a role in splitting of the 70S ribosome into subunits during ribosome recycling [40, 69]. By preventing premature binding of the 50S subunit and thus shifting the 30S + 50S $\rightleftarrows$ 70S equilibrium toward free subunits, IF3 supplies the cell with a pool of free 30S subunits on which the other initiation factors, mRNA, and fMet-tRNA^{fMet} can assemble to form the 30S IC. Light

scattering studies have shown that the presence of IF3 prevents the formation of aberrant 70S ICs lacking initiator tRNA: an IF3-dependent blockage of 50S subunit joining to 30S ICs containing only mRNA, IF1, IF2, and IF3 is alleviated through the additional inclusion of fMet-tRNA^{fMet} [70]. The location of IF3's binding site at the interface side of the 30S platform suggested that IF3 may sterically occlude the formation of several intersubunit bridges, in particular bridge B2b comprising RNA-RNA contacts between H69 of the 23S rRNA with h23, h24, and h25 of the 16S rRNA [65]. Given the critical role of intersubunit bridges in subunit association (see [71-73], Figure 1.6, and Table 1.1), this was proposed as a mechanism to explain IF3's anti-association properties.

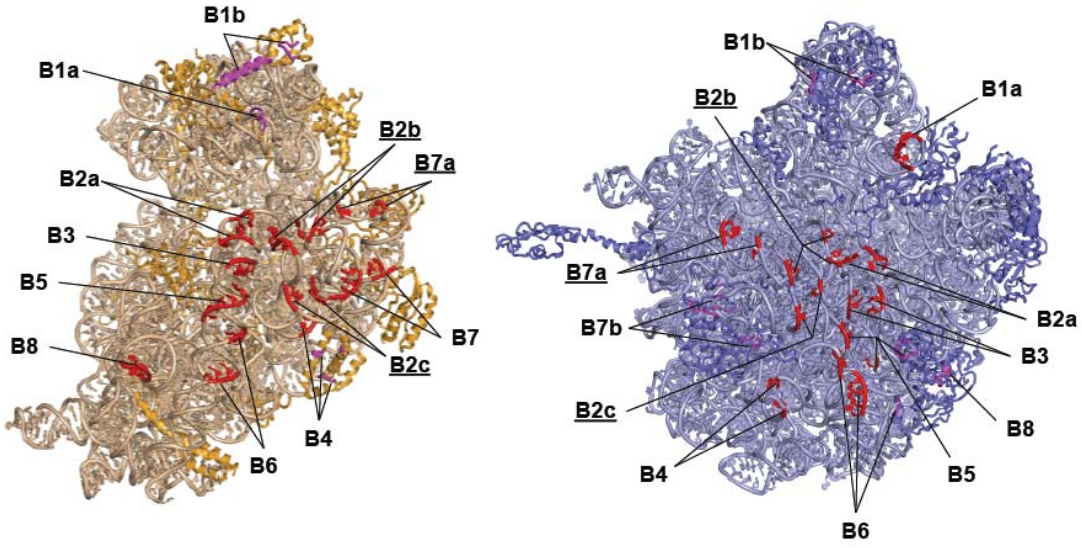

Figure 1.6: Intersubunit bridges play an important role in subunit association.
Interface view of the 30S (left) and 50S (right) subunits highlighting the twelve intersubunit bridges in the *E. coli* ribosome (PDB ID: 2AVY and 2AW4). rRNA components of the bridges are colored red and protein components are colored pink. IF3 has been proposed to exert its anti-association function by sterically blocking formation of B2b, B2c, and/or B7a (underlined). Figure provided by Mr. Wei Ning.

Table 1.1: Catalog of intersubunit bridges in the *E. coli* 70S ribosome.
List of the 30S and 50S components comprising the twelve intersubunit bridges of the *E. coli* ribosome. Interaction types: P-P, protein-protein; P-R, protein-RNA; R-R, RNA-RNA. Table reproduced from reference [71].

Interaction type	30 S component (16 S rRNA helix/nt or S protein)	50 S component (23 S 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

During assembly of the 30S IC, IF3 additionally plays a prominent role in maintaining the fidelity of initiator tRNA and start codon selection. The presence of IF3 causes preferential destabilization of initiation complexes containing non-canonical codon-anticodon interactions at the P site, on account of incorrectly bound elongator tRNAs or start codons other than AUG, GUG, and UUG [74, 75]. The *infC* gene encoding IF3 possesses a non-canonical AUU start codon, which allows for autoregulation of IF3 expression at the level of translation initiation *in vivo* [76, 77]. IF3 has also been shown to dissociate 30S ICs formed at the 5’ AUG start codon of so-called leaderless mRNAs [78].

1.4 IF2-catalyzed 50S subunit joining

Docking of the 50S ribosomal subunit to the 30S IC represents a critical regulatory checkpoint in the translation initiation pathway. The 50S docking event leads to formation of a 70S IC that can enter into the elongation phase of protein synthesis. It is therefore essential that the formation of aberrant 70S complexes—either missing tRNA or containing an incorrect tRNA and/or start codon at the P site—be prevented. Regulation of 50S subunit joining involves the interplay of IF2 and IF3 activities, which oppose each other in promoting and inhibiting this process, respectively. As described above, IF3 prevents the docking of 50S subunits to 30S ICs lacking initiator tRNA, while IF2-GTP is required for acceleration of this process following correct assembly of a complete 30S IC [70].

Translation initiation efficiency can be affected by differences in components within the mRNA's translation initiation region (TIR), such as variations in the Shine-Dalgarno sequence and initiation codon. These features of the TIR have been shown to modulate the efficiency of 70S IC formation by altering the kinetics of 50S subunit joining [29]. Thus, mRNAs with favorable TIRs undergo subunit association more rapidly than mRNAs with unfavorable TIRs. These differing kinetics were attributed to IF1- and IF3-induced conformational rearrangements of the 30S ribosomal subunit that regulate 50S docking as well as IF3 dissociation. Similarly, another study concluded that IF3-mediated inhibition of translation initiation at the non-canonical AUU start codon occurs not at the level of 30S IC assembly, but rather due to a reduction in the rate of 70S IC formation [79].

IF2 counterbalances the inhibitory effect of IF3 by serving to accelerate subunit joining to the 30S IC. Subunit joining is most rapid in the presence of both IF2 and fMet-tRNA^{fMet} [30, 56, 70]. Cryo-EM reconstructions of IF2 bound to the 30S IC and 70S IC offer insight into the mechanism by which this occurs (Figure 1.7). The 30S IC reconstruction contains IF2 in its GTP-bound form [54], while the 70S IC reconstruction contains IF2 bound to the non-hydrolyzable GTP analog GDPNP, thus trapping it in the pre-GTP hydrolysis state [80]. In the context of the 30S IC, an intermolecular interaction between IF2's C-terminal domain VI-2 and the aminoacyl-acceptor stem of fMet-tRNA^{fMet} positions the tRNA so that it adopts a binding configuration that the authors refer to as the 30S P/I state. In this configuration, the tRNA's anticodon stem-loop is bound to the 30S P site, but the tRNA body is tilted such that its elbow region and acceptor stem lie between the P and E sites. This orientation of the IF2–fMet-tRNA^{fMet} sub-complex was proposed to optimally position both the tRNA body and IF2 surface for interfacing with the 50S subunit to guide rapid subunit association. The presence of IF2 on the 30S IC increases the surface area available for interactions with the 50S subunit by approximately 25%, and a high degree of shape complementarity between IF2 and the 50S subunit interface was observed, helping to explain how IF2 accelerates this process. The cryo-EM reconstruction of the 70S IC reveals that, following subunit association, IF2 buries ~2600 Å^2 of the 50S subunit surface area that was formerly solvent accessible. The authors thus liken IF2-catalyzed subunit joining to a high-affinity dimerization process in which the burial of large surfaces at the dimerization interface is driven by a large, favorable change in Gibb's free energy (ΔG < -10 kcal/mole).

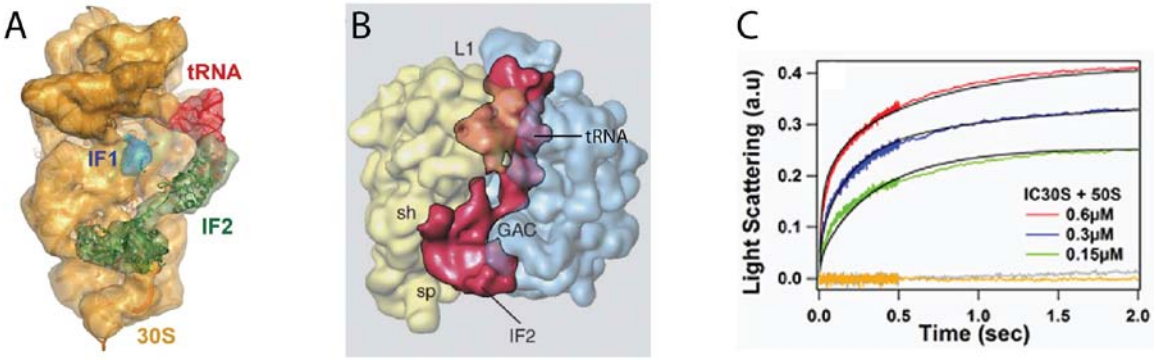

Figure 1.7: The IF2–fMet-tRNA^{fMet} subcomplex promotes 50S subunit joining to the 30S IC. (A) Cryo-EM reconstruction of the 30S IC from *T. thermophilus*. IF2 (green) and fMet-tRNA^{fMet} (red) bind to the interface side of the 30S subunit where they are oriented so as to promote rapid docking of the 50S subunit. Figure reproduced from [54]. (B) Cryo-EM reconstruction of the 70S IC from *E. coli*, representing the state immediately following 50S subunit joining. Electron density corresponding to IF2 and fMet-tRNA^{fMet} (red) is observed spanning the interface between the 30S (yellow) and 50S (blue) subunits, suggesting that the IF2–fMet-tRNA^{fMet} sub-complex promotes subunit docking in a process akin to protein-protein dimerization. Figure adapted from [80]. (C) Light scattering experiments demonstrate the dependence of 70S IC formation on IF2 and fMet-tRNA^{fMet}. An increase in light scattering upon rapid mixing of 30S ICs with 50S subunits indicates the formation of 70S particles. For the red, blue, and green curves, varying concentrations of 50S subunit (0.6 μM, 0.3 μM, and 0.15 μM, respectively) were mixed with 30S ICs containing IF1, IF2, IF3, mRNA, fMet-tRNA^{fMet}, and GTP. The grey and orange curves arise from experiments conducted at 3.0 μM 50S subunits and in the absence of either fMet-tRNA^{fMet} or IF2, respectively. Figure reproduced from [56].

The L7/L12 protein stalk of the 50S subunit's GAC was recently shown to play a key role in IF2-catalyzed subunit joining [81]. Association of 50S subunits with 30S ICs containing IF2-GTP was found to be ~40-fold slower in the presence of 50S subunit cores depleted of L7/L12, and this effect could be reversed by reconstitution of the 50S cores with L7/L12. No effect of L7/L12 depletion was seen in the rate of 50S subunit association with 30S ICs in the absence of IF2 or in the presence of IF2 in its GDP-bound or nucleotide-free forms. Therefore, it was concluded that rapid subunit association depends on the formation of specific interactions between IF2-GTP and the L7/L12 stalk [81]. These interactions are likely mediated by a conserved region of the L7/L12 C-

terminal domain shown previously by NMR to bind to IF2, EF-Tu, EF-G, and RF3 [82]. Analogous to the proposal that the L7/L12 stalk serves as an initial ribosomal binding site for EF-G and EF-Tu:GTP:aa-tRNA ternary complex preceding their accommodation into the ribosome's A site [21], interactions between L7/L12 and IF2 may correspond to a first step in the subunit association process that would rapidly bring the 30S IC and 50S subunit into close proximity in roughly the correct relative orientation. Formation of specific interactions between IF2 and the 50S subunit core could then more readily occur to lock into place the precise inter-subunit orientation required to form an elongation-competent 70S IC.

The notion that the presence of IF2-GTP, as opposed to IF2-GDP or the nucleotide-free form of the factor, is required for rapid 50S subunit docking represents the consensus view in the literature [57, 81]. (It should be noted, however, that there are conflicting reports which state that the guanine-nucleotide state of IF2 has no effect on 70S IC formation [58].) This suggests that binding of GTP to IF2, which is favored in the presence of 30S subunits and fMet-tRNA^{fMet} [57], results in a conformational change of the factor that is required to accelerate the subunit joining reaction. Recently, IF2 mutations were identified outside of its C-terminal tRNA-binding domain that allow it to bypass the dual requirement of formylated initiator tRNA and GTP for the switch from its inactive to its active form [83, 84]. Several of these mutants exhibited high levels of subunit joining activity even in the absence of initiator tRNA. These findings suggested that under normal conditions, fMet-tRNA^{fMet} and GTP indirectly affect subunit joining by promoting "activation" of IF2 on the 30S IC.

1.5 Ribosome-dependent GTP hydrolysis by IF2

IF2 has no intrinsic GTP hydrolysis activity, and stimulation of this activity requires the presence of 70S ribosomes. Thus, during initiation, the GTP molecule is hydrolyzed only upon association of the 30S–IF2-GTP complex with the 50S subunit during 70S IC formation. Interactions between IF2-GTP and the 50S subunit's GAC lead to rapid hydrolysis of GTP to GDP and P_i, with a rate of 30 ± 5 sec^{-1} at room temperature. Subsequent release of P_i occurs more slowly at a rate of 1.5 ± 0.5 sec^{-1} [58]. Components of the GAC such as L7/L12, L11 and its associated rRNA, and/or the SRL probably interact with the guanine-nucleotide binding pocket of IF2 in order to induce structural rearrangements required to activate hydrolysis. Based on the position and orientation of IF2-GTP seen in the cryo-EM reconstruction of the 30S IC, it was proposed that domains IV and V of IF2 interact with the GAC immediately upon 50S subunit docking, which would rationalize the rapid rate of GTP hydrolysis observed biochemically (Figure 1.7A) [54].

1.5.1 Mechanism of GTP hydrolysis

The hydrolysis of GTP to GDP and P_i take place through in-line attack on the γ-phosphate by an activated nucleophilic water molecule. It proceeds through a pentacoordinate transition state in which the γ-phosphate is surrounded by a trigonal bipyramid of oxygen atoms. Important elements of the catalytic mechanism include activation of the water molecule and its correct positioning for direct, in-line attack, and stabilization of the transition state through neutralization of negative charge build-up on

the β- and γ-phosphates [85, 86]. Magnesium is an essential cofactor for catalysis of GTP hydrolysis, and in the X-ray structure of aIF5B/IF2-GDPNP, the catalytic Mg^{2+} ion is located in a cleft between the GTP binding site and the switch 2 loop. It is stabilized by contacts with the β- and γ-phosphates of GTP, a hydroxyl group of Thr19 within the P loop, and a water molecule coordinated by Asp76 of switch 2 [47].

The GTPase activity of translation factors is dependent on their binding to the ribosome, and the role of ribosomal components and ribosome-translation factor interactions in the catalytic mechanism has been investigated in some detail for the case of EF-Tu. The results from these studies may apply to IF2 as well, as it has been suggested that ribosome-catalyzed GTP hydrolysis may occur via a similar mechanism for all of the translation factor GTPases, which have overlapping binding sites at the 50S GAC. The structural similarity between the G domain and domain II of EF-Tu and EF-G with the corresponding domains of IF2 points towards a shared mechanism of GTPase activation via a common set of interactions with the ribosome [11].

An invariant histidine residue has been identified within the G-domain of the translation factors (His84 in EF-Tu and His448 in IF2) which may act as a general base during catalysis that activates the water molecule to OH^- through abstraction of a proton [87-89], though it should be noted that this proposal has recently been challenged [90]. The crystal structure of Trp-$tRNA^{Trp}$, EF-Tu, and the antibiotic paromomycin bound to the 70S ribosome, stalled by the non-hydrolyzable GTP analog GDPCP, depicts the state just before GTP hydrolysis by EF-Tu [89] (Figure 1.8).

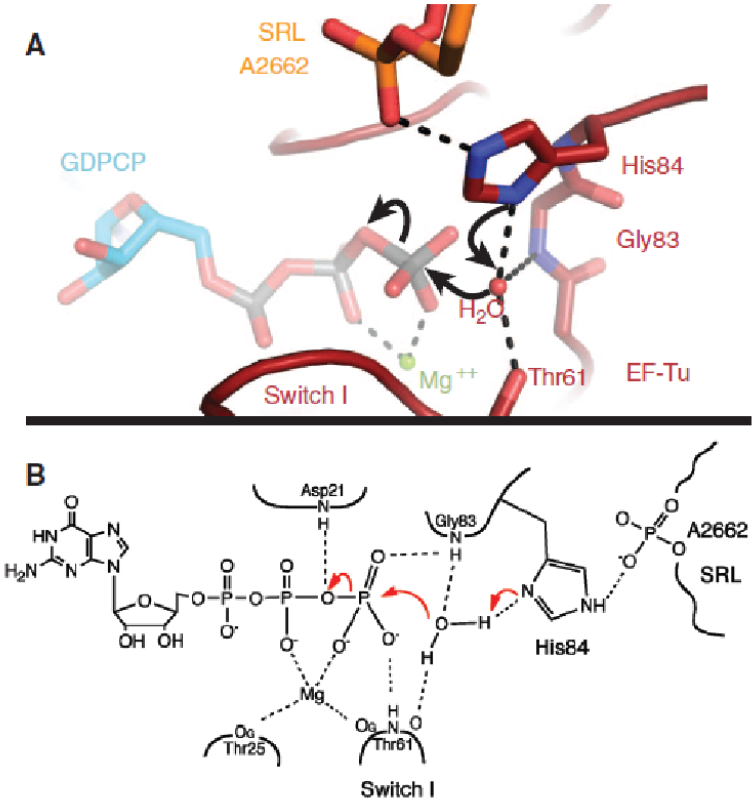

Figure 1.8: Proposed mechanism for ribosome- and translation-factor mediated GTP hydrolysis.
(A) View of EF-Tu's GTP-binding pocket in the state just prior to GTP hydrolysis, from the X-ray crystal structure of Trp-tRNA^{Trp}, EF-Tu-GDPCP, and paromomycin bound to the 70S ribosome. The catalytic water molecule is positioned by interactions with Thr61, Gly83, and His84 for inline attack on the γ-phosphate. His84 (His448 in IF2) was proposed to act as a general base that deprotonates and thus activates the water molecule for nucleophilic attack. His84 is stabilized in its active conformation through an interaction with A2662 of the SRL. (B) Chemical structure diagram of the interactions between EF-Tu and GTP which help position and activate the catalytic water molecule and stabilize the transition state for GTP hydrolysis through neutralization of negative charge build-up at the β- and γ-phosphates. Figure reproduced from [89].

In this structure, EF-Tu's catalytic His84 is stabilized in its active conformation via an interaction with A2662 of the SRL. Prior to ribosome binding, His84 is rotated away from the GTP binding site, where its access to GTP is likely blocked by a "hydrophobic gate" formed by Val20 of the P loop and Ile60 within switch 1 [91].

Therefore, when EF-Tu binds to the ribosome, interactions between His84 and A2662 of the SRL promote the reorientation of His84 into the active site where it can interact with the catalytic water molecule [89]. Interactions between A2662 of the SRL and His448 of IF2 likely play an analogous role in inducing the catalytic conformation of His448 during GTPase activation of IF2. Mutation of His448 to Glu in IF2 results in severely impaired ribosome-dependent GTP hydrolysis without affecting IF2's affinity for GTP, which underscores the importance of this residue for GTPase activation [88]. Interestingly, a mutation of the corresponding residue in human eIF5B (His706 to Gln mutation) also causes a substantial reduction in GTP hydrolysis activity, suggesting a highly conserved mechanism [92].

1.5.2 Role of GTP hydrolysis during initiation

There is strong evidence that GTP hydrolysis is necessary for IF2 dissociation from the 70S IC. This was suggested by early experiments which showed that, in the presence of GDPNP, formation of 70S ICs containing fMet-tRNA^{fMet} is stoichiometric with IF2 concentration, but that in the presence of GTP, IF2 can function catalytically [93]. Dissociation of IF2 from the ribosome is usually considered to be a prerequisite for accommodation of the first elongator aa-tRNA into the ribosomal A site and formation of the first peptide bond. Luchin, et al. showed that peptide bond formation was blocked completely in the presence of the GTPase-deficient His448Glu IF2 mutant, despite the fact that this mutant promoted fMet-tRNA^{fMet} binding to 70S ICs at near-wild type levels [88]. Combined with the observation that His448Glu IF2, but not wild-type IF2, stably

associates with the ribosome following 70S IC formation, this suggested that the dominant negative phenotype of His488Glu IF2 observed *in vivo* [94] is a consequence of its inability to recycle off of the ribosome. The same conclusion was reached by Antoun, et al., whose light-scattering data indicated that 50S subunit joining is fast in the presence of both IF2-GDPNP and IF2-GTP, but that subsequent initiation dipeptide formation is inhibited with IF2-GDPNP [57].

The explanation for why GTP hydrolysis is required for dissociation of IF2 is likely tied to the conformational changes expected to occur following P_i release as IF2 transitions from its GTP- to its GDP-bound form [47]. Comparison of two cryo-EM reconstructions of the *T. thermophilus* 70S IC—one trapped in the pre-hydrolysis state with IF2-GMPPCP and the other in the post-hydrolysis state with IF2-GDP—suggested that the transition of IF2 to the GDP-bound form involves conformational changes of both IF2 and the ribosome which alter their intermolecular contacts and cause IF2 to undergo a shift of ~10 Å outwards from the intersubunit space, adopting a "ready-to-leave" conformation [95].

GTP hydrolysis, transition to the GDP-bound state, and release of IF2 from the ribosome may have additional mechanistic consequences. For example, the cryo-EM reconstruction of the 70S IC from *E. coli* containing IF2-GDPNP shows that the position of IF2's C-terminal domain sterically occludes fMet-tRNA^{fMet} binding to the 50S P site, such that the tRNA is bound in a P/I hybrid configuration in which its anticodon sits at the 30S P site and its 3' aminoacyl-acceptor stem resides between the 50S subunit P and E sites [80]. Accommodation of fMet-tRNA^{fMet} into the classical P/P configuration,

which would require a movement of its acceptor stem by ~28 Å, is thus prevented prior to GTP hydrolysis by IF2. Adjustment of tRNA into the P/P configuration may be coupled to the conformational changes of IF2 following GTP hydrolysis and/or its dissociation from the ribosome. In a separate study, single-molecule data suggested that GTP hydrolysis by IF2 is additionally required to promote an intersubunit rearrangement of the ribosome necessary for forming a 70S IC that can efficiently bind the first EF-Tu:GTP:aa-tRNA ternary complex and enter into the elongation cycle [96].

1.6 Interactions between IF2 and the GTPase-associated center

IF2's binding site on the 70S ribosome overlaps with that of numerous other translation factors (e.g. EF-Tu, EF-G, RF1, RF2, and RF3). Cryo-EM density maps have provided a low-resolution view of IF2 at the factor-binding site following 50S subunit joining to the 30S IC [80, 95]. The relative positioning of IF2's domains with respect to structural features of the 30S and 50S ribosomal subunits can be gleaned, including a close approach between IF2's G domain and the 50S subunit's GAC. Chemical probing and crosslinking data have proven useful in identifying specific interactions between IF2 and the individual structural components of the GAC. IF2 was crosslinked to L7/L12 [97], and binding of IF2 protects residues in the SRL (G2655, A2665, and G2661) from chemical modification [98]. In another study, chemical nucleases tethered to a cysteine residue introduced into domain VI-1 of IF2 cleaved positions C1076 and G1068 within the L11 region of 23S rRNA [99].

Despite these studies, a detailed structural understanding of IF2's interactions with the GAC during initiation is lacking. A major reason for this is the inherently dynamic nature of the GAC, in particular the L11 region and the L10-(L7/L12) protein stalk, which are poorly resolved in the available cryo-EM reconstructions. Similarly, clear electron density for the GAC is lacking in the majority of X-ray crystal structures of 50S subunits and the 70S ribosome. This implies a high degree of conformational flexibility of L11 and L10-(L7/L12), and suggests that conformational rearrangements of the GAC may play an important functional role during the interaction of translation factors with the ribosome. For example, conformational changes of the GAC may help promote and/or be coupled to IF2 activities during initiation, such as GTP hydrolysis, P_i release, and factor dissociation.

Efforts to better understand the mechanistic role of GAC components during IF2-catalyzed 50S subunit joining and 70S IC formation will benefit from a characterization of their structural dynamics and the timing of their interactions with IF2. While IF2's GAC interaction partners can be ascertained from the cryo-EM, chemical probing, and crosslinking studies, other methods will be required to gain access to this dynamic information. It is likely that IF2's interaction with the 50S subunit during initiation is far more complex than a simple one-step binding reaction, and that it instead comprises a complex sequence of interactions between IF2 and different GAC components. I envision a scenario in which intermolecular contacts between IF2 and the different GAC components are formed, broken, and rearranged over the course of subunit joining, GTP hydrolysis, P_i release, and factor dissociation. In this thesis, I describe the development of

single-molecule methods to monitor interactions between IF2 and the GAC during real-time initiation reactions, in an attempt to characterize these dynamics and gain insights into the mechanistic role of the GAC during 50S subunit joining and 70S IC formation. Specifically, I have focused on the interactions between IF2 and r-protein L11.

1.7 L11 structure and function

As described above, the L7/L12 stalk has been shown to play a prominent role in IF2-GTP–dependent recruitment of the 50S subunit to the 30S IC, and the SRL is probably involved in the mechanism of GTPase activation. The mechanistic function of L11 and its associated rRNA helices H43 and H44, however, is less clear. L11 binds cooperatively with L10-(L7/L12) to 23S rRNA at the stalk base [100]. It is anchored to the H43/44 platform at the tip of the stalk base through interactions with its C-terminal domain (CTD). L11 is disordered in the majority of X-ray crystal structures of the ribosome, but a crystal structure of the isolated complex between L11 and H43/44 from *Thermotoga maritima* has been solved at 2.6 Å resolution, offering a glimpse of the full-length protein and its interactions with the rRNA [101] (Figure 1.9). The L11 protein is composed of two globular domains connected by a short linker region. The CTD binds to H43/44 through recognition of the A1067 stem-loop's minor groove, and in so doing stabilizes a compact rRNA tertiary fold [102]. The N-terminal domain (NTD), on the other hand, makes very few contacts with H43/44, resulting in a gap between the NTD and the rRNA in the crystal structure of the isolated L11-rRNA complex [101]. This lack of stabilizing interactions allows for conformational flexibility of the NTD, as evidenced

by relatively poor electron density and a higher average B factor for this domain (Figure 1.9A). This conformational flexibility was proposed to play a functional role during the interaction of translation factors with the ribosome [101]. Specifically, it was suggested that reversible association of the NTD with the H43/44 rRNA platform could serve as a molecular switch; conformational switching of the NTD between open and closed states, for example, might serve to alter the accessibility or conformation of the rRNA and thus promote or inhibit binding of the translation factors.

The L11 region of the ribosome constitutes the binding site for the thiazole family of antibiotics, the most well-studied being thiostrepton, initially identified as an inhibitor of mRNA-tRNA translocation [103]. Thiostrepton resistance mutations cluster around a cleft between the A1067/A1095 stem-loops of rRNA and the proline-rich helix 1 of the L11 NTD [101]. Thiostrepton-producing *Streptomyces* strains possess a 2'-O-methylation of A1067 as a natural resistance mechanism, and mutations within the rRNA (transversion mutations of A1067 or A1095) as well as the L11 NTD (Pro22 to Ser/Thr substitutions) have additionally been shown to confer resistance [104-106]. These results suggested that thiostrepton binds within the cleft between L11's NTD and H43/44. This was recently confirmed by X-ray crystal structures of thiostrepton, as well as the related drugs micrococcin and nosiheptide, in complex with the *Deinococcus radiodurans* large ribosomal subunit [107]. Thiostrepton may inhibit protein synthesis by sterically blocking the formation of interactions between translation factors and L11 and/or H43/44; alternatively, the drug could prevent conformational switching of the L11 NTD by stabilizing it in a fixed conformation with respect to the rRNA.

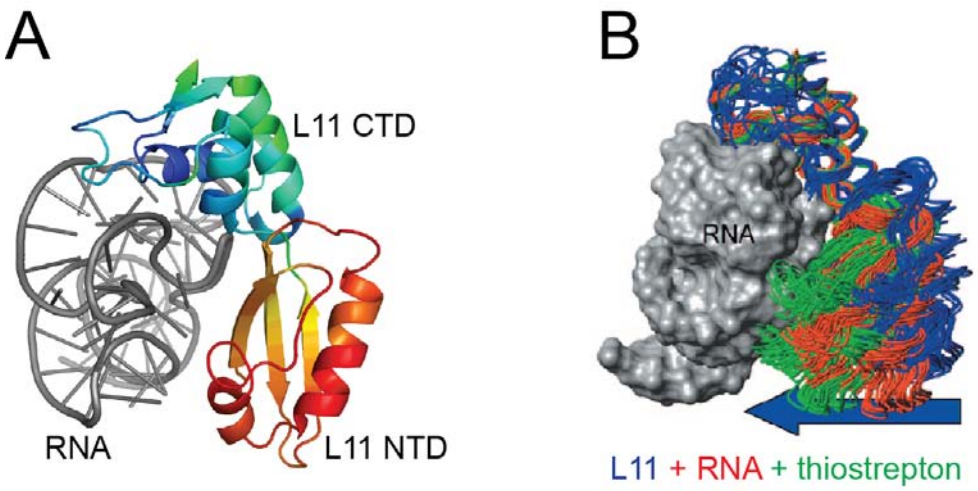

Figure 1.9: Inter-domain flexibility of r-protein L11.
(A) X-ray crystal structure of L11 from *Thermotoga maritima* bound to a 58-nucleotide stretch of 23S rRNA (nucleotides 1051-1108 in *E. coli*) corresponding to helices H43/44 (PDB ID: 1MMS). L11 residues are colored according to B-factor, with a color scale ranging from blue ($B \leq 20$ Å^2) to red ($B \geq 100$ Å^2). The average main chain B-factor is 40 Å^2 for the CTD and 85 Å^2 for the NTD. (B) Different L11 inter-domain configurations determined using NMR spectroscopy. Structures were calculated for free L11 (blue), L11 bound to a 60-nucleotide rRNA fragment (nucleotides 1050-1109, red), and L11 in complex with rRNA and the antibiotic thiostrepton (green). Bundles of the twenty best structures were aligned based on stable secondary structural elements within the CTD and are shown here as backbone traces. Figure reproduced from [108].

Since thiostrepton binds to the L11 region of the ribosome, the effect of this drug on IF2 activities could indirectly shed light on the mechanistic role of L11 during translation initiation. There are multiple reports that thiostrepton does indeed affect IF2 functions on the ribosome, but unfortunately, the data are conflicting: independent studies have concluded that thiostrepton enhances [109, 110], inhibits [111], or has no effect [112] on the multiple-turnover, ribosome-dependent GTP hydrolysis activity of IF2. Similarly, there are studies indicating that the ability of IF2 to function catalytically in 70S IC formation is prevented [112] [113] or enhanced [110] in the presence of

thiostrepton. The reason for the conflicting results, and thus a clear description of the effect of thiostrepton on IF2-ribosome interactions, remains to be resolved.

Several investigations making use of L11-depleted ribosomes have also yielded data attesting to the functional importance of L11-IF2 interactions. Sucrose density gradient analysis showed that efficient association of the 30S IC with the 50S subunit requires the presence of both L11 and IF2, suggesting that L11 represents an important binding partner for IF2-GTP during subunit joining [114]. In another study, the multiple-turnover GTP hydrolysis activity of IF2 elicited by L11(-) ribosomes was found to be four-fold less than with wild-type ribosomes, and near wild-type levels of activity could be restored through reconstitution of the L11(-) ribosomes with purified L11 [110]. The lower GTP hydrolysis activity observed with L11(-) ribosomes was not due to a lower affinity of IF2 for the ribosome, as indistinguishable binding curves were generated for IF2 in the presence of L11(-) and wild-type ribosomes.

The function of the L11 region is expected to be intimately tied to conformational change of its rRNA and/or protein components. Several modes of conformational flexibility have been proposed for the L11 region. As noted above, movements of the flexible L11 NTD could result in different orientations with respect to the CTD and H43/44. Solution-state NMR data provide evidence for a relative reorientation of the CTD and NTD upon binding of the free protein to RNA, and again upon addition of thiostrepton, the latter of which causes the NTD to bend closer to the CTD and RNA (Figure 1.9B) [108]. Similarly, a comparative analysis of L11 NTD conformations among numerous X-ray crystal structures and cryo-EM reconstructions representing different

functional states of the ribosome suggested a wide dynamic range of L11 movements [115]. The specific configuration of the NTD with respect to the CTD and RNA was found to be correlated with the ligand-bound state of the ribosome. In general, translation factor binding was accompanied by movement of the L11 NTD towards the body of the 50S subunit, while in the factor-free "resting state" the L11 NTD moved away from the 50S subunit. These movements are apparently facilitated by a rotation around the protein's hinge region and an accompanying twist of the NTD [115]. The NTD orientation may also change depending on the nucleotide bound state of the translation factor, as was noted by Agrawal, et al. based on a comparison of cryo-EM reconstructions of the 70S ribosome bound to either EF-G–GDPNP or EF-G–GDP. Following GTP hydrolysis by EF-G, a ~5 Å downwards movement of the NTD occurs, resulting in the formation of an intermolecular "arc-like connection" between the tip of the NTD and the G' subdomain of EF-G–GDP [116].

The 23S rRNA components of the L11 region may also undergo functionally important conformational changes. For example, molecular dynamics simulations suggest that nucleotide A1067 within the H43/44 L11-binding platform may reversibly flip out and adopt a solvent-exposed conformation in order to contact the elbow region of the A/T hybrid-state tRNA during aa-tRNA accommodation [117].

Finally, large-scale movements of the entire L11 arm with respect to the body of the 50S subunit may occur during translation. The crystal lattice of the vacant *E. coli* 70S ribosome contained two independent copies of the ribosome per asymmetric unit, which differed from each other by, among other things, a ~15 Å movement of the L11 arm

towards the A site (Figure 1.10). The point of flexibility permitting this structural rearrangement was identified as the region comprising base pairs U1035/G1120 through C1041/G1114 at the base of H42 [5]. This region includes two consecutive G-U wobble base pairs (U1035/G1120 and G1036/U1121), which are known to promote conformational flexibility of RNA [4].

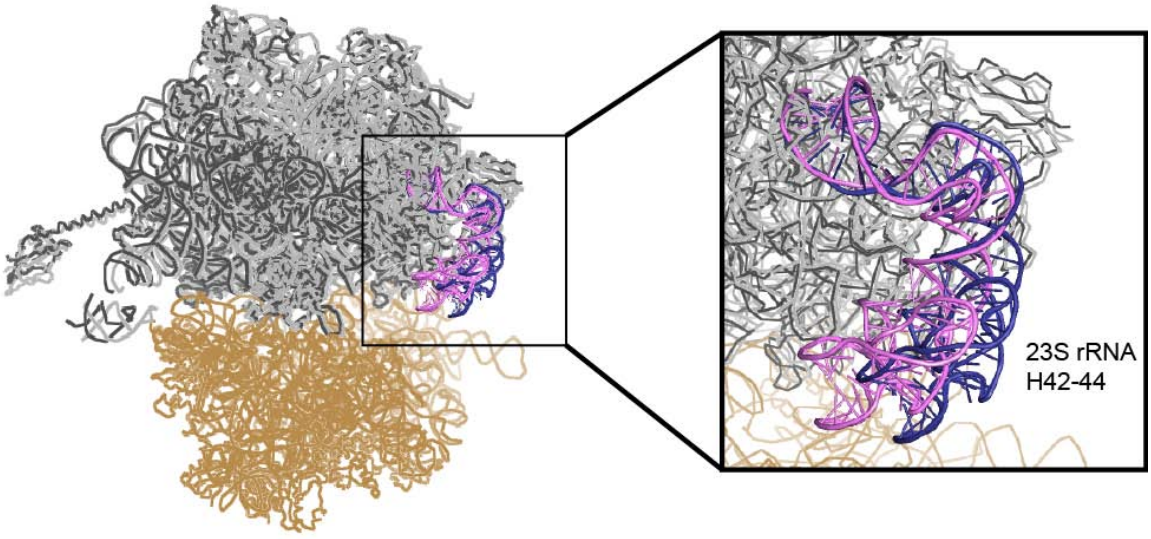

Figure 1.10: Conformational flexibility of the L11 arm.
Different conformations of the L11 arm (23S rRNA helices H42-44) were observed for two unique copies of the ribosome within the asymmetric unit of crystals formed with vacant *E. coli* ribosomes. The structures of ribosome I (PDB ID: 2AVY and 2AW4) and ribosome II (PDB ID: 2AW7 and 2AWB) were aligned based on total 23S rRNA using PyMOL [121]. Compared to the L11 arm in ribosome II (dark blue), the L11 arm in ribosome I (pink) has moved ~15 Å towards the A site. The L11 protein has been removed from the figure for clearer visualization of the rRNA.

Notably, the L11 arm contains additional rRNA structural motifs known to facilitate conformational flexibility, namely a kink-turn motif at the internal loop of H42, and a G-ribo motif at nucleotide 1042 near the junction between H41 and H42 [4, 118]. The large-scale rearrangements of the L11 arm seen in the crystal structures are likely related to the

"inward curling" of the L11 region observed by cryo-EM to occur upon binding of translation factors RF1/2 [119], EF-G [116], and EF-Tu:GTP:aa-tRNA [120].

The conformational changes described above may be required for L11 to interact with each translation factor in a unique way and/or to change its interaction pattern with translation factors over the course of their residency on the ribosome. Conformational changes of the L11 region could be coupled to and/or actively promote IF2 conformational rearrangements associated with factor binding, GTP hydrolysis, P_i release, or dissociation. Similarly, these conformational changes could facilitate the reversibility of L11-IF2 interactions; reversible formation and disruption of IF2-L11 contacts could stabilize or destabilize IF2's binding to the ribosome and thus serve to modulate its association and dissociation kinetics. The observation that IF2 increases the accessibility of H43 and H44 to cleavage by hydroxyl radicals strongly suggests that conformational changes within the L11 region do in fact occur upon IF2 binding to the ribosome [110].

In summary, the L11 region of the 50S subunit's GAC likely plays an important role in regulating IF2-mediated processes during translation initiation, but its precise mechanistic function remains to be clarified. There is evidence for conformational changes of the L11 protein, its associated rRNA binding platform, and the entire L11 arm, which are presumably intimately tied to L11 function. A better understanding of the mechanistic role of L11 during initiation will thus require a detailed characterization of L11 dynamics, as well as the timing of its interactions with IF2. Investigations into the nature and timescale of IF2-L11 interactions, including how these interactions change as

a function of the individual biochemical steps of the initiation pathway, should prove useful in building a detailed mechanistic model for IF2-mediated formation of the 70S IC.

1.8 smFRET and TIRF microscopy: tools for studying dynamics

Single-molecule fluorescence resonance energy transfer (smFRET) is a powerful biophysical technique for studying conformational dynamics [122, 123]. FRET is based upon energy transfer between two fluorophores, a donor and an acceptor [124, 125]. The energy transfer is a non-radiative process that occurs via dipole-dipole interactions between the donor in its excited electronic state and the acceptor in its ground state. It depends upon significant overlap between the emission spectrum of the donor and the absorbance spectrum of the acceptor (Figure 1.11A). The rate of energy transfer from donor to acceptor, $k_T(r)$, depends on the inter-fluorophore distance according to the equation:

$$k_T(r) = \frac{1}{\tau_D}\left(\frac{R_0}{r}\right)^6 \quad (1)$$

where r is the distance between the two fluorophores, τ_D is the lifetime of the donor in the absence of acceptor, and R_0 is the Förster distance, the distance at which the energy transfer efficiency is 50%. The value of R_0 is different for each donor-acceptor pair and is described by the following equation:

$$R_0 = 0.211(\kappa^2 n^{-4} Q_D J(\lambda))^{1/6} \quad \text{(in Å)} \quad (2)$$

Here, κ^2 is a geometric factor that describes the relative orientation of the donor and acceptor transition dipoles; n is the refractive index of the medium; Q_D is the quantum yield of the donor in the absence of acceptor; and $J(\lambda)$ is the "overlap integral", a measure of the amount of spectral overlap between the donor emission and acceptor absorption profiles. Values of R_0 typically range from 20 to 60 Å, with $R_0 \approx 55$ Å for the commonly used Cy3-Cy5 donor-acceptor pair. This means that the efficiency of energy transfer E_{FRET}, given by

$$E_{FRET} = \frac{R_0^6}{R_0^6 + r^6} \tag{3}$$

will be most sensitive to distance changes in the range of ~35 to 75 Å for Cy3-Cy5 (Figure 1.11B). This window of sensitivity is often ideal for probing structure and dynamics of biological molecules. Thus, by attaching donor and acceptor fluorophores to appropriate regions of a biomolecule of interest, measurements of E_{FRET} can be used to monitor conformational changes occurring on the Å length scale. FRET can also be used to study the association and dissociation of two binding partners, with non-zero FRET efficiencies reporting on the bound state.

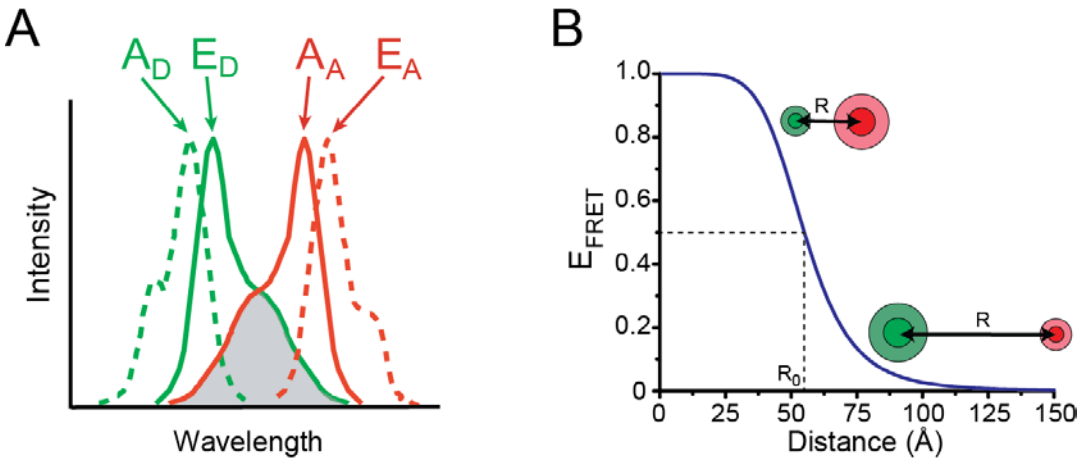

Figure 1.11: Physical principles of fluorescence resonance energy transfer.
(A) FRET is dependent on significant spectral overlap between the donor emission spectrum and the acceptor absorbance spectrum. A_D: donor absorbance, E_D: donor emission, $A_{A:}$ acceptor absorbance, E_A: acceptor emission. (B) Plot of FRET efficiency versus distance for a donor-acceptor pair with R_0 = 55 Å.

Probing conformational dynamics and substrate-ligand binding at the level of single molecules permits access to mechanistic information that is often difficult or impossible to extract from ensemble measurements. For example, the ensemble may be composed of two subsets of molecules, those in conformational state A and those in conformational state B. This so-called static heterogeneity would be clearly identifiable from smFRET measurements, and the relative occupancy of each subpopulation could be readily tabulated. In contrast, the population-weighted average FRET value generated by an ensemble measurement may hide the presence of subpopulations completely. Similarly, dynamic heterogeneity might exist, in which individual molecules within the ensemble fluctuate stochastically between two or more conformational states. Asynchronous fluctuations would be masked by ensemble averaging, but are readily observed in single-molecule time trajectories, allowing for a straightforward analysis of

their underlying thermodynamic and kinetic parameters. Finally, rare and short-lived events, such as infrequent substrate-ligand binding events or transiently sampled conformational intermediates, might only be observable using techniques with single-molecule resolution [126, 127].

Detection of fluorescence from single molecules is often accomplished using total internal reflection fluorescence (TIRF) microscopy [128]. In a commonly implemented TIRF microscope configuration, a focused laser beam is introduced onto a quartz microfluidic flowcell through a prism attached to the flowcell surface. When the laser beam encounters the interface between the quartz microscope slide (index of refraction $n_1 \approx 1.5$) and aqueous buffer ($n_2 \approx 1.3$), it will either be transmitted into the sample or, at incidence angles past a certain critical angle θ_C (Equation 4), totally reflected from the interface back into the quartz slide.

$$\theta_C = \sin^{-1}\left(\frac{n_1}{n_2}\right) \tag{4}$$

Total internal reflection of the incident laser beam generates a weak electromagnetic field within the sample called the evanescent field. The intensity of the evanescent field decays exponentially with distance from the quartz-buffer interface, and can be used to selectively excite fluorescence from molecules within ~100 nm of the surface. TIRF thereby increases the signal-to-noise of the measurements and permits detection of fluorescence from single molecules near the microscope slide surface.

Fluorescently labeled molecules can be immobilized on the surface, thus preventing diffusion and permitting imaging of individual molecules with observation

times (~10s of sec to min) limited by fluorophore photobleaching. Fluorescently labeled ribosomal complexes are often assembled on an mRNA with a 5’ biotin modification, which allows surface tethering via a biotin-streptavidin interaction. Using wide-field optics combined with an electron-multiplying charge coupled device (EMCCD) camera as the detector, fluorescence emission from hundreds of spatially separated ribosomal complexes can be collected simultaneously as a function of time [126, 127].

In a typical smFRET experiment, the Cy3 donor fluorophore is directly excited with green, 532 nm laser illumination, and can either fluoresce itself or transfer energy through FRET to a nearby Cy5 acceptor. Cy3 and Cy5 emission are collected simultaneously and separated using dual-view optics onto two halves of the EMCCD detector. Following data acquisition, the Cy3 and Cy5 fields of view are aligned to identify co-localized Cy3 and Cy5 fluorescence spots originating from single ribosomal complexes. From this data, fluorescence intensity versus time trajectories are produced, as well as the corresponding FRET efficiency versus time trajectories (Figure 1.12).

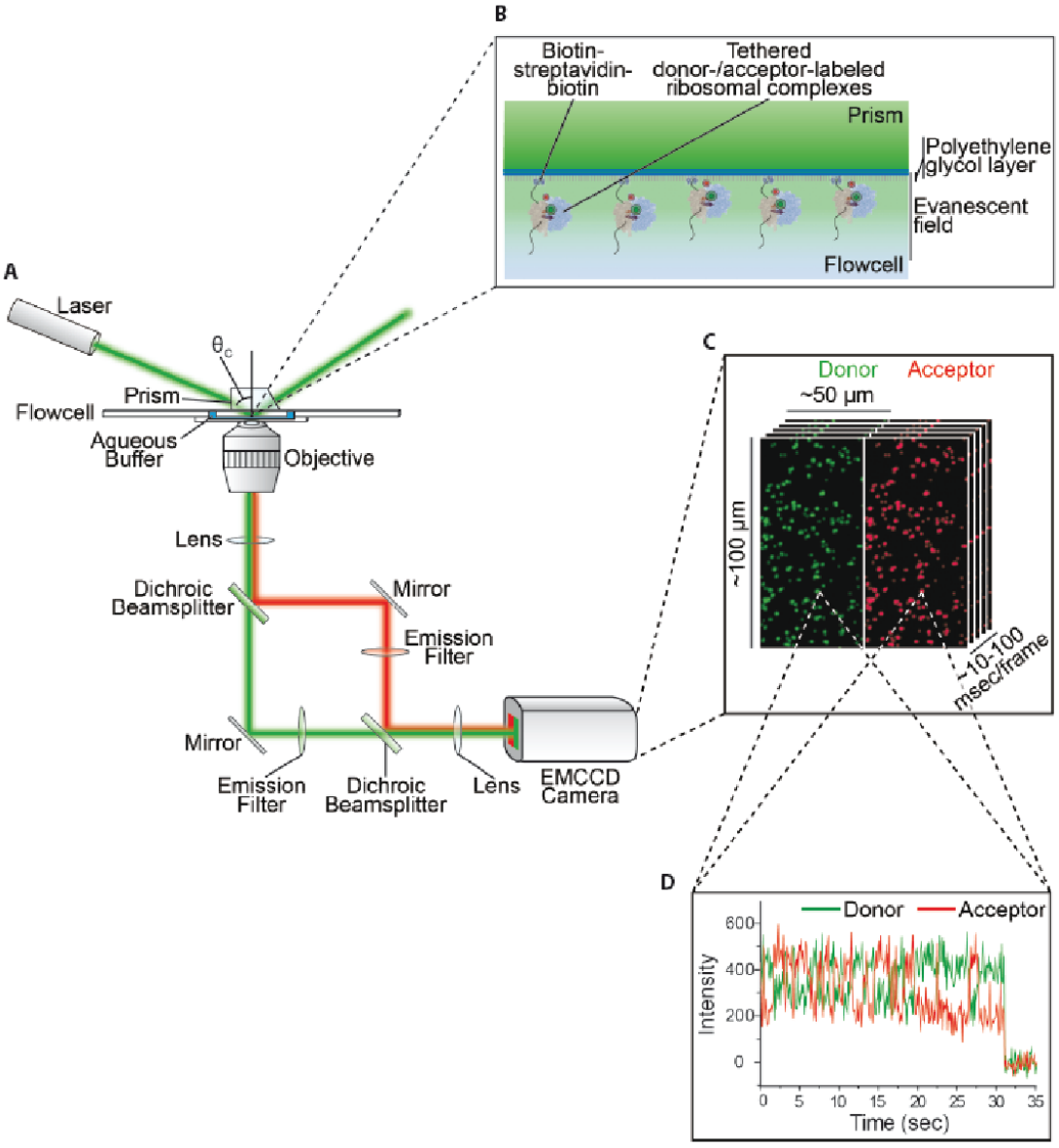

Figure 1.12: TIRF microscopy and smFRET data collection.
(A) Principles of operation and typical optical setup of a prism-based TIRF microscope. See text for detailed description. (B) Inset showing an enlargement of the quartz-buffer interface of the sample flowcell. Fluorescently labeled ribosomal complexes are tethered to the polyethylene glycol (PEG)/biotin PEG-passivated quartz surface through a biotin-streptavidin-biotin bridge and are thereby confined within the effective excitation volume of the evanescent field. (C) Inset showing an enlargement of a typical dual-view fluorescence image recorded by the EMCCD detector. Individual frames contain donor and acceptor signals from hundreds of spatially localized molecules. Images are collected at a frame rate of 10s to 100s of msec to follow the time evolution of the fluorescence signals. (D) Representative donor and acceptor emission intensities versus time trajectory derived from a single fluorescently labeled ribosomal complex within the field-of-view. Figure reproduced from [127].

The same experimental set-up can also be used for single-molecule fluorescence co-localization measurements. In this case, both Cy3- and Cy5-labeled species are directly excited using a combination of 532 nm and 635 nm laser illumination. Detection of co-localized Cy3 and Cy5 fluorescent spots indicates binding of Cy3- and Cy5-labeled species to the same surface-immobilized ribosome. Co-localization experiments lack the structural information inherent to FRET, but can be useful in investigating, for example, the temporal organization of ligand binding events or the assembly/disassembly of multi-component macromolecular complexes. Notably, this technique has recently been applied to dissect the spliceosome assembly pathway by following the ordered association of fluorescently labeled spliceosomal sub-complexes onto a pre-mRNA substrate in real time [129].

1.9 Summary and motivation for my Ph.D. work

Docking of the 50S subunit to the 30S IC is a critical checkpoint along the translation initiation pathway. This event is highly regulated by the initiation factors in order to ensure efficient formation of a 70S IC that is correctly assembled, with fMet-tRNA^{fMet} bound at the mRNA's start codon, and is primed for in-frame synthesis of the encoded protein. IF2 plays a central role in accelerating subunit joining and guiding formation of the elongation-competent 70S IC. During its catalytic cycle, IF2 interacts with multiple components of the GAC; how each of these components collaborates with IF2 in order to direct subunit joining, GTP hydrolysis, and IF2 recycling is a question of key mechanistic interest for understanding ribosome function. Particularly unclear is the

mechanistic role of the L11 region in these processes, due in part to conflicting reports in the literature regarding the effect of L11-targeting antibiotics on IF2 activities. However, based on its structurally dynamic nature, it is likely that conformational rearrangements of the L11 region help to coordinate one or more of IF2's activities. It is unlikely that IF2's interaction with the GAC can be depicted as a one-step binding event; instead, it probably involves a stepwise series of interactions with the different GAC components. In other words, the binding site of IF2 at the 50S GAC is likely dynamically remodeled during subunit joining and 70S IC formation, with the formation, breakage, and rearrangement of intermolecular contacts between IF2 and different GAC components serving to direct and regulate biochemical and mechanical events.

I thus hypothesized that conformational rearrangements between IF2 and L11 help guide IF2 activities during the late stages of translation initiation. This thesis presents the development of an smFRET signal between IF2 and L11 to test this hypothesis. Based on smFRET measurements, it may be possible to characterize the interaction patterns between IF2 and L11, including the timing of the formation and disruption of IF2-L11 intermolecular contacts and how they change along the reaction pathway leading to 70S IC formation. This approach could ultimately help clarify the mechanistic function of the L11 region during initiation. Development and validation of the smFRET signal is presented in Chapter 2 of this thesis, and its sensitivity to IF2-L11 conformational dynamics is demonstrated. An experimental platform is then described which was implemented to follow real-time, IF2-catalyzed subunit joining reactions. The

smFRET signal is additionally shown to be useful in characterizing the lifetime of IF2 on the 70S IC prior to its release.

The ability to observe 50S subunit joining to the 30S IC in real time with my smFRET assay opens the door to investigations of how other components of the translational machinery regulate 70S IC formation. In Chapter 3, the regulatory effect of IF3 on this process is specifically addressed. IF3 is known to have anti-subunit association properties, and observation of this activity at the level of single molecules has yielded insights into IF3's mechanism of action. This approach has allowed direct observation of reversible 50S subunit docking to the 30S IC, and in so doing, provides a unique perspective with which to interpret bulk biochemical data on subunit joining. Additionally, identification of subunit joining intermediates emphasizes the idea that formation of an elongation-competent 70S IC is a multi-step process, with opportunities for positive or negative regulation at each point along the pathway. Finally, the data have implications with regards to the stimulus and timing of IF3 dissociation from the ribosome during initiation.

In general, the order and timing of ligand binding and dissociation during translation initiation has not been well characterized, though this information is crucial for a detailed understanding of the initiation mechanism and its regulation. Of particular interest is the relative timing of IF2 dissociation with respect to the arrival of the first EF-Tu:GTP:aa-tRNA ternary complex. IF2 and ternary complex bind to partially overlapping sites at the 50S GAC, and it is often assumed that IF2 dissociation from the ribosome must precede ternary complex binding. To directly test this idea, I have designed single-

molecule fluorescence co-localization experiments, which are described in Chapter 4. By investigating the temporal relationship between IF2 release and ternary complex arrival, I hope to gain insight into how the ribosome coordinates the transition from initiation into elongation. My single-molecule approach could yield insights into questions such as: Does the presence of ribosome-bound IF2 preclude ternary complex binding completely, or are transient binding events possible? Does ternary complex affect the dissociation rate of IF2 from the 70S IC? More generally, these experiments may provide a clearer picture of how the ribosome efficiently coordinates the sequential binding of translation factors to overlapping sites at the GAC over the course of protein synthesis.

1.10 References

1. Green, R. and H.F. Noller, *Ribosomes and translation.* Annu Rev Biochem, 1997. **66**: p. 679-716.

2. Ramakrishnan, V., *Ribosome structure and the mechanism of translation.* Cell, 2002. **108**(4): p. 557-72.

3. Fei, J., *Coupling of Ribosome and tRNA Dynamics during Protein Synthesis*. 2010, Columbia University: New York.

4. Korostelev, A., D.N. Ermolenko, and H.F. Noller, *Structural dynamics of the ribosome.* Curr Opin Chem Biol, 2008. **12**(6): p. 674-83.

5. Schuwirth, B.S., M.A. Borovinskaya, C.W. Hau, W. Zhang, A. Vila-Sanjurjo, J.M. Holton, and J.H. Cate, *Structures of the bacterial ribosome at 3.5 A resolution.* Science, 2005. **310**(5749): p. 827-34.

6. 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.

7. Gao, H., J. Sengupta, M. Valle, A. Korostelev, N. Eswar, S.M. Stagg, P. Van Roey, R.K. Agrawal, S.C. Harvey, A. Sali, M.S. Chapman, and J. Frank, *Study of*

the structural dynamics of the E coli 70S ribosome using real-space refinement. Cell, 2003. **113**(6): p. 789-801.

8. Frank, J. and R.K. Agrawal, *A ratchet-like inter-subunit reorganization of the ribosome during translocation.* Nature, 2000. **406**(6793): p. 318-22.

9. Horan, L.H. and H.F. Noller, *Intersubunit movement is required for ribosomal translocation.* Proc Natl Acad Sci U S A, 2007. **104**(12): p. 4881-5.

10. Janosi, L., H. Hara, S. Zhang, and A. Kaji, *Ribosome recycling by ribosome recycling factor (RRF)--an important but overlooked step of protein biosynthesis.* Adv Biophys, 1996. **32**: p. 121-201.

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. Steitz, T.A., *A structural understanding of the dynamic ribosome machine.* Nat Rev Mol Cell Biol, 2008. **9**(3): p. 242-53.

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

14. Agirrezabala, X. and M. Valle, *Structure and Dynamics of the Ribosome as Revealed by Cryo-Electron Microscopy*, in *Molecular Machines in Biology*, J. Frank, Editor. 2011, Cambridge University Press. p. 117-41.

15. Yusupova, G.Z., M.M. Yusupov, J.H. Cate, and H.F. Noller, *The path of messenger RNA through the ribosome.* Cell, 2001. **106**(2): p. 233-41.

16. Ogle, J.M., D.E. Brodersen, W.M. Clemons, Jr., M.J. Tarry, A.P. Carter, and V. Ramakrishnan, *Recognition of cognate transfer RNA by the 30S ribosomal subunit.* Science, 2001. **292**(5518): p. 897-902.

17. Voorhees, R.M., A. Weixlbaumer, D. Loakes, A.C. Kelley, and V. Ramakrishnan, *Insights into substrate stabilization from snapshots of the peptidyl transferase center of the intact 70S ribosome.* Nat Struct Mol Biol, 2009. **16**(5): p. 528-33.

18. Valle, M., A. Zavialov, J. Sengupta, U. Rawat, M. Ehrenberg, and J. Frank, *Locking and unlocking of ribosomal motions.* Cell, 2003. **114**(1): p. 123-34.

19. Andersen, C.B., T. Becker, M. Blau, M. Anand, M. Halic, B. Balar, T. Mielke, T. Boesen, J.S. Pedersen, C.M. Spahn, T.G. Kinzy, G.R. Andersen, and R.

Beckmann, *Structure of eEF3 and the mechanism of transfer RNA release from the E-site.* Nature, 2006. **443**(7112): p. 663-8.

20. Rodnina, M.V., H. Stark, A. Savelsbergh, H.J. Wieden, D. Mohr, N.B. Matassova, F. Peske, T. Daviter, C.O. Gualerzi, and W. Wintermeyer, *GTPases mechanisms and functions of translation factors on the ribosome.* Biol Chem, 2000. **381**(5-6): p. 377-87.

21. Diaconu, M., U. Kothe, F. Schlunzen, N. Fischer, J.M. Harms, A.G. Tonevitsky, H. Stark, M.V. Rodnina, and M.C. Wahl, *Structural basis for the function of the ribosomal L7/12 stalk in factor binding and GTPase activation.* Cell, 2005. **121**(7): p. 991-1004.

22. Wahl, M.C. and W. Moller, *Structure and function of the acidic ribosomal stalk proteins.* Curr Protein Pept Sci, 2002. **3**(1): p. 93-106.

23. Mohr, D., W. Wintermeyer, and M.V. Rodnina, *GTPase activation of elongation factors Tu and G on the ribosome.* Biochemistry, 2002. **41**(41): p. 12520-8.

24. Endo, Y. and I.G. Wool, *The site of action of alpha-sarcin on eukaryotic ribosomes. The sequence at the alpha-sarcin cleavage site in 28 S ribosomal ribonucleic acid.* J Biol Chem, 1982. **257**(15): p. 9054-60.

25. Endo, Y., K. Mitsui, M. Motizuki, and K. Tsurugi, *The mechanism of action of ricin and related toxic lectins on eukaryotic ribosomes. The site and the characteristics of the modification in 28 S ribosomal RNA caused by the toxins.* J Biol Chem, 1987. **262**(12): p. 5908-12.

26. 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.

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

28. Kozak, M., *Regulation of translation via mRNA structure in prokaryotes and eukaryotes.* Gene, 2005. **361**: p. 13-37.

29. 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.

30. 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.

31. Rodnina, M.V., K.B. Gromadski, U. Kothe, and H.J. Wieden, *Recognition and selection of tRNA in translation.* FEBS Lett, 2005. **579**(4): p. 938-42.

32. 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.

33. 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.

34. 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.

35. 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.

36. Myasnikov, A.G., A. Simonetti, S. Marzi, and B.P. Klaholz, *Structure-function insights into prokaryotic and eukaryotic translation initiation.* Curr Opin Struct Biol, 2009. **19**(3): p. 300-9.

37. 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.

38. 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.

39. 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.

40. 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.

41. 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.

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

43. 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.

44. 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.

45. Nyengaard, N.R., K.K. Mortensen, S.F. Lassen, J.W. Hershey, and H.U. Sperling-Petersen, *Tandem translation of E. coli initiation factor IF2 beta: purification and characterization in vitro of two active forms.* Biochem Biophys Res Commun, 1991. **181**(3): p. 1572-9.

46. 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.

47. 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.

48. Meunier, S., R. Spurio, M. Czisch, R. Wechselberger, M. Guenneugues, C.O. Gualerzi, and R. Boelens, *Structure of the fMet-tRNA(fMet)-binding domain of B. stearothermophilus initiation factor IF2.* Embo J, 2000. **19**(8): p. 1918-26.

49. Wienk, H., S. Tomaselli, C. Bernard, R. Spurio, D. Picone, C.O. Gualerzi, and R. Boelens, *Solution structure of the C1-subdomain of Bacillus stearothermophilus translation initiation factor IF2.* Protein Sci, 2005. **14**(9): p. 2461-8.

50. 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.

51. 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.

52. 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.

53. Sundari, R.M., E.A. Stringer, L.H. Schulman, and U. Maitra, *Interaction of bacterial initiation factor 2 with initiator tRNA.* J Biol Chem, 1976. **251**(11): p. 3338-45.

54. 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.

55. 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.

56. 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.

57. 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.

58. 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.

59. La Teana, A., C.L. Pon, and C.O. Gualerzi, *Late events in translation initiation. Adjustment of fMet-tRNA in the ribosomal P-site.* J Mol Biol, 1996. **256**(4): p. 667-75.

60. 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.

61. 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.

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. 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.

64. 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.

65. 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.

66. 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.

67. 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.

68. Elvekrog, M.M., *The Role of Initiation Factor Dynamics in Translation Initiation.* 2011, Columbia University.

69. Hirokawa, G., M.C. Kiel, A. Muto, M. Selmer, V.S. Raj, A. Liljas, K. Igarashi, H. Kaji, and A. Kaji, *Post-termination complex disassembly by ribosome recycling factor, a functional tRNA mimic.* Embo J, 2002. **21**(9): p. 2272-81.

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

71. 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.

72. Maivali, U. and J. Remme, *Definition of bases in 23S rRNA essential for ribosomal subunit association.* Rna, 2004. **10**(4): p. 600-4.

73. Ali, I.K., L. Lancaster, J. Feinberg, S. Joseph, and H.F. Noller, *Deletion of a conserved, central ribosomal intersubunit RNA bridge.* Mol Cell, 2006. **23**(6): p. 865-74.

74. 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.

75. 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.

76. 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.

77. Butler, J.S., M. Springer, and M. Grunberg-Manago, *AUU-to-AUG mutation in the initiator codon of the translation initiation factor IF3 abolishes translational autocontrol of its own gene (infC) in vivo.* Proc Natl Acad Sci U S A, 1987. **84**(12): p. 4022-5.

78. 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.

79. 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.

80. 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.

81. Huang, C., C.S. Mandava, and S. Sanyal, *The ribosomal stalk plays a key role in IF2-mediated association of the ribosomal subunits.* J Mol Biol, 2010. **399**(1): p. 145-53.

82. Helgstrand, M., C.S. Mandava, F.A. Mulder, A. Liljas, S. Sanyal, and M. Akke, *The ribosomal stalk binds to translation factors IF2, EF-Tu, EF-G and RF3 via a conserved region of the L12 C-terminal domain.* J Mol Biol, 2007. **365**(2): p. 468-79.

83. Pavlov, M.Y., A. Zorzet, D.I. Andersson, and M. Ehrenberg, *Activation of initiation factor 2 by ligands and mutations for rapid docking of ribosomal subunits.* Embo J, 2010. **30**(2): p. 289-301.

84. Zorzet, A., M.Y. Pavlov, A.I. Nilsson, M. Ehrenberg, and D.I. Andersson, *Error-prone initiation factor 2 mutations reduce the fitness cost of antibiotic resistance.* Mol Microbiol, 2010. **75**(5): p. 1299-313.

85. Sprang, S.R., *G proteins, effectors and GAPs: structure and mechanism.* Curr Opin Struct Biol, 1997. **7**(6): p. 849-56.

86. Kjeldgaard, M., J. Nyborg, and B.F. Clark, *The GTP binding motif: variations on a theme.* Faseb J, 1996. **10**(12): p. 1347-68.

87. Cool, R.H. and A. Parmeggiani, *Substitution of histidine-84 and the GTPase mechanism of elongation factor Tu.* Biochemistry, 1991. **30**(2): p. 362-6.

88. Luchin, S., H. Putzer, J.W. Hershey, Y. Cenatiempo, M. Grunberg-Manago, and S. Laalami, *In vitro study of two dominant inhibitory GTPase mutants of Escherichia coli translation initiation factor IF2. Direct evidence that GTP hydrolysis is necessary for factor recycling.* J Biol Chem, 1999. **274**(10): p. 6074-9.

89. Voorhees, R.M., T.M. Schmeing, A.C. Kelley, and V. Ramakrishnan, *The mechanism for activation of GTP hydrolysis on the ribosome.* Science, 2010. **330**(6005): p. 835-8.

90. Liljas, A., M. Ehrenberg, and J. Aqvist, *Comment on "The mechanism for activation of GTP hydrolysis on the ribosome".* Science, 2011. **333**(6038): p. 37; author reply 37.

91. Berchtold, H., L. Reshetnikova, C.O. Reiser, N.K. Schirmer, M. Sprinzl, and R. Hilgenfeld, *Crystal structure of active elongation factor Tu reveals major domain rearrangements.* Nature, 1993. **365**(6442): p. 126-32.

92. Lee, J.H., T.V. Pestova, B.S. Shin, C. Cao, S.K. Choi, and T.E. Dever, *Initiation factor eIF5B catalyzes second GTP-dependent step in eukaryotic translation initiation.* Proc Natl Acad Sci U S A, 2002. **99**(26): p. 16689-94.

93. Benne, R., N. Naaktgeboren, J. Gubbens, and H.O. Voorma, *Recycling of initiation factors IF-1, IF-2 and IF-3.* Eur J Biochem, 1973. **32**(2): p. 372-80.

94. Laalami, S., A.V. Timofeev, H. Putzer, J. Leautey, and M. Grunberg-Manago, *In vivo study of engineered G-domain mutants of Escherichia coli translation initiation factor IF2.* Mol Microbiol, 1994. **11**(2): p. 293-302.

95. 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.

96. 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.

97. 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.

98. La Teana, A., C.O. Gualerzi, and A.E. Dahlberg, *Initiation factor IF 2 binds to the alpha-sarcin loop and helix 89 of Escherichia coli 23S ribosomal RNA.* Rna, 2001. **7**(8): p. 1173-9.

99. Marzi, S., W. Knight, L. Brandi, E. Caserta, N. Soboleva, W.E. Hill, C.O. Gualerzi, and J.S. Lodmell, *Ribosomal localization of translation initiation factor IF2.* Rna, 2003. **9**(8): p. 958-69.

100. Iben, J.R. and D.E. Draper, *Specific interactions of the L10(L12)4 ribosomal protein complex with mRNA, rRNA, and L11.* Biochemistry, 2008. **47**(9): p. 2721-31.

101. Wimberly, B.T., R. Guymon, J.P. McCutcheon, S.W. White, and V. Ramakrishnan, *A detailed view of a ribosomal active site: the structure of the L11-RNA complex.* Cell, 1999. **97**(4): p. 491-502.

102. Xing, Y. and D.E. Draper, *Stabilization of a ribosomal RNA tertiary structure by ribosomal protein L11.* J Mol Biol, 1995. **249**(2): p. 319-31.

103. Pestka, S., *Thiostrepton: a ribosomal inhibitor of translocation.* Biochem Biophys Res Commun, 1970. **40**(3): p. 667-74.

104. Rosendahl, G. and S. Douthwaite, *The antibiotics micrococcin and thiostrepton interact directly with 23S rRNA nucleotides 1067A and 1095A.* Nucleic Acids Res, 1994. **22**(3): p. 357-63.

105. Porse, B.T., I. Leviev, A.S. Mankin, and R.A. Garrett, *The antibiotic thiostrepton inhibits a functional transition within protein L11 at the ribosomal GTPase centre.* J Mol Biol, 1998. **276**(2): p. 391-404.

106. Thompson, J., F. Schmidt, and E. Cundliffe, *Site of action of a ribosomal RNA methylase conferring resistance to thiostrepton.* J Biol Chem, 1982. **257**(14): p. 7915-7.

107. Harms, J.M., D.N. Wilson, F. Schluenzen, S.R. Connell, T. Stachelhaus, Z. Zaborowska, C.M. Spahn, and P. Fucini, *Translational regulation via L11: molecular switches on the ribosome turned on and off by thiostrepton and micrococcin.* Mol Cell, 2008. **30**(1): p. 26-38.

108. Jonker, H.R., S. Ilin, S.K. Grimm, J. Wohnert, and H. Schwalbe, *L11 domain rearrangement upon binding to RNA and thiostrepton studied by NMR spectroscopy.* Nucleic Acids Res, 2007. **35**(2): p. 441-54.

109. Cameron, D.M., J. Thompson, P.E. March, and A.E. Dahlberg, *Initiation factor IF2, thiostrepton and micrococcin prevent the binding of elongation factor G to the Escherichia coli ribosome.* J Mol Biol, 2002. **319**(1): p. 27-35.

110. Brandi, L., S. Marzi, A. Fabbretti, C. Fleischer, W.E. Hill, C.O. Gualerzi, and J. Stephen Lodmell, *The translation initiation functions of IF2: targets for thiostrepton inhibition.* J Mol Biol, 2004. **335**(4): p. 881-94.

111. Grunberg-Manago, M., J. Dondon, and M. Graffe, *Inhibition by thiostrepton of the IF-2-dependent ribosomal GTPase.* FEBS Lett, 1972. **22**(2): p. 217-221.

112. Lockwood, A.H., P. Sarkar, U. Maitra, N. Brot, and H. Weissbach, *Effect of thiostrepton on polypeptide chain initiation in Escherichia coli.* J Biol Chem, 1974. **249**(18): p. 5831-4.

113. Mazumder, R., *Effect of thiostrepton on recycling of Escherichia coli initiation factor 2.* Proc Natl Acad Sci U S A, 1973. **70**(7): p. 1939-42.

114. Naaktgeboren, N., P. Schrier, W. Moller, and H.O. Voorma, *The involvement of protein L11 in the joining of the 30-S initiation complex to the 50-S subunit.* Eur J Biochem, 1976. **62**(1): p. 117-23.

115. Kavran, J.M. and T.A. Steitz, *Structure of the base of the L7/L12 stalk of the Haloarcula marismortui large ribosomal subunit: analysis of L11 movements.* J Mol Biol, 2007. **371**(4): p. 1047-59.

116. Agrawal, R.K., J. Linde, J. Sengupta, K.H. Nierhaus, and J. Frank, *Localization of L11 protein on the ribosome and elucidation of its involvement in EF-G-dependent translocation.* J Mol Biol, 2001. **311**(4): p. 777-87.

117. Li, W., J. Sengupta, B.K. Rath, and J. Frank, *Functional conformations of the L11-ribosomal RNA complex revealed by correlative analysis of cryo-EM and molecular dynamics simulations.* Rna, 2006. **12**(7): p. 1240-53.

118. Klein, D.J., T.M. Schmeing, P.B. Moore, and T.A. Steitz, *The kink-turn: a new RNA secondary structure motif.* Embo J, 2001. **20**(15): p. 4214-21.

119. Rawat, U., H. Gao, A. Zavialov, R. Gursky, M. Ehrenberg, and J. Frank, *Interactions of the release factor RF1 with the ribosome as revealed by cryo-EM.* J Mol Biol, 2006. **357**(4): p. 1144-53.

120. Valle, M., A. Zavialov, W. Li, S.M. Stagg, J. Sengupta, R.C. Nielsen, P. Nissen, S.C. Harvey, M. Ehrenberg, and J. Frank, *Incorporation of aminoacyl-tRNA into the ribosome as seen by cryo-electron microscopy.* Nat Struct Biol, 2003. **10**(11): p. 899-906.

121. DeLano, W., *The PyMOL Molecular Graphics System, Version 0.99, Schrodinger, LLC*.

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

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

124. Lakowicz, J.R., *Energy Transfer*, in *Principles of Fluorescence Spectroscopy, Third Edition*. 2006, Springer Science+Business Media, LLC: New York. p. 443-75.

125. Cantor, C.R. and P.R. Schimmel, *Fluorescence Spectroscopy*, in *Biophysical Chemistry*. 1980, W. H. Freeman and Company. p. 433-65.

126. MacDougall, D.D., J. Fei, and R.L. Gonzalez, Jr., *Single-Molecule Fluorescence Resonance Energy Transfer Investigations of Ribosome-Catalyzed Protein Synthesis*, in *Molecular Machines in Biology*, J. Frank, Editor. 2011, Cambridge University Press. p. 93-116.

127. 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/Wein. p. 273-294.

128. Axelrod, D., N.L. Thompson, and T.P. Burghardt, *Total internal inflection fluorescent microscopy.* J Microsc, 1983. **129**(Pt 1): p. 19-28.

129. 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.

Chapter 2

Interaction of IF2 with the GTPase-associated center during 70S IC formation

2.1 Introduction

Following correct assembly of a 30S IC containing fMet-tRNA^{fMet} bound to the mRNA start codon, 50S subunit joining is catalyzed by IF2 [1, 2]. The subunit joining event represents an important step in the translation initiation pathway that is regulated in order to prevent formation of improperly assembled ribosomal complexes and to selectively accelerate the formation of properly assembled 70S ICs [3-5]. This chapter reports the development of an experimental platform to monitor IF2-catalyzed subunit joining in real time at the single-molecule level, based upon a FRET signal between IF2 and r-protein L11. The design and generation of fluorescently labeled constructs is discussed (Sections 2.2 and 2.3), as well as the results from biochemical assays that demonstrate full biochemical activity of these components (Section 2.4). The smFRET labeling scheme was initially validated in the context of a stable 70S IC containing IF2 bound to the non-hydrolyzable GTP analog GDPNP (Section 2.5). Then, an experimental platform was developed for observing real-time subunit joining reactions. The results from these experiments demonstrate that this approach is useful for studying the rate of subunit joining, the lifetime of IF2 on the ribosome prior to dissociation, and conformational dynamics between L11 and IF2 within the 70S IC (Section 2.6). Finally, IF2 has been shown to hydrolyze GTP in a ribosome-dependent, multiple-turnover reaction that is uncoupled from translation initiation [6]; using the IF2-L11 smFRET

signal, I demonstrate that IF2's interactions with the GAC during multiple-turnover GTP hydrolysis are much different from those that occur during 50S subunit joining and 70S IC formation (Section 2.7).

Bulk fluorescence experiments have suggested that, following subunit joining, large-scale conformational changes of the ribosome and its initiation factor and tRNA ligands are involved in formation of the elongation-competent 70S IC [7, 8]. The known conformational flexibility of the L11 region of the 50S GAC (see Chapter 1 and references cited therein) suggested that L11 dynamics may be involved in this process. By monitoring the interactions between L11 and IF2, therefore, I hoped to take a first step toward characterizing GAC dynamics during 70S IC formation and understanding the mechanistic role that L11 plays during this process. Following 70S IC assembly one molecule at a time has provided direct access to static and dynamic heterogeneity within the population of initiating ribosomes, allowing, for example, detection and characterization of stochastic conformational fluctuations between L11 and IF2 prior to IF2 dissociation.

2.2 Design of smFRET probes

In order to monitor the interaction between IF2 and the ribosome's GAC during translation initiation, smFRET probes were site-specifically attached to appropriate positions on the surface of IF2 and r-protein L11. The distance between donor and acceptor fluorophores should ideally be close to the Förster distance (R_0, ~60 Å for the Cy3-Cy5 FRET pair) [9, 10], where small changes in inter-fluorophore distance result in

large changes in FRET efficiency. Currently, only low-resolution cryo-EM reconstructions of IF2 in the context of a 70S IC are available [11, 12], which allowed only an approximate prediction of the distance between L11 and IF2. Therefore, to maximize the probability of obtaining a mechanistically informative labeling scheme, three candidate labeling positions were chosen on IF2 at residues Arg561, Ser566, and Ser672 (the amino acid numbering convention used here assigns as residue 1 the N-terminal methionine of IF2-α) (Figure 2.1).

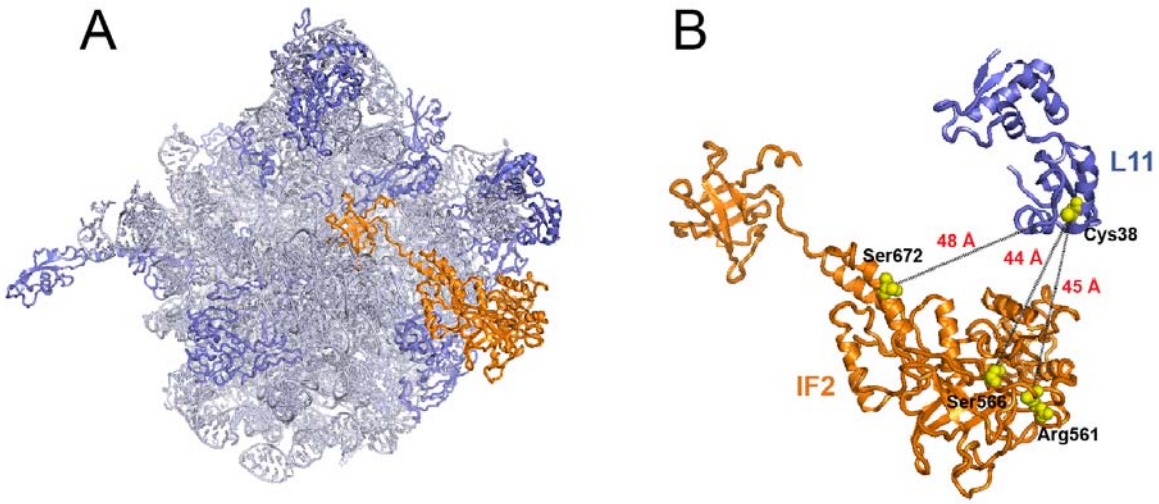

Figure 2.1: Fluorophore labeling positions on L11 and IF2.
(A) Model of the relative orientation of IF2 and the 50S subunit within the 70S IC. The model was constructed through the superposition of multiple cryo-EM and X-ray crystallographic structures in order to approximate distances between potential fluorophore labeling positions on L11 and IF2. PDB files used to construct the model were 1ZO1, 1ZO3, 2J00, 2J01, 2AW7, and 2AWB. The 50S subunit (PDB ID: 2AWB) is shown in an interface view with rRNA colored gray and r-proteins colored in lavender. IF2 (PDB ID: 1ZO1) is colored in orange. (B) Based on this model, approximate distances between labeling positions on L11 and IF2 were ~45 Å from Cys38 to Arg561, ~44 Å from Cys38 to Ser566, and ~48 Å from Cys38 to Ser672.

All three residues are located within IF2's domain V, which is C-terminal to the GTP-binding domain [13]. Labeling of the G domain itself was avoided in order to minimize

the risk of fluorophore labeling interfering with IF2's GTP hydrolysis activity. These three residues were chosen due to their low level of conservation in bacteria, as well as their expected surface accessibility and proximity to L11 in the 70S IC (Figure 2.1). r-protein L11 contains a single, wild-type cysteine residue (Cys38) in its NTD, which was chosen as the L11 labeling position.

2.3 Generation of fluorescently labeled constructs

2.3.1 (Cy3/Cy5)-L11 50S subunits

Preparation of site-specifically labeled 50S subunits was a two-step process that involved fluorescent labeling of L11 with Cy3 or Cy5 followed by *in vitro* reconstitution of (Cy3/Cy5)-L11 with 50S subunits lacking L11, termed L11(-). First, the *rplK* gene encoding r-protein L11 was PCR-amplified from *E. coli* genomic DNA and cloned into an overexpression vector that introduces a six-histidine (6xHis) affinity tag followed by a tobacco etch virus (TEV) protease cleavage site at the N-terminus of the cloned protein (Section 5.1.4.1). Recombinant, 6xHis-tagged L11 was overexpressed and affinity-purified over a Ni^{2+}-nitrilotriacetic acid (NTA) column under denaturing conditions, followed by protein renaturation, cleavage of the 6xHis-tag with TEV protease, and removal of the cleaved 6xHis-tags with a second Ni^{2+}-NTA column (Section 5.1.4.2). Purified L11 was then fluorescently labeled at its unique cysteine residue (Cys38) by reaction with Cy3 or Cy5-maleimide conjugated fluorophores to generate (Cy3/Cy5)-L11. Separation of (Cy3/Cy5)-L11 from free, unreacted dye was achieved by gel

filtration chromatography, which was incidentally found to separate unlabeled- and (Cy3/Cy5)-labeled species (Section 5.1.4.3).

L11(-) ribosomes were purified from the strain NVD005, a derivative of *E. coli* K-12 from which the *rplK* gene has been deleted from the chromosome [14, 15]. The corresponding wild-type strain NVD001 was used to purify wild-type ribosomes using the same methodology. Tight-coupled 70S ribosomes were first purified using sucrose density gradient ultracentrifugation. These were subsequently split into 30S and 50S subunits by resuspension in low (1 mM) Mg^{2+} buffer and isolated through another round of sucrose density gradient ultracentrifugation. L11(-) 50S subunits were then reconstituted with (Cy3/Cy5)-L11 and purified through another sucrose density gradient, thereby generating 50S subunits site-specifically labeled at Cys38 of L11's NTD (Section 5.1.5). L11(-) 50S subunits were also reconstituted with recombinant, unlabeled L11 as a control to test for the effect of fluorophore labeling in the biochemical assays described below.

2.3.2 (Cy3/Cy5)-IF2

Site-directed mutagenesis of wild-type IF2 (γ isoform) was performed in order to generate R561C, S566C, and S672C mutants (Section 5.1.3.1). These point mutants were purified using a combination of Ni^{2+}-NTA affinity chromatography and cation-exchange chromatography (Section 5.1.3.2). Wild-type IF2 contains three cysteine residues, at positions 599 (domain V), 815 (domain VI-2), and 861 (domain VI-2), which have previously been shown in our laboratory to be inaccessible to the fluorophore labeling

reaction under particular conditions [16]. These conditions were employed to fluorescently label each of the three mutants with Cy3/Cy5; labeled IF2 was then separated from free, unreacted dye by gel filtration chromatography (Section 5.1.3.3). A negative-control labeling reaction using wild-type IF2 demonstrates highly specific labeling of the engineered cysteines (Figure 2.2). Fluorophore labeling efficiencies were estimated to be ~60% for IF2 R561C, ~85% for IF2 S566C, and ~90% for IF2 S672C.

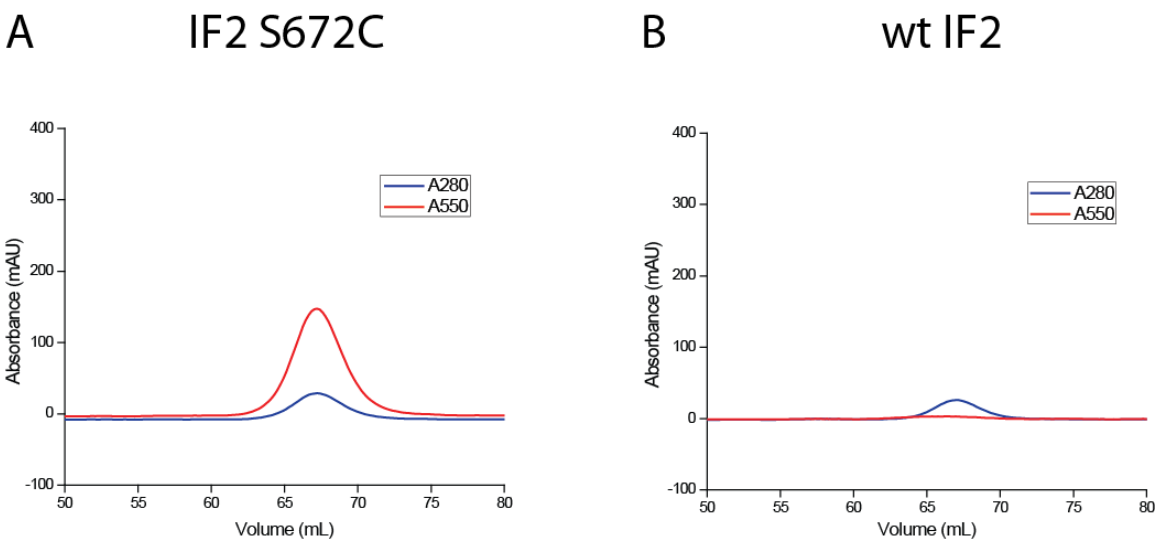

Figure 2.2: Site-specific fluorescent labeling of IF2.
Mutant and wild-type IF2 were incubated with Cy3-maleimide under identical reaction conditions, followed by separation of IF2 from free, unreacted dye using gel filtration chromatography. IF2 elutes at ~67 mL and was detected by absorbance at 280 nm (blue curve). Cy3 was detected by absorbance at 550 nm (red curve). AU=arbitrary units. (A) Chromatogram from the IF2 S672C labeling reaction. Based on integration of the A280 and A550 peaks, the labeling efficiency for IF2 S672C was ~90%. (B) Chromatogram from the wild-type IF2 labeling reaction. The absence of an A550 peak co-migrating with wild-type IF2 demonstrates that the three wild-type cysteines are labeled with very low efficiency, if at all.

2.4 Biochemical testing of fluorescently labeled components

Following generation and purification of Cy3/Cy5-labeled IF2 and 50S subunits, it was necessary to test whether the fluorophores and/or labeling procedures affected their biochemical activities. Three independent activity assays were performed that tested IF2's ribosome-dependent GTP hydrolysis activity (Section 2.4.1), IF2-promoted selection of fMet-tRNA^{fMet} on the 30S IC (Section 2.4.2), and IF2-mediated formation of an elongation-competent 70S IC (Section 2.4.3). The fluorescently labeled constructs exhibited wild-type levels of activity in all three assays. A low-salt version of the standard Tris-polymix buffer system developed for *in vitro* translation work [17-19] was employed in all biochemical experiments. It contains 10 mM Tris-acetate ($pH_{25°C}$ = 7.5), 20 mM KCl, 5-15 mM $Mg(OAc)_2$, 1 mM NH_4OAc, 0.1 mM $Ca(OAc)_2$, 0.1 mM EDTA, 6 mM β-mercaptoethanol, 5 mM putrescine-HCl, and 1 mM spermidine free-base. Detailed protocols for these assays can be found in Section 5.2 of the Materials and Methods.

2.4.1 GTP hydrolysis assay

A GTP-hydrolysis assay was performed to test the ability of fluorescently labeled IF2 to hydrolyze GTP in a ribosome-dependent manner [20]. In this assay, radiolabeled [α-^{32}P]GTP is incubated with IF2 in the presence or absence of ribosomes. GTP hydrolysis by IF2 leads to the accumulation of [α-^{32}P]GDP, which can be separated from unreacted [α-^{32}P]GTP by thin layer chromatography (TLC) and quantified by phosphorimaging. This is a multiple-turnover assay, in which IF2 recycles on and off the ribosome. The multiple-turnover GTP hydrolysis activity of IF2 has been shown to be

affected by the presence or absence of L11, with L11(-) ribosomes eliciting less GTP hydrolysis than wild-type ribosomes, apparently without affecting IF2's ribosomal binding affinity [21]. Therefore, this assay can report on the biochemical activity of fluorescently labeled IF2 as well as (Cy3/Cy5)-L11 reconstituted ribosomes.

The results from a timecourse are shown in Figure 2.3. High levels of GTP hydrolysis requires the presence of both ribosomes and IF2, while incubation of [α-^{32}P]GTP with ribosomes or IF2 alone results in only basal levels of hydrolysis. L11(-) ribosomes were found to be impaired in their ability to stimulate multiple-turnover GTP hydrolysis, with the observed four-fold effect in complete agreement with that reported by Brandi, et al [21]. The dual-labeled system, consisting of (Cy3)-IF2 and ribosomes reconstituted with (Cy5)-L11, exhibits wild-type levels of GTP hydrolysis, indicating both a lack of interference of the fluorescent labels with this biochemical activity and a high (Cy5)-L11 reconstitution efficiency.

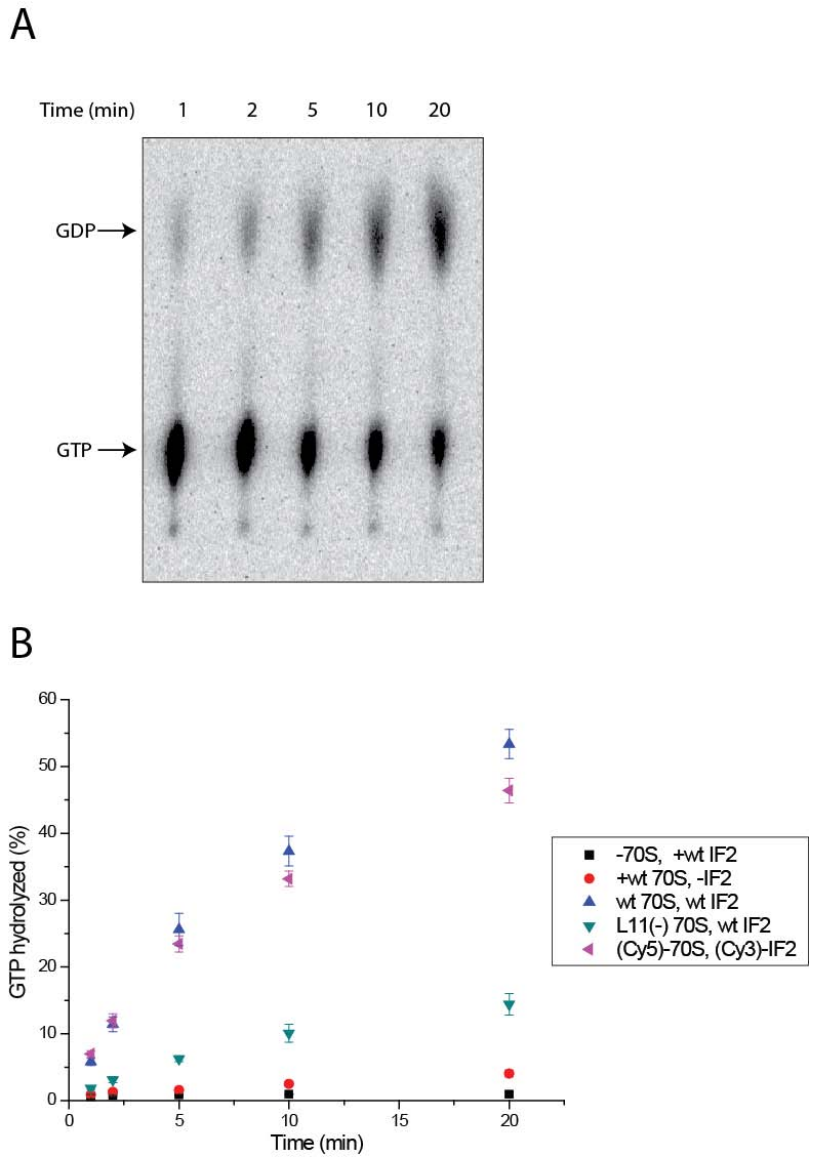

Figure 2.3: GTP hydrolysis assay.
(A) Sample phosphor image showing the conversion of [α-^{32}P]GTP to [α-^{32}P]GDP. 30S and 50S ribosomal subunits (0.4 μM) were incubated with IF2 (0.8 μM) and [α-^{32}P]GTP (25 μM) in Low-Salt Tris-polymix buffer at room temperature. Reactions were quenched at the indicated time points, and [α-^{32}P]GDP was separated from [α-^{32}P]GTP by TLC. (B) Timecourse of GTP hydrolysis with different combinations of ribosomes and IF2. wt70S refers to ribosomes from strain NVD001, L11(-) 70S refers to ribosomes purified from strain NVD005, (Cy5)-70S refers to NVD005 ribosomes reconstituted with (Cy5)-L11, and (Cy3)-IF2 corresponds to the Cy3-labeled S672C mutant. Phosphor images were quantified using ImageQuant, and percent GTP hydrolyzed was calculated as intensity of the GTP spot divided by the sum of GTP and GDP spots, multiplied by 100. Error bars represent the standard deviation of three independent measurements.

2.4.2 Selection of fMet-tRNA^{fMet} on the 30S IC

IF2 plays an important role in the recruitment of fMet-tRNA^{fMet} to the 30S ribosomal subunit as well as its stabilization on the 30S IC [1, 22]. Even in the absence of the other initiation factors, IF2 imparts selection of fMet-tRNA^{fMet} over elongator tRNAs at the level of 30S IC assembly. The ability of IF2 to select fMet-tRNA^{fMet} on the 30S IC can be tested using a primer extension inhibition assay called "toeprinting" [23, 24]. In this assay, the 30S IC is assembled on an mRNA that has been pre-annealed with a ^{32}P-labeled DNA primer complementary to the mRNA's 3' end. Following 30S IC assembly, reverse transcriptase is added to extend the ^{32}P-labeled DNA primer, thus generating ^{32}P-labeled cDNA products. Extension of the cDNA is halted when the reverse transcriptase encounters a ribosomal complex bound to the mRNA; thus, the length of cDNA product generated provides a readout for the position of the 30S IC on the mRNA (Figure 2.4).

The mRNA used in the toeprinting experiments contains an AUG start codon, which codes for tRNA^{fMet}, followed by a UUC codon, which codes for tRNA^{Phe}. All mRNAs used in this thesis were derived from the mRNA encoding gene product 32 from T4 bacteriophage (T4gp32), and their full sequences are reported in Appendix A. In the absence of initiation factors, the 30S subunit can form a complex on the mRNA with either tRNA^{fMet} or tRNA^{Phe} bound to their cognate codons in the P site. If the 30S IC contains tRNA^{fMet} bound to the AUG start codon at the P site, reverse transcription is halted at a position 15 nucleotides downstream of the first nucleotide of the start codon, thereby generating a +15 toeprint. If, instead, the 30S IC contains tRNA^{Phe} bound to the

UUC codon at the P site, reverse transcription is halted three nucleotides sooner, thereby generating a +18 toeprint (Figure 2.4A).

In the toeprinting reactions, 30S subunits and ^{32}P-primer annealed T4gp32 mRNA were incubated with an equimolar mixture of tRNA^{Phe} and fMet-tRNA^{fMet} in the absence or presence of IF2. Subsequent reverse transcription yielded a mixture of ^{32}P-labeled cDNA products, which were separated on a 9% sequencing PAGE gel (Figure 2.4B). The relative intensities of the +15 and +18 toeprints reports on the relative efficiency of initiation with fMet-tRNA^{fMet} versus tRNA^{Phe}. In the absence of IF2, roughly equal intensities of the +15 and +18 toeprints were observed (lane 1), demonstrating that the 30S subunit by itself does not discriminate fMet-tRNA^{fMet} from tRNA^{Phe}. Addition of IF2, however, leads to preferential selection of fMet-tRNA^{fMet}, as indicated by a large shift in the +15/+18 ratio towards the +15 toeprint (lane 2). Furthermore, the intensity of the +15 toeprint in the presence of IF2 is higher than the total initiation signal (i.e. the sum of +15 and +18 toeprints) in the absence of IF2, which is consistent with an enhancement of fMet-tRNA^{fMet} binding and 30S IC stability promoted by IF2. Similar results were obtained in the presence of either mutant IF2 (lane 3) or fluorescently labeled IF2 (lane 4) constructs, demonstrating that neither mutation nor fluorophore labeling impair the ability of IF2 to select initiator tRNA on the 30S IC.

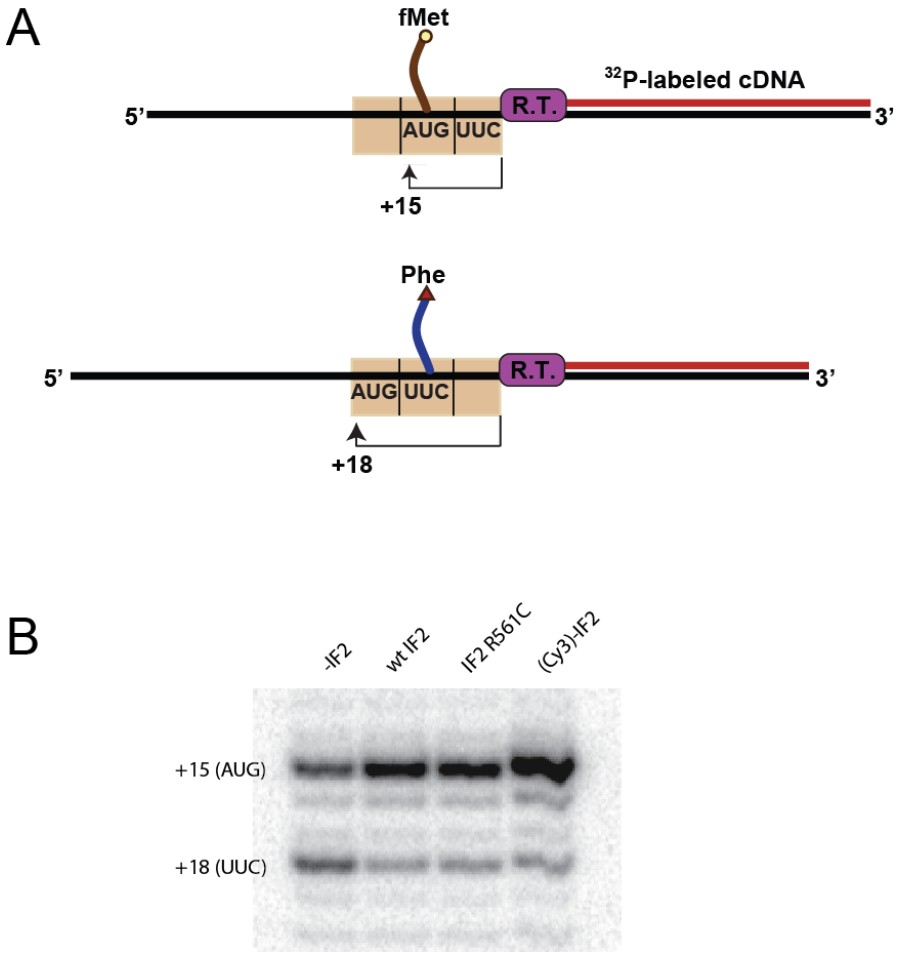

Figure 2.4: Toeprinting activity assay.
(A) Cartoon depiction of the toeprinting assay. See text for details. "R.T." denotes reverse transcriptase. (B) Initiation complex formation reactions were performed by incubating 30S ribosomal subunits (0.7 μM) with ^{32}P-primer annealed mRNA (0.3 μM), fMet-tRNA^{fMet} (1 μM) and tRNA^{Phe} (1 μM) for 10 min at 37°C in the presence or absence of IF2 (7 μM). Reverse transcription reactions were then performed by adding reverse transcriptase and dNTPs and incubating for 15 min at 37°C. cDNA products from the reverse transcription reaction were resolved on a 9% sequencing PAGE gel, which was dried and used to expose a Phosphor Imaging screen. The region of the resulting phosphor image containing the +15 and +18 bands, corresponding to initiation at the AUG codon with fMet-tRNA^{fMet} and initiation at the UUC codon with tRNA^{Phe}, respectively, is shown. Lane 1: Reaction was performed in the absence of IF2. Lanes 2-4: Reactions were performed in the presence of wild-type IF2, unlabeled IF2 R561C, or Cy3-labeled R561C IF2 as indicated.

2.4.3 Dipeptide formation assays

The ability of IF2 to promote formation of a correctly assembled 70S IC capable of entering into elongation was tested using a dipeptide formation assay. First, 30S ICs were prepared by incubating 30S ribosomal subunits (0.9 μM), IF1 (0.9 μM), mRNA (1.8 μM), and radiolabeled ^{35}S-fMet-tRNA^{fMet} (0.6 μM), in the presence or absence of IF2 (0.9 μM), in Low-Salt Tris-polymix buffer supplemented with GTP (1 mM). The procedure for 30S IC formation employed here was the same as that used to prepare fluorescently labeled 30S ICs for microscope experiments, with the exception that non-biotin mRNA was used and ^{35}S-fMet-tRNA^{fMet}, rather than (Cy3)-IF2, was limiting. The pre-formed 30S ICs were then reacted with a mixture of 50S subunits and either EF-Tu:GTP:Phe-tRNA^{Phe} ternary complex or the antibiotic puromycin. Puromycin is an analog of the 3'-acceptor end of aminoacyl-tRNA. It binds to the 50S subunit at the peptidyl transferase center, participates in peptide bond formation, thereby deacylating the P-site tRNA, and subsequently dissociates from the ribosome [25]. In this assay, before peptidyl transfer to puromycin can occur, the 50S subunit must dock with the 30S IC and ^{35}S-fMet-tRNA^{fMet} must be accommodated into the P site. Thus, formation of fMet-puromycin is commonly used to monitor the completion of 70S IC formation and the proper positioning of fMet-tRNA^{fMet} in the peptidyl transferase center [6, 26]. When EF-Tu:GTP:Phe-tRNA^{Phe} ternary complex is used instead of puromycin, peptide bond formation requires, in addition to 50S subunit joining and placement of fMet-tRNA^{fMet} in the P site, EF-Tu–catalyzed accommodation of Phe-tRNA^{Phe} into the A site.

Peptidyl transfer results in deacylation of the P-site ^{35}S-fMet-tRNA^{fMet} and formation of either ^{35}S-fMet-puromycin or ^{35}S-fMet-Phe dipeptide. The radiolabeled dipeptide products can be separated from unreacted ^{35}S-fMet using electrophoretic TLC (eTLC) [27]. The results from eTLC analysis of dipeptide formation time courses are shown in Figure 2.5. The reactions were largely complete at the earliest, 15 sec time point, which precluded analysis of their initial rates. The extent of both fMet-puromycin and fMet-Phe formation, however, was shown to depend strongly on IF2, whose presence stimulated dipeptide formation ~5-fold. The low levels of dipeptide formation observed in the absence of IF2 are likely a consequence of inefficient 50S subunit joining. This interpretation is consistent with light scattering data showing that formation of 70S complexes from 30S and 50S subunits is highly dependent on the presence of both IF2 and fMet-tRNA^{fMet} [4, 8]. Fluorescent labeling of neither IF2 nor L11 interferes with 70S IC formation, as the dipeptide formation time courses generated using (Cy3)-IF2 and (Cy5)-L11 reconstituted 50S subunits are indistinguishable from those obtained with wild-type ribosomes and IF2. Furthermore, the extent of fMet-puromycin and fMet-Phe formation observed here is comparable to that observed previously under similar conditions [6, 26]. The eTLC results thus demonstrate full activity of fluorescently labeled components in all of the biochemical steps leading to formation of a 70S IC that is competent to enter into the elongation phase of protein synthesis.

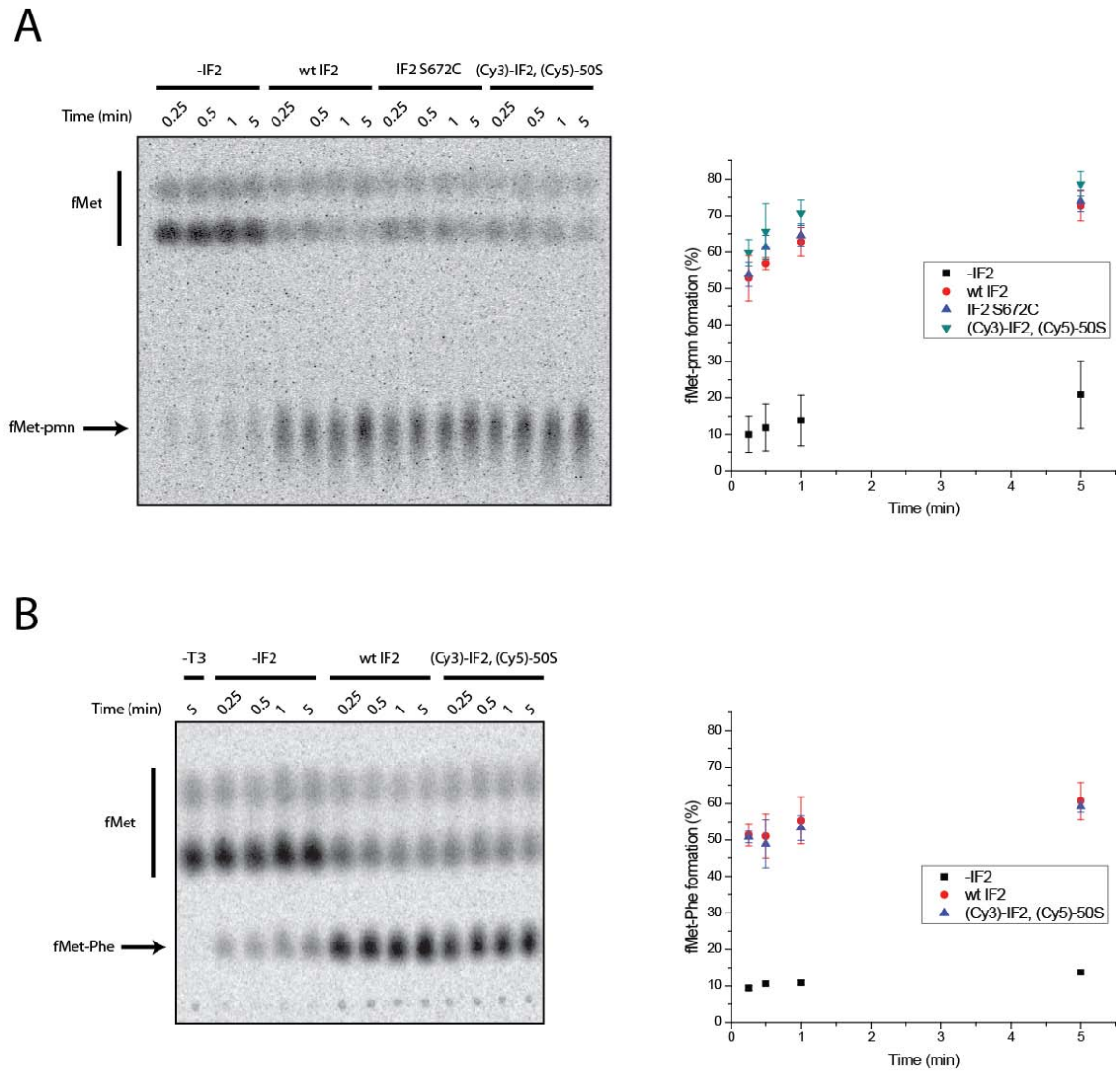

Figure 2.5: eTLC analysis of dipeptide formation.
(A) fMet-puromycin reaction. Preformed 30S ICs (1.5 pmol) were mixed with 50S subunits (2.25 pmol) and puromycin (10 nmol) in Low-Salt Tris-polymix buffer. The reaction was incubated at room temperature and quenched with base at 15 sec, 30 sec, 1 min, and 5 min time points. Dipeptide fMet-puromycin (fMet-pmn) was separated from unreacted fMet by eTLC (left panel). The two fMet spots correspond to oxidized and reduced forms of the amino acid. Reactions contained 30S ICs formed in the absence of IF2 (Lanes 1-4), wild-type IF2 (Lanes 5-8), IF2 S672C (Lanes 9-12), and (Cy3)-IF2 S672C (Lanes 13-16). Wild-type 50S subunits were used in all reactions except those in Lanes 13-16, where (Cy5)-L11 reconstituted 50S subunits were used. Phosphor images were quantified and the percent of fMet converted to dipeptide was calculated as intensity of the fMet-pmn spot divided by the sum of fMet-pmn and unreacted fMet spots, multiplied by 100 (right panel). Plotted data represents the mean and standard deviation from three independent experiments. (B) fMet-Phe reaction. Preformed 30S ICs (1.5 pmol) were mixed with 50S subunits (2.25 pmol) and preformed EF-Tu:GTP:Phe-tRNA^{Phe} ternary complex (6 pmol)

in Low-Salt Tris-polymix buffer. The reaction was incubated at room temperature and quenched with base at 15 sec, 30 sec, 1 min and 5 min time points, and formation of fMet-Phe was monitored by eTLC (left panel). Reactions contained 30S ICs formed in the absence of IF2 (Lanes 2-5), wild-type IF2 (Lanes 6-9), and (Cy3)-IF2 S672C (Lanes 10-13). Wild-type 50S subunits were used in all reactions except those in Lanes 10-13, where (Cy5)-L11 reconstituted 50S subunits were used. Lane 1 is a negative control in which Phe-tRNA^{Phe} ternary complex (T3) was omitted. Phosphor images were quantified and the percent of fMet converted to dipeptide was calculated as intensity of the fMet-Phe spot divided by the sum of fMet-Phe and unreacted fMet spots, multiplied by 100 (right panel). Plotted data represents the mean and standard deviation from three independent experiments.

2.5 Characterization of L11-IF2 smFRET signals within 70SIC$_{GDPNP}$

Initial steady-state smFRET measurements were performed on 70S ICs formed using (Cy3)-IF2 bound to the non-hydrolyzable GTP analog GDPNP. The GDPNP-bound form of IF2 promotes 50S subunit joining to the 30S IC, but remains stably bound to the ribosome following 70S IC formation [12, 28]. It was thus expected that (Cy3)-IF2-GDPNP should bind stably enough to 70S ICs to allow surface immobilization of ribosomal complexes and smFRET imaging for an extended period of time (~minutes) prior to (Cy3)-IF2-GDPNP dissociation. These 70S ICs, referred to here as 70SIC$_{GDPNP}$, were formed using (Cy5)-L11 reconstituted 50S subunits and provided a robust platform for initial validation and preliminary characterization of the three IF2-L11 smFRET signals.

70SIC$_{GDPNP}$ complexes were formed in a two-step process comprising 30S IC assembly and 50S subunit joining (Section 5.3.2). They were then surface immobilized and imaged under steady-state conditions (Section 5.4.2) in order to confirm that the chosen fluorophore labeling positions on the surface of IF2 are indeed within FRET distance of the fluorophore attached to L11's NTD in the context of a 70S IC. Complexes

were immobilized via their biotinylated-mRNA to the surface of a streptavidin-coated flowcell and (Cy3)-IF2 was directly excited with 532 nm TIR illumination. Cy5 emission via FRET was observed for $70SIC_{GDPNP}$ complexes formed with all three of the individual (Cy3)-IF2 constructs (Figure 2.6). The E_{FRET} distributions were centered at 0.64 FRET for (Cy3)-IF2 R561C, 0.74 FRET for (Cy3)-IF2 S566C, and 0.68 FRET for (Cy3)-IF2 S672C. The similarity in E_{FRET} distributions among the three (Cy3)-IF2/(Cy5)-L11 signals suggests that the distance between the L11 NTD and the three IF2 labeling positions is similar within $70SIC_{GDPNP}$; assuming a Förster distance of 60 Å and free rotation of the fluorophores (i.e. $\kappa^2 = 2/3$) [29, 30], these E_{FRET} ratios suggest an inter-fluorophore distance of ~50-55 Å.

Based on comparison of the data collected for the three (Cy3)-IF2 constructs within $70SIC_{GDPNP}$, (Cy3)-IF2 S672C was chosen for the majority of the smFRET experiments presented in this thesis, for two main reasons. First, the (Cy3)-IF2 R561C construct was ruled out because of its relatively low brightness and signal-to-noise ratio, which is probably caused by an unfavorable local environment of the Cy3 fluorophore that leads to more rapid fluorophore quenching. Second, the average E_{FRET} ratio of 0.68 observed for (Cy3)-IF2 S672C was preferred over the E_{FRET} ratio of 0.74 observed for (Cy3)-IF2 S566C, since it is closer to 0.5, where E_{FRET} is most sensitive to small changes in inter-fluorophore distance. Thus, the smFRET signal between (Cy3)-IF2 S672C and (Cy5)-L11 might be expected to provide a more sensitive probe of conformational change between IF2 and the GAC.

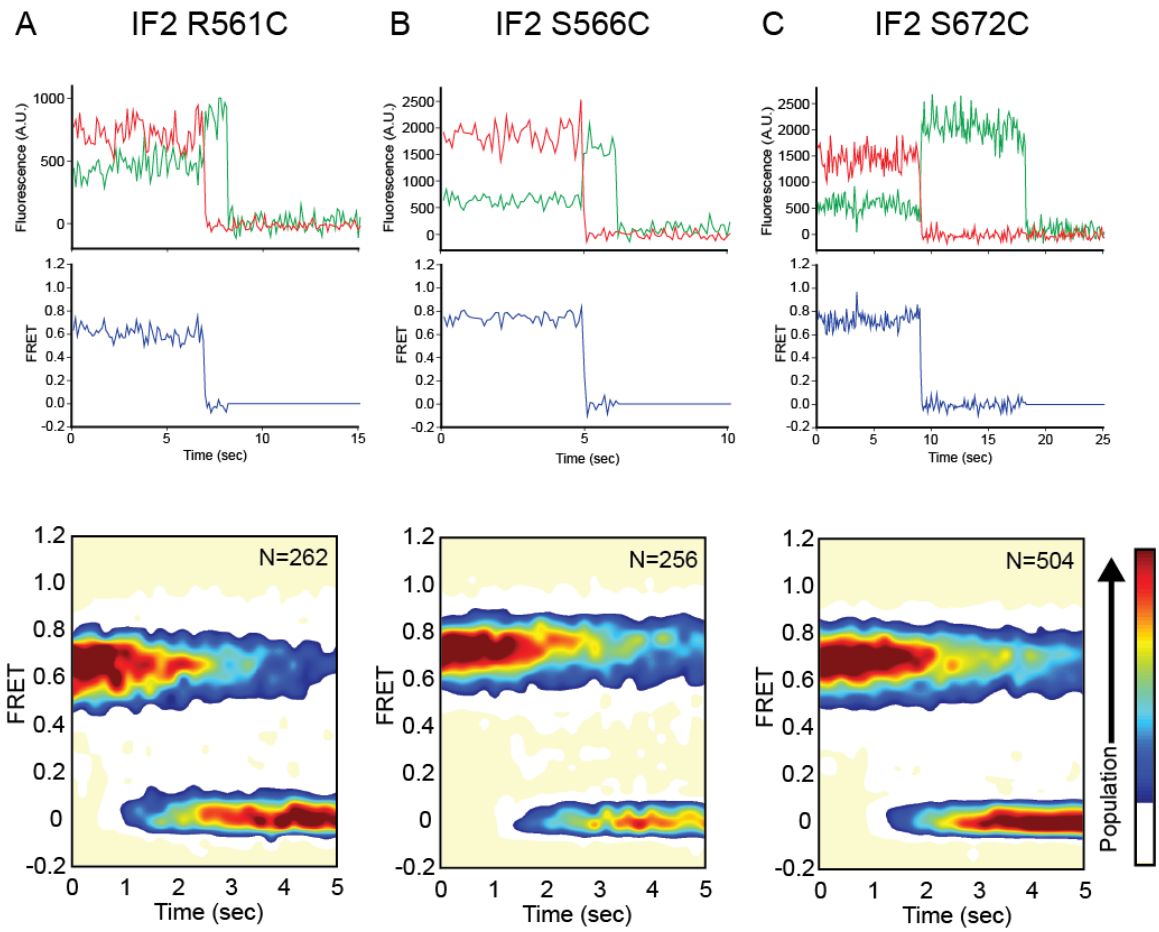

Figure 2.6: Steady-state FRET measurements on $70SIC_{GDPNP}$.
$70SIC_{GDPNP}$ complexes were prepared which contained 30S subunits, (Cy5)-L11 reconstituted 50S subunits, biotin-mRNA, fMet-$tRNA^{fMet}$, IF1, and either (Cy3)-IF2-GDPNP R561C (A), S566C (B), or S672C (C). The preformed complexes were then surface immobilized and imaged using TIRF microscopy. Top Row: Representative Cy3 (green) and Cy5 (red) fluorescence intensity versus time trajectories. Second row: The corresponding smFRET versus time trajectories, where FRET is calculated as $I_{Cy5}/(I_{Cy3} + I_{Cy5})$. Third row: Time evolution of population FRET histograms, generated by superimposing all of the individual smFRET versus time trajectories. Surface contours are plotted from tan (lowest population) to red (highest population) as indicated by the color bar. The number of trajectories used to construct each contour plot is indicated by "N."

2.6 Real-time observation of IF2-catalyzed 50S subunit joining

2.6.1 Development of experimental platform

Having validated the IF2-L11 smFRET signal, I next aimed to use it for characterization of the interaction between IF2 and the GAC during IF2-catalyzed docking of the 50S subunit to the 30S IC and subsequent formation of the 70S IC. This was accomplished by surface-immobilization of 30S ICs containing (Cy3)-IF2-GTP, followed by stopped-flow delivery of (Cy5)-L11 labeled 50S subunits and observation of the time-evolution of the smFRET signal. Within this experimental set-up, the 50S subunit docking event should be signaled by the appearance of Cy5 fluorescence resulting from FRET between (Cy3)-IF2 and (Cy5)-L11. Subsequent fluctuations of the smFRET signal would report on conformational dynamics between IF2 and L11 within the 70S IC, and loss of the fluorescence signal would contain information on the lifetime of (Cy3)-IF2 on the 70S IC prior to dissociation.

30S ICs were prepared which contained IF1, (Cy3)-IF2-GTP, biotin-mRNA with an AUG start codon, and fMet-tRNA^{fMet} (Section 5.3.1). It should be noted that for the experiments described in this chapter, 30S ICs contained all components of the canonical 30S IC with the exception of IF3. 30S ICs formed in the absence of IF3 (i.e. 30SIC$_{-IF3}$) exhibit full biochemical activity with regards to formation of a 70S IC that is competent for initiation dipeptide formation (see Section 2.4.3) and are thus reasonable substrates to use for studying 50S subunit joining and 70S IC assembly. Furthermore, 30SIC$_{-IF3}$ possesses the advantage of having higher stability than complexes formed in the presence of all initiation components including IF3 (i.e. 30SIC$_{+IF3}$). The regulatory effect of IF3 on

the assembly and conformational dynamics of the 70S IC was investigated in detail and will be discussed separately in Chapter 3.

In order to make measurements of 50S subunit joining, it was first necessary to establish conditions under which (Cy3)-IF2 is stably bound to the surface-immobilized 30S IC on a timescale long enough to allow rinsing of the flowcell, assembly of the stopped-flow apparatus, and other steps preceding smFRET imaging (~5-10 min). It is known that the presence of IF1 and GTP enhance IF2's affinity for the 30S subunit [28, 31-33], which suggested that (Cy3)-IF2 could be kept stably bound to surface-immobilized 30S ICs by including high concentrations of IF1 (0.9 μM) and GTP (1 mM) in all dilution, wash, and imaging buffers. Under these conditions, the 30S ICs could be diluted to ~100s of pM and surface immobilized to yield ~200-300 spatially separated fluorescence spots per field-of-view, corresponding to stably and specifically bound (Cy3)-IF2-GTP (Figure 2.7).

Following surface-immobilization of 30S ICs containing (Cy3)-IF2-GTP, (Cy5)-L11 labeled 50S subunits were stopped-flow delivered into the flowcell. The resulting smFRET trajectories exhibit an initial dwell with an E_{FRET} ratio of zero, corresponding to the waiting time before 50S subunit joining, followed by a sharp transition to non-zero E_{FRET} ratios, which indicates the subunit joining event. Similar data was obtained with all three of the (Cy3)-IF2 constructs (Figure 2.8). Thus, surface-immobilized 30S ICs are capable of participating in IF2-catalyzed 50S subunit joining. The smFRET versus time trajectories obtained from these experiments contain a wealth of mechanistic information on the rate of subunit joining, conformational dynamics between IF2 and the GAC during

70S IC formation, and the lifetime of IF2 on the 70S IC, as described in the following sections.

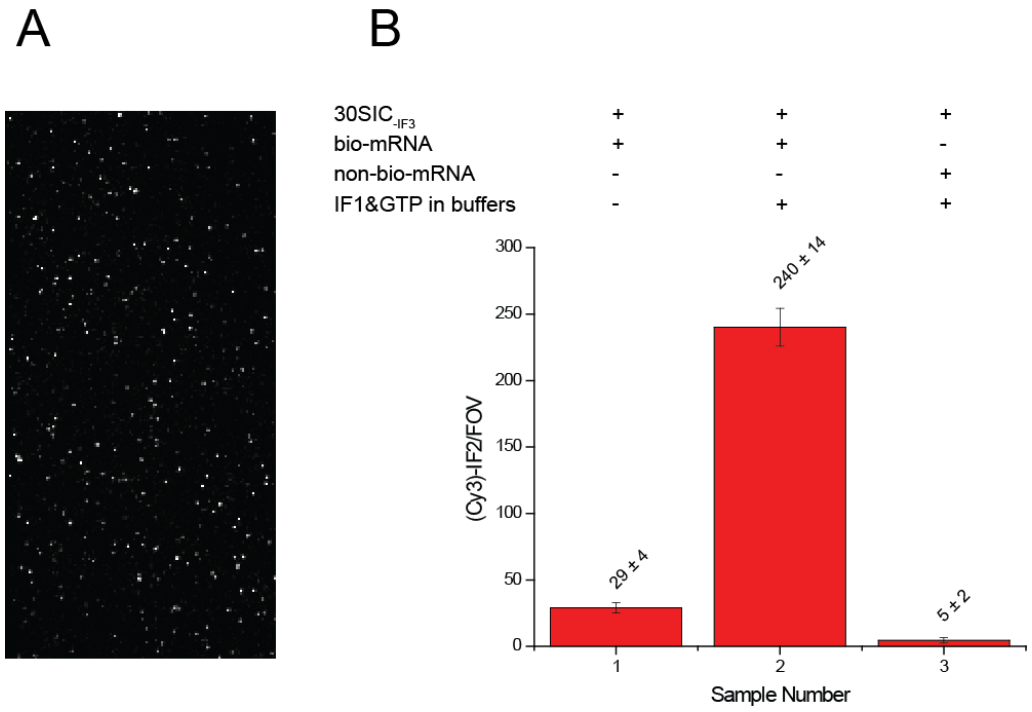

Figure 2.7: Stability of (Cy3)-IF2-GTP binding to 30SIC$_{-IF3}$ and biotin specificity of surface immobilization.
(A) Sample field-of-view (FOV) depicting several hundred fluorescent spots originating from individual surface-bound, (Cy3)-IF2–containing 30S ICs. (B) The number of (Cy3)-IF2 molecules per FOV was quantified under various conditions. Following surface immobilization of 30SIC$_{-IF3}$ and rinsing of the flowcell, ten separate FOVs were imaged. Fluorescent spots were identified by applying an intensity threshold and selecting regions containing at least two contiguous pixels above the threshold. The average and standard deviation for the number of (Cy3)-IF2 spots per FOV is depicted in the bar graph. 30SIC$_{-IF3}$ complexes contained fMet-tRNA^{fMet}, IF1, (Cy3)-IF2, GTP, and either biotin- or non-biotin mRNA as indicated. Inclusion of IF1 (0.9 μM) and GTP (1 mM) in the dilution, wash, and imaging buffers caused an ~8-fold increase in the number of surface-bound (Cy3)-IF2 molecules observed. Only background levels of fluorescence were observed for 30SIC$_{-IF3}$ formed with non-biotin mRNA. This suggests that ~98% of the (Cy3)-IF2 spots correspond to (Cy3)-IF2 that is bound to the 30S IC, which in turn is specifically tethered to the flowcell surface via the biotin-streptavidin interaction.

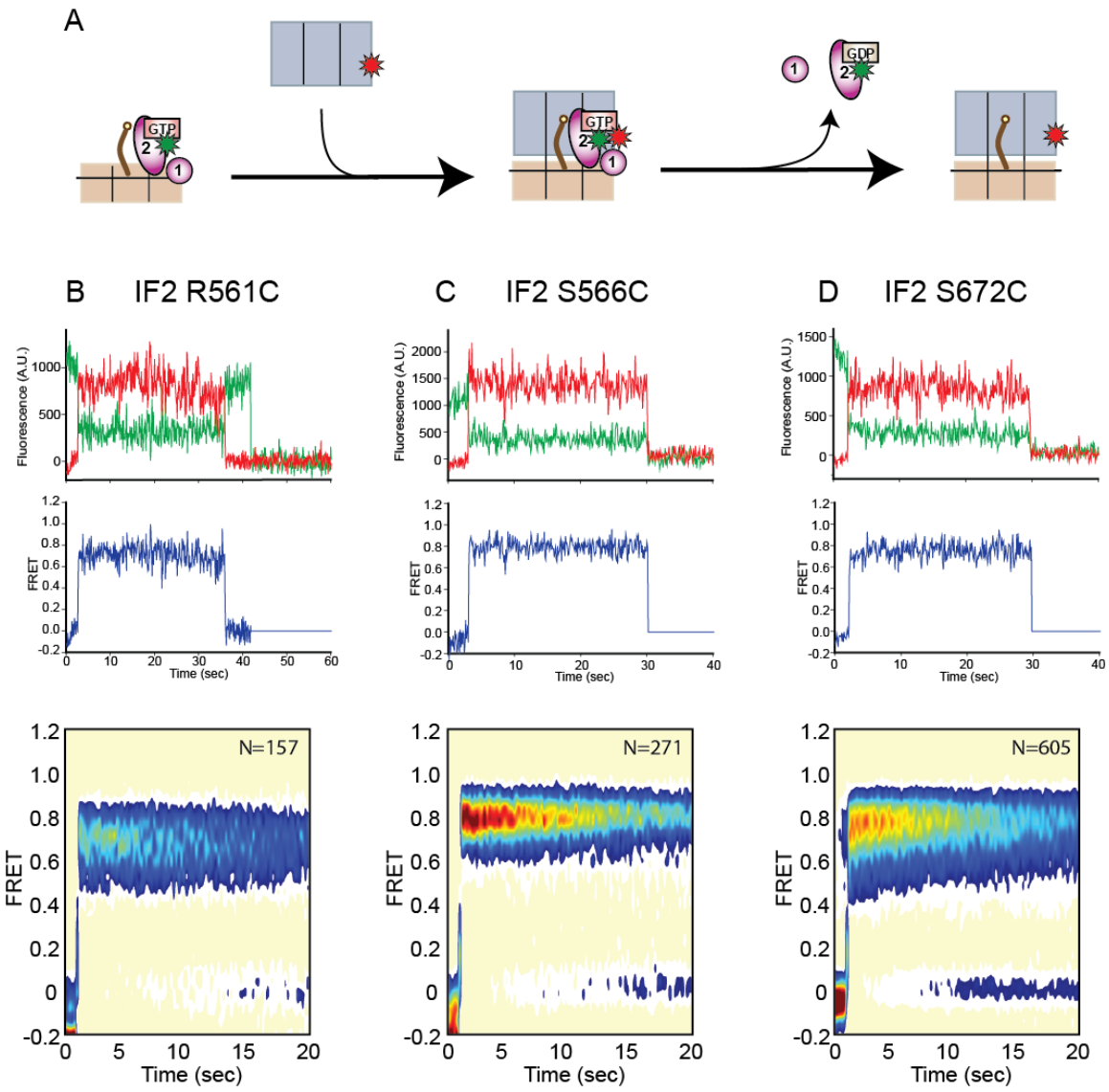

Figure 2.8: Real-time smFRET measurements of IF2-catalyzed 50S subunit joining.
(A) Cartoon depicting the stopped-flow delivery of (Cy5)-L11 labeled 50S subunits to 30S ICs carrying (Cy3)-IF2-GTP, IF1, mRNA, and fMet-tRNA^{fMet}. Three different IF2 constructs were used, with data for (Cy3)-IF2 R561C, (Cy3)-IF2 S566C, and (Cy3)-IF2 S672C shown in panels (B), (C), and (D), respectively. Second row: Representative Cy3 (green) and Cy5 (red) fluorescence intensity versus time trajectories. Third row: The corresponding smFRET versus time trajectories, where FRET is calculated as $I_{Cy5}/(I_{Cy3} + I_{Cy5})$. Fourth row: Post-synchronized time evolution of population FRET histogram, made by superimposing individual smFRET trajectories after synchronizing the first FRET event > 0.2 to time = 1 sec. The number of trajectories used to construct each contour plot is indicated by "N." Contours are plotted from tan (lowest population) to red (highest population).

2.6.2 Rate of subunit joining

50S subunit joining to 30S ICs has previously been investigated at the single-molecule level by Marshall, et al. using smFRET probes attached to the 30S and 50S ribosomal subunits [34, 35]. In these studies, fluorescently labeled DNA oligonucleotides were hybridized to rRNA hairpin extensions that had been engineered into helix 44 on the 30S subunit and helix 101 on the 50S subunit, thus generating a (Cy3)-h44/(Cy5)-H101 intersubunit FRET signal. Stopped-flow delivery of (Cy5)-H101 50S subunits to surface-tethered (Cy3)-h44 30S ICs led to onset of FRET upon 50S subunit docking. Since my 50S subunit joining experiments were conducted under similar Tris-polymix buffer conditions and at the same concentration of Cy5-labeled 50S subunits (20 nM) used by Marshall, et al., a comparison of their measurements on the apparent rate of subunit joining with my measurements using the (Cy3)-IF2/(Cy5)-L11 smFRET signal could provide insight into the mechanism of the IF2-catalyzed reaction.

Subunit joining times were defined as the FRET arrival time corrected for the dead-time of our stopped-flow apparatus (see Section 5.5.1.3), and were measured for hundreds of individual 30S ICs under a range of magnesium ion concentrations (3.5 to 15 mM Mg^{2+}, Figure 2.9 and Table 2.1). The apparent first order rate of subunit joining, calculated as the inverse of the mean FRET arrival time, was ~0.4 sec^{-1} at 20 nM 50S subunit concentration and room temperature, and does not exhibit a strong magnesium dependence. This corresponds to a bimolecular rate of ~20 $\mu M^{-1}sec^{-1}$, which falls within the range of 12 to 120 $\mu M^{-1}sec^{-1}$ reported based on ensemble light scattering measurements of 50S subunit docking to 30S subunits in the presence of IF1, IF2,

mRNA, fMet-tRNA^{fMet}, and GTP [3, 4]. The fact that the rate measured here lies near the lower end of the range of ensemble values is likely due to steric and surface effects associated with tethering of the 30S ICs [36]. Subunit joining times were also measured for 30SIC$_{+IF3}$ complexes (see Chapter 3 and Appendix B). The mean subunit joining time was found to decrease as a function of increasing (Cy5)-50S concentrations, consistent with a bimolecular association reaction.

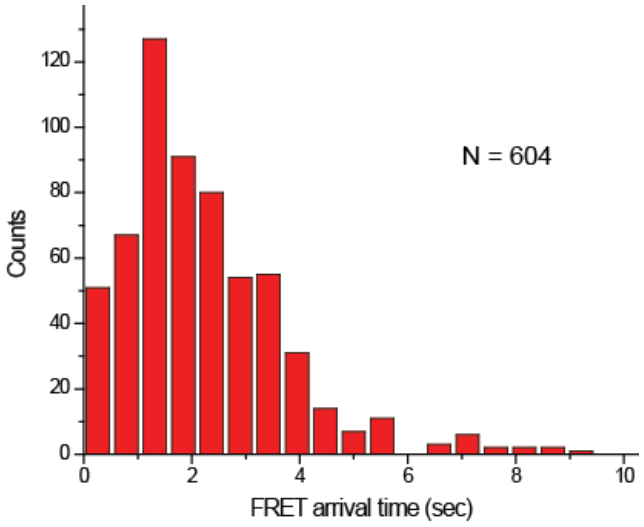

Figure 2.9: Histogram of subunit joining times.
Subunit joining times were determined for experiments in which (Cy5)-L11 reconstituted 50S subunits were stopped-flow delivered to surface-immobilized 30S ICs containing (Cy3)-IF2 S672C-GTP in Low-Salt Tris-polymix buffer with 15 mM Mg^{2+}. FRET arrival times were calculated as the time of the first data point > 0.2 FRET, minus the estimated dead time of our stopped-flow instrument (~2.0 sec). Data from 604 subunit joining events was distributed among 20 equally spaced bins spanning the range of 0 to 10 sec.

There are two notable differences between the data presented in Table 2.1 and that reported by Marshall, et al. using the (Cy3)-h44/(Cy5)-H101 smFRET signal [35]. First, the mean FRET arrival times reported here are generally faster, for example, ~3.5-fold

faster at 5 mM Mg^{2+} using similarly prepared 30S ICs. Second, whereas Marshall, et al. report a ~3-fold acceleration of subunit joining at 15 mM versus 5 mM Mg^{2+}, my results indicated little difference in the apparent rate of subunit joining over an even wider concentration range (3.5 mM to 15 mM Mg^{2+}). These differences may reflect the fact that all subunit joining events detected using my (Cy3)-IF2/(Cy5)-L11 smFRET signal arise from surface-immobilized 30S ICs that already contain IF2, which, once bound to the 30S subunit, can rapidly promote subunit joining. Subunit association, however, does not explicitly require IF2; it does occur, albeit more slowly, in the absence of initiation factors [37]. Therefore, it seems likely that a non-negligible subset of the subunit joining events observed by Marshall et al. correspond to 50S subunits that associate with IF2-free 30S ICs. These 30S ICs would exhibit a slower rate of subunit joining, thereby shifting the mean FRET arrival time for the entire population towards higher values. In support of this interpretation, 50S subunit joining events were observed with the (Cy3)-h44/(Cy5)-H101 signal even in the absence of initiation factors, but the association rate was slowed ~3-fold [35].

This interpretation would suggest that, under certain conditions, the binding of IF2 to the 30S IC may limit the apparent rate of 50S subunit joining. Indeed, there is evidence from our laboratory that tuning IF2's binding kinetics to the 30S IC may constitute a regulatory mechanism for modulating the rate of subunit joining. IF2's binding affinity was found to be dramatically weakened on incorrectly assembled 30S ICs, which may inhibit the 50S subunit joining reaction and thus discourage assembly of aberrant 70S ICs [16]. The (Cy3)-IF2/(Cy5)-L11 smFRET signal reported here, however,

requires that (Cy3)-IF2 be bound to the 30S IC at the onset of data collection, and consequently, is not sensitive to this level of regulation.

Table 2.1: Subunit joining times at different Mg^{2+} ion concentrations.
Mean FRET arrival times were calculated at four different Mg^{2+} concentrations (3.5, 5, 10, and 15 mM Mg^{2+}) as described in the text. Data were compiled from at least two independent experiments. Errors were estimated by splitting the data into three equal parts and calculating the average and standard deviation.

[Mg^{2+}] (mM)	Number of Molecules	Mean FRET Arrival Time (sec)
3.5	245	2.1 ± 0.2
5	258	3.9 ± 0.6
10	320	3.5 ± 0.2
15	604	2.0 ± 0.2

The higher degree of Mg^{2+}-dependence observed by Marshall et al. on the rate of subunit joining can be explained in a similar light. Mg^{2+} ions act to shield the negative charge of rRNA phosphate groups that come into close contact at the interface between 30S and 50S ribosomal subunits, thereby reducing the energetic barrier for subunit association [37, 38]. Even in the absence of initiation factors, the equilibrium between 70S complexes and free subunits can be shifted towards full association at high (>10 mM) concentrations of Mg^{2+} [39]. As noted above, it is possible that a significant fraction of the subunit joining events observed using the (Cy3)-h44/(Cy5)-H101 signal occur when a 50S subunit docks to a 30S IC that does not contain IF2. It may be that these slower, uncatalyzed docking events are more sensitive to Mg^{2+} concentration than the IF2-catalyzed events, and that the observed ~3-fold acceleration of subunit joining at high Mg^{2+} observed by Marshall, et al. largely reflects the effect of Mg^{2+} on this

subpopulation of complexes. When bound to the 30S IC, IF2 likely reduces the activation energy for subunit joining to such an extent that it can occur rapidly regardless of the specific Mg^{2+} concentration.

2.6.3 Lifetime of IF2 on the 70S IC

The smFRET versus time trajectories shown in Figure 2.8 contain information about the lifetime of IF2 on the 70S IC following 50S subunit joining. Dissociation of (Cy3)-IF2 from surface-immobilized ribosomes should lead to loss of spatially localized Cy3 fluorescence and termination of the smFRET signal. Interpretation of loss of Cy3 fluorescence as corresponding to the dissociation of (Cy3)-IF2, however, is complicated by the fact that signal loss may also occur as a result of fluorophore photobleaching. Nevertheless, it was evident from the individual smFRET trajectories that, on average, IF2 remains bound to the 70S IC for a longer period of time (~10s of sec) than anticipated. Based on ensemble biochemical measurements, GTP hydrolysis by IF2 occurs rapidly following 50S subunit joining, with an apparent rate of 30 ± 5 sec^{-1}. This is followed by a slower release of inorganic phosphate (P_i), which occurs with a rate of 1.5 ± 0.5 sec^{-1} after a lag phase of 200 msec [6]. These biochemical events are thought to result in a conformational change of IF2 from its GTP-bound form to its GDP-bound form, initiated by changes in the Switch 2 loop of IF2's G-domain, and propagated via coupled domain movements to affect a large-scale rearrangement of the entire IF2 structure [40]. The extended residency of IF2 on the 70S IC observed here suggests that even after GTP hydrolysis, P_i release, and IF2's conformational change, IF2-GDP

maintains a sufficient number of contacts with the 30S and 50S ribosomal subunits, and possibly fMet-tRNA^{fMet}, which slow its passive dissociation from the ribosome.

To better characterize the lifetime of IF2 on the 70S IC, it was necessary to limit the effects of photobleaching on the observed lifetime of the Cy3 signal. This was achieved by shuttering the laser excitation source at regular intervals, which allows extended observation times at the expense of reduced time resolution. The strategy that was employed involved collecting data frames (100 msec exposure) continuously under constant laser excitation for five seconds (i.e. 50 frames) at the beginning of the experiment before starting the shuttering routine, at which point single, 100 msec exposure data frames were collected at regular intervals with the laser light blocked in between. Continuous data collection at the beginning of the experiment was necessary to resolve the rapid subunit joining event, which occurs within the first five seconds for the vast majority (92%) of (Cy3)-IF2–bound 30S ICs (Figure 2.9). Additionally, it minimizes the risk of failing to observe fast-dissociating (Cy3)-IF2 molecules, in case this subset of molecules were to become significantly populated under certain experimental conditions.

The data show that, as the time interval between data frames is increased, the observed lifetime of the (Cy3)-IF2 fluorescence signal following subunit joining initially increases, confirming the hypothesis that the original measurements were limited by photobleaching (Figure 2.10 and Table 2.2). However, as the time interval is further increased past ~1 sec, the measured lifetime begins to plateau at values of ~100 sec, indicating that, after subunit joining, IF2 likely remains bound to the ribosome on the minutes timescale prior to passive dissociation. As a control, the same experiment was

performed using 30S ICs containing (Cy3)-IF2-GDPNP (with a time delay between data frames of 4 sec). GTP hydrolysis is a prerequisite for dissociation of IF2 from the ribosome [28, 41]; by preventing this biochemical step using a non-hydrolyzable GTP analog, the observed lifetime of (Cy3)-IF2 was extended by an order of magnitude to ~1200 sec. This measurement is likely limited by photobleaching and thus probably represents a lower limit to the actual residency time of IF2-GDPNP on the ribosome following subunit joining.

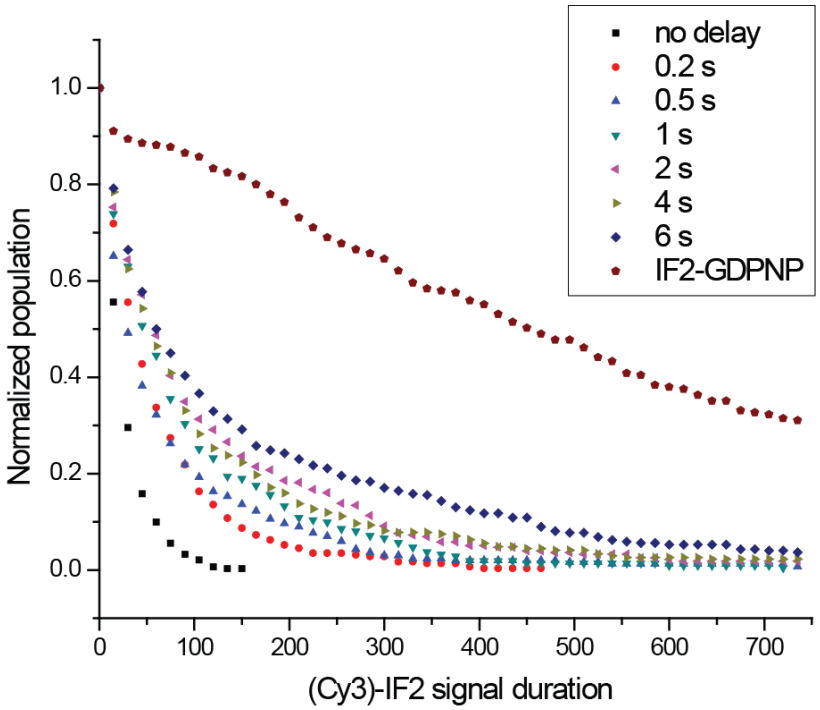

Figure 2.10: Lifetime of (Cy3)-IF2 on the 70S IC following 50S subunit joining.
(Cy5)-L11 labeled 50S subunits were stopped-flow delivered to surface-immobilized 30S ICs containing (Cy3)-IF2. All experiments were performed in the presence of GTP, except as indicated in the legend. 100 msec exposures were collected under continuous laser excitation for 5 sec, followed by 100 msec exposures separated by time delays during which the laser light was shuttered. The time delay between exposures used for each experiment is indicated in the legend. Molecules with an observed FRET event > 0.2 occurring within the first 5 sec of data acquisition were selected for analysis. The time from the FRET event to the loss of spatially localized Cy3 fluorescence signal was calculated for each molecule and data were plotted as a normalized population decay histogram (bin size = 15 sec). Data for each shuttering condition were compiled from three independent experiments comprising hundreds of single molecules.

In summary, my results confirm that GTP hydrolysis by IF2 is required for its dissociation from the 70S IC, but suggest that even in the presence of GTP hydrolysis, passive dissociation is quite slow. This is consistent with the results from a cryo-EM reconstruction of *Thermus thermophilus* IF2 bound to the 70S ribosome in the post-GTP hydrolysis state, in which stable binding of IF2-GDP was a necessary requirement for data collection [11]. A comparison of the reconstruction with IF2-GDP to one with IF2-GMPPCP (mimicking the pre-GTP hydrolysis state) revealed conformational rearrangements of IF2 following GTP hydrolysis which involved a ~20° rotation of IF2 relative to the ribosome and a ~10 Å shift outwards from the intersubunit space. These movements break contacts between IF2's G domain and 30S subunit rRNA helices h8 and h14, between domain VI-1 and h5 and h15, and between domain VI-2 and h44 and fMet-tRNA^{fMet} [11]. Apparently, despite the disruption of IF2-ribosome and IF2-tRNA contacts during the transition of IF2 from its GTP- to GDP-bound conformation, the remaining contacts between IF2-GDP and the ribosome are sufficient to keep it bound for an extended period of time (~100 sec), at least at room temperature. In the future, it may be interesting to study the dissociation time of IF2 as a function of temperature in order to estimate the free energy barrier to IF2 release following GTP hydrolysis.

Based on these results, I speculated that the next event in the translation initiation pathway, binding of EF-Tu:GTP:aa-tRNA ternary complex to the ribosome, might accelerate IF2 dissociation. To test this possibility, ternary complexes were prepared containing Phe-tRNA^{Phe}, which is encoded by the second, UUC codon of the T4gp32 mRNA. Phe-tRNA^{Phe} ternary complex was then co-delivered with (Cy5)-50S subunits to

surface-immobilized 30S ICs containing (Cy3)-IF2-GTP, and the loss of (Cy3)-IF2 fluorescence after subunit joining was followed as before. In the presence of ternary complex, the measured (Cy3)-IF2 residency time decreased only modestly, by ~1.4-fold at 1 μM ternary complex (Table 2.2). These data suggest that ternary complex can in fact bind to a 70S IC containing IF2-GDP and promote its dissociation, but that this effect is small. There are several potential reasons for this, including slow and/or inefficient binding of ternary complex to 70S ICs assembled on the flowcell surface. Nevertheless, these data imply that simultaneous binding of IF2 and ternary complex can occur, a notion that is somewhat surprising considering their partially overlapping binding sites on the ribosome [12, 42]. I sought to test this idea by fluorescently labeling the ternary complex so as to directly monitor its binding to the 70S IC. These experiments, which are ongoing, aim to explore the relative timing of ternary complex binding with respect to IF2 dissociation, and will be discussed later in Chapter 4.

While my data suggest that ternary complex might be able to bind to the 70S IC prior to IF2 dissociation, the position of IF2 observed in the cryo-EM reconstructions [11, 12] certainly precludes final accommodation of aa-tRNA into the ribosomal A site and participation in peptide bond formation at the peptidyl transferase center. The observation of a delayed release of IF2 from the ribosome following 50S subunit joining is thus consistent with the idea that dissociation of initiation factors from the 70S IC may limit the rate of initiation dipeptide formation [6, 28].

Table 2.2: Lifetime of the (Cy3)-IF2 fluorescence signal following 50S subunit joining.

Complex	Time delay (sec)[a]	[Ternary complex] (μM)[b]	(Cy3)-IF2 fluorescence signal lifetime (sec)[c]
30S IC with (Cy3)-IF2-GTP	None	-	25
	0.2	-	56
	0.5	-	58
	1	-	83
	2	-	102
	4	-	87
	6	-	115
	4	0.25	73
	4	0.5	68
	4	1	61
30S IC with (Cy3)-IF2-GDPNP	4	-	1186

(a) Experiments began with 50 frames of continuous data acquisition (100 msec exposure), followed by a shuttering routine in which 100 msec exposures were separated by time delays of varying length during which the laser light was blocked.

(b) When included, preformed ternary complex consisting of EF-Tu, GTP, and Phe-tRNA^{Phe} was co-delivered with (Cy5)-labeled 50S subunits to surface-immobilized 30S ICs.

(c) For all 30S ICs that show a FRET event >0.2 within the first 5 sec of data acquisition, the time from the onset of FRET to loss of the (Cy3)-IF2 fluorescence signal was calculated, and these values were plotted as a normalized population decay histogram (see Figure 2.10). These curves were fit with a single exponential decay of the form $y=A_1*\exp(-x/t_1) + y_0$, and the average lifetime of the decay, t_1, is reported here.

2.6.4 Conformational dynamics within the 70S IC

Following subunit joining to 30S ICs containing IF2-GTP (referred to in this section as $30SIC_{GTP}$), fluctuations were observed between at least two discrete non-zero FRET states in a subpopulation of the time trajectories (Figure 2.11). Fluctuations between different non-zero FRET states were interpreted as corresponding to transitions between different conformational states of the 70S IC encompassing different distances between IF2's domain VI-1 and r-protein L11. The dynamic transitioning between

different FRET states within the 70S IC indicates that conformational rearrangements of IF2 with respect to L11 occur following 50S subunit joining to $30SIC_{GTP}$ and prior to release of IF2. In other words, the intermolecular interactions between IF2 and the GAC are dynamically remodeled during IF2's residency on the 70S IC.

This interpretation is in accord with results from a bulk FRET study in which *Bacillus stearothermophilus* IF2, labeled at amino acid position 378 with Cy3, was used in conjunction with an *E. coli* translation system that included 50S subunits reconstituted with L11, labeled at position 38 with Cy5. The authors observe an increase in E_{FRET} following 70S IC formation, which they conclude results from a relative movement of the L11 NTD towards the G domain of IF2 following GTP hydrolysis and preceding P_i release [43]. My single-molecule data also support the idea that a conformational rearrangement of the GAC occurs with respect to IF2; furthermore, they provide evidence that this rearrangement is reversible during the residency of IF2 on the 70S IC.

Since my one-dimensional smFRET signal reports only on the relative distance change between IF2 and L11, it is not possible, based on the present data, to unambiguously determine whether movements of IF2, the GAC, or both, are involved. However, based on the known conformational flexibility of the L11 region [44, 45], it seems likely that the FRET fluctuations arise at least in part from dynamics of the GAC. This possibility receives support from the observation that the conformation of the L11 arm is altered in response to binding of other translation factors to the ribosome: an "inward curling" of the L11 arm has been observed structurally, which helps to facilitate contacts of L11 with the elbow region of aa-tRNA during decoding at the A site [42],

with EF-G's G' domain during translocation [46], and with domain 1 of the Class 1 release factors RF1 and RF2 during translation termination [47]. To more clearly define the conformational changes occurring between IF2 and the GAC within the 70S IC, it will be necessary to devise and implement additional labeling schemes to probe the transitions from different structural perspectives. This approach could additionally prove useful in exploring if and how the conformational fluctuations between IF2 and L11 are related to other important conformational events occurring during 70S IC formation, such as adjustment of fMet-tRNA^{fMet} from the P/I hybrid configuration into the P/P configuration [12, 26].

In an attempt to gain insight into the nature and mechanistic role of the FRET fluctuations, I performed experiments monitoring subunit joining to 30SIC$_{GDPNP}$, in which the guanine nucleotide bound to IF2 was changed from GTP to GDPNP. I hypothesized that the conformational fluctuations between IF2 and L11 might be dependent on the GTP hydrolysis event and/or on the nucleotide bound to IF2. The results from these experiments revealed that, even in the presence of GDPNP, similar fluctuations between non-zero FRET states could be observed within individual 70S ICs (Figure 2.11). This indicates that the FRET fluctuations do not require GTP hydrolysis to be activated. Despite this, I suspected that there may be subtle differences in the dynamic behavior of 70SIC$_{GTP}$ and 70SIC$_{GDPNP}$ and thus sought to characterize the FRET fluctuations more quantitatively.

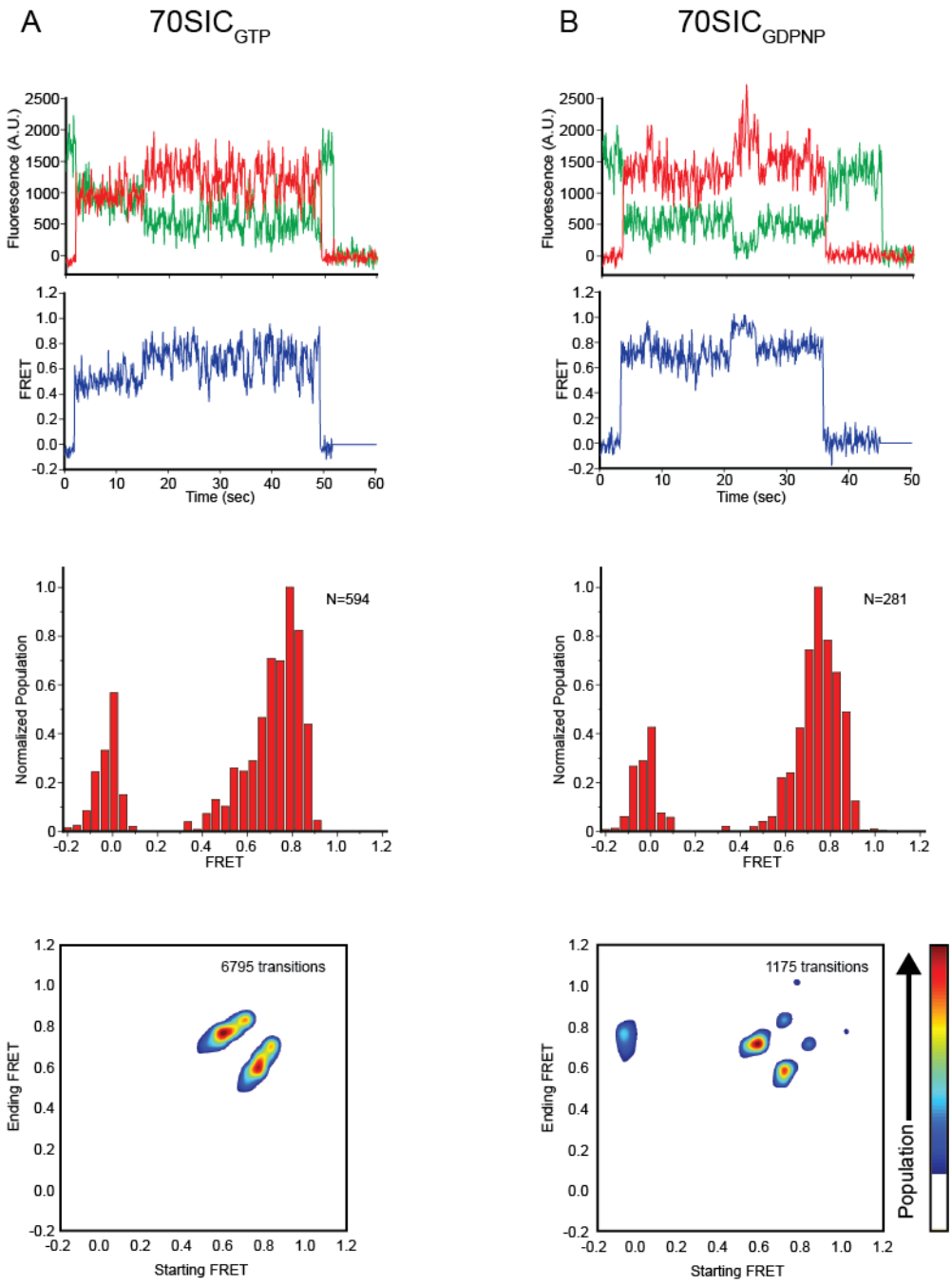

Figure 2.11: Conformational fluctuations within 70SIC$_{GTP}$ and 70SIC$_{GDPNP}$
Top Panel: Sample Cy3 (green) and Cy5 (red) fluorescence intensity versus time trajectories and the corresponding FRET versus time trajectory (blue) for molecules exhibiting discrete transitions between at least two non-zero FRET states. FRET is calculated as $I_{Cy5}/(I_{Cy3}+I_{Cy5})$. Middle Panel: One-dimensional histogram of the population's FRET distribution. Following idealization of smFRET trajectories with vbFRET, the histogram was constructed using all data points prior to Cy3 photobleaching for all trajectories in the dataset. Data was distributed among 35 equally

spaced bins spanning the range of -0.2 to 1.2 FRET and normalized to the most populated bin. The number of idealized trajectories used to construct the histograms is indicated by "N." Bottom Panel: Transition density plots (TDPs) were generated by plotting the "Starting FRET" versus "Ending FRET" for transitions within the idealized smFRET versus time trajectories as a contour plot representation of a two-dimensional histogram. The lower- and upper-bound thresholds for plotting the data were 30% and 100% of the most populated bin, respectively, which were chosen so as to highlight the most frequently observed transitions. Data are plotted from white (lowest population) to red (highest population) as indicated by the color bar. The number of transitions used to construct the TDP is indicated in the upper right corner.

In order to compare conformational dynamics within $70SIC_{GTP}$ and $70SIC_{GDPNP}$, the individual smFRET versus time trajectories were first idealized to a hidden Markov model using the vbFRET software package [48]. This software uses a maximum-evidence based algorithm to calculate the most probable number of FRET states, and the path through these FRET states, for each time trajectory. The analysis settings used for the fitting procedure instructed the software to fit a minimum of one state and a maximum of five states to the data, with 25 fitting attempts per trajectory. Using this approach, the majority (57%) of smFRET versus time trajectories for $70SIC_{GDPNP}$ were best fit with a two-state model, while 39% were fit with a three-state model and 4% were fit with a four-state model. In all cases, at least one of the states corresponds to the 30S IC prior to docking of the 50S subunit (i.e. a zero-FRET state). Therefore, these results suggested that, following subunit joining, the majority of $70SIC_{GDPNP}$ complexes sample either one or two non-zero FRET states during the observation time. Similar results were obtained from vbFRET modeling of smFRET trajectories corresponding to formation of $70SIC_{GTP}$, in which 44% of trajectories were fit with a two-state model, 46% with a three-state model, and 9% with a four-state model.

The idealized smFRET trajectories generated by vbFRET allow construction of transition density plots (TDPs) [49], which provide a graphical representation of the frequency of transitions between different FRET states (Figure 2.11). The most prominent peaks on the TDP for $70SIC_{GTP}$ are those corresponding to transitions between E_{FRET} ratios of ~0.6 and ~0.8. For $70SIC_{GDPNP}$, the most prominent peaks correspond to transitions between E_{FRET} ratios of ~0.7 and ~0.8. Since the non-zero FRET fluctuations generally occur between such closely spaced FRET states, although discrete transitions are clearly identifiable in individual smFRET versus time trajectories, individual states are not separable in the one-dimensional histogram of E_{FRET} ratios (Figure 2.11). For example, the distribution of non-zero FRET values for $70SIC_{GTP}$ is asymmetrical, with a large peak at ~0.8 FRET and a shoulder extending toward ~0.5-0.6 FRET, which likely corresponds to two FRET states centered at E_{FRET} ratios of ~0.8 and ~0.6 that have broad, overlapping distributions. Assuming $R_0 \approx 60$ Å for this Cy3/Cy5 labeling scheme, transitions between states centered at E_{FRET} ratios of ~0.8 and ~0.6 would correspond to distance changes of ~9 Å.

It is possible that separation between FRET states could be improved by using different fluorophore labeling positions on IF2 and/or L11 that report on the same dynamics but exhibit a larger change in distance as a result of the conformational changes. Similarly, alternative labeling schemes might be found that reduce the width of the FRET distribution for one or more states, thus leading to better peak definition. For example, it is possible that the flexibility of the L11 NTD gives rise to conformational noise that broadens the FRET peaks, and that better peak resolution would be obtained by

moving the Cy5 fluorophore to L11's CTD. Regardless, with the current labeling scheme, the fact that discrete states are not resolved in the histogram of E_{FRET} ratios complicates detailed kinetic analysis of transition rates from one well-defined state to another. Instead, the overall dynamic behavior of $70SIC_{GTP}$ and $70SIC_{GDPNP}$ was compared in a more general way, as described in the following paragraph.

Trajectories for which vbFRET identified transitions between at least two non-zero FRET states with $\Delta E_{FRET} \geq 0.05$ were selected so that they could be analyzed separately (see Section 5.5.1.8). For $70SIC_{GDPNP}$, 29% of the trajectories were found to undergo transitions between at least two non-zero FRET states. The remaining 71% of the trajectories for $70SIC_{GDPNP}$, therefore, sample only one non-zero FRET state prior to loss of the fluorescence signal. For $70SIC_{GTP}$, 43% of the trajectories were found to exhibit transitions between at least two non-zero FRET states, with the remaining 57% of trajectories sampling only one non-zero FRET state prior to signal loss. Thus, the presence of GDPNP in place of GTP led to a slight decrease in the subpopulation of ribosomes that exhibited conformational fluctuations within the 70S IC during the experimental observation window. Furthermore, within this fluctuating subpopulation, the average transition rate between non-zero FRET states (defined as the total number of transitions divided by the total time spent in non-zero FRET states, see Section 5.5.1.8) was more than two-fold slower in the presence of GDPNP (0.06 transitions sec^{-1}) versus GTP (0.15 transitions sec^{-1}).

Taken together, the observations that FRET fluctuations within $70SIC_{GDPNP}$ occur more slowly and in a smaller fraction of the molecules (within the experimental

observation window) compared with $70SIC_{GTP}$ suggest that the presence of GDPNP may dampen dynamic behavior of the 70S IC and reduce the probability that fluctuations will be observed. In other words, IF2-dependent GTP hydrolysis somehow increases the probability of activating conformational fluctuations within the 70S IC. GTP hydrolysis by IF2 occurs rapidly upon subunit joining and is followed shortly thereafter by P_i release [6]; thus, IF2 is expected to be present in the GDP-bound form for the majority of the experimental observation window. This would suggest that the transition of IF2 from the GTP- to GDP-bound form is accompanied by an increase in conformational dynamics of the 70S IC.

IF2-GDP adopts an alternate conformation within the 70S IC corresponding to a different network of IF2-ribosome interactions [11]. Based on my smFRET data, I propose that one of the structural changes that occurs after GTP hydrolysis and P_i release involves a reconfiguration of L11-IF2 intermolecular contacts. The observed increase of conformational fluctuations in the presence of GTP hydrolysis (i.e. within $70SIC_{GTP}$) may arise from fewer and/or weaker contacts formed between L11 and IF2-GDP in the post-hydrolysis state. This could provide a mechanistic explanation for why IF2 dissociates more quickly from the ribosome in the presence of GTP versus GDPNP (Figure 2.10) and would suggest that one function of L11 is to regulate the stability of IF2 on the ribosome. This model would predict that the rate of IF2 dissociation from the 70S IC should be accelerated, even in the absence of GTP hydrolysis, by weakening or disrupting the interactions between IF2 and L11. Future experiments to test this hypothesis will make use of L11-binding thiazole antibiotics such as thiostrepton, as well as an L11-NTD

truncation mutant, both of which should lead to altered and presumably weakened interactions between L11 and the GAC.

2.7 Interaction of IF2 with the GAC during multiple-turnover GTP hydrolysis

Even in the absence of fMet-tRNA^{fMet} and mRNA, IF2 can bind to the 70S ribosome and hydrolyze GTP catalytically [6]. In other words, IF2 can hydrolyze GTP in a multiple-turnover reaction that is uncoupled from the initiation pathway. Such uncoupled GTP hydrolysis is also observed for EF-G and EF-Tu in the presence of vacant 70S ribosomes [50]. I designed an experiment to characterize IF2's interactions with the GAC during multiple-turnover GTP hydrolysis, to see if and how they differ from the IF2-GAC interactions established during IF2-catalyzed subunit joining and 70S IC formation. To do so, (Cy5)-labeled IF2 and 50S subunits reconstituted with (Cy3)-L11 were prepared, thus reversing the fluorophore labeling scheme used for the 50S subunit joining assays. The (Cy3)-L11 reconstituted 50S subunits were used to enzymatically prepare 70S ICs containing biotin-mRNA and fMet-tRNA^{fMet} in the P site, which were purified by sucrose-density gradient ultracentrifugation prior to immobilization on the surface of the flowcell (Section 5.3.3). Following surface immobilization of (Cy3)-L11 labeled 70S ICs, (Cy5)-IF2 was introduced into the flowcell at nanomolar concentrations along with 1 mM GTP, and steady-state smFRET data was collected under continuous 532 nm laser illumination.

The smFRET versus time trajectories collected under these conditions revealed extended dwells in a zero-FRET state with intermittent and transient excursions to non-zero E_{FRET} ratios. Within a given smFRET trajectory, multiple such excursions could be observed (Figure 2.12). This data was initially interpreted as corresponding to successive binding and dissociation events of different (Cy5)-IF2 molecules from the same ribosomal complex, with the free state of the complex indicated by dwells with an E_{FRET} ratio of zero and the bound state indicated by dwells with non-zero E_{FRET} ratios.

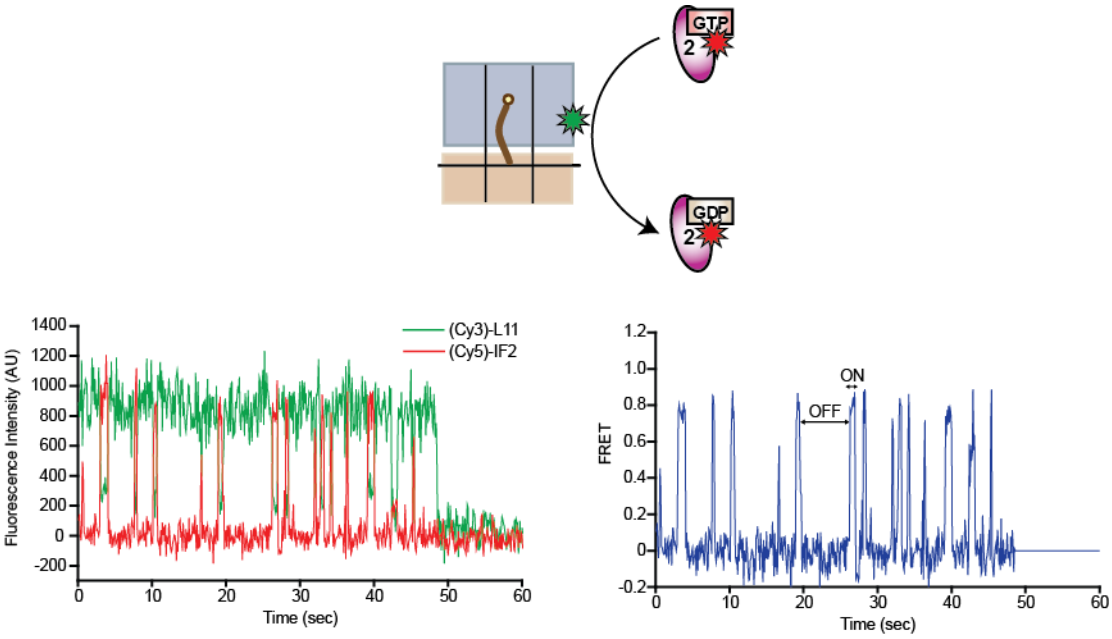

Figure 2.12: Repetitive binding and dissociation of (Cy5)-IF2 on pre-formed 70S ICs.
Top row: Cartoon depicting the experimental set-up used to monitor IF2's interaction with the GAC during multiple-turnover GTP hydrolysis. 70S ICs containing (Cy3)-L11, biotin-mRNA, and fMet-tRNA^{fMet} in the P site were assembled, purified, and tethered to the surface of the flowcell. (Cy5)-IF2 and GTP were introduced into the flowcell and data were collected under steady-state conditions. Bottom row: Sample (Cy3)-L11 and (Cy5)-IF2 fluorescence versus time trajectory (left) and the corresponding smFRET versus time trajectory (right). These data were collected with 10 nM free (Cy5)-IF2 present in the flowcell. Dwells in the zero-FRET state correspond to the unbound state of the 70S IC (OFF dwells), while dwells of non-zero FRET correspond to the (Cy5)-IF2 bound state (ON dwells). Multiple transitions between the bound and unbound state were observed for individual 70S ICs, indicative of repetitive binding and dissociation cycles of (Cy5)-IF2.

To verify this interpretation, a (Cy5)-IF2 titration was performed in which the free (Cy5)-IF2 concentration within the flowcell was varied from 2.5 to 40 nM (Figure 2.13). The data reveal that as the free concentration of (Cy5)-IF2 is increased, dwell times in the zero-FRET state decrease, as would be expected for a bimolecular binding reaction. Conversely, dwell times in the non-zero FRET state were concentration independent, as predicted for a unimolecular dissociation reaction. The smFRET data are thus consistent with repetitive binding and dissociation of (Cy5)-IF2 to the ribosome, with each cycle likely resulting in the hydrolysis of one molecule of GTP.

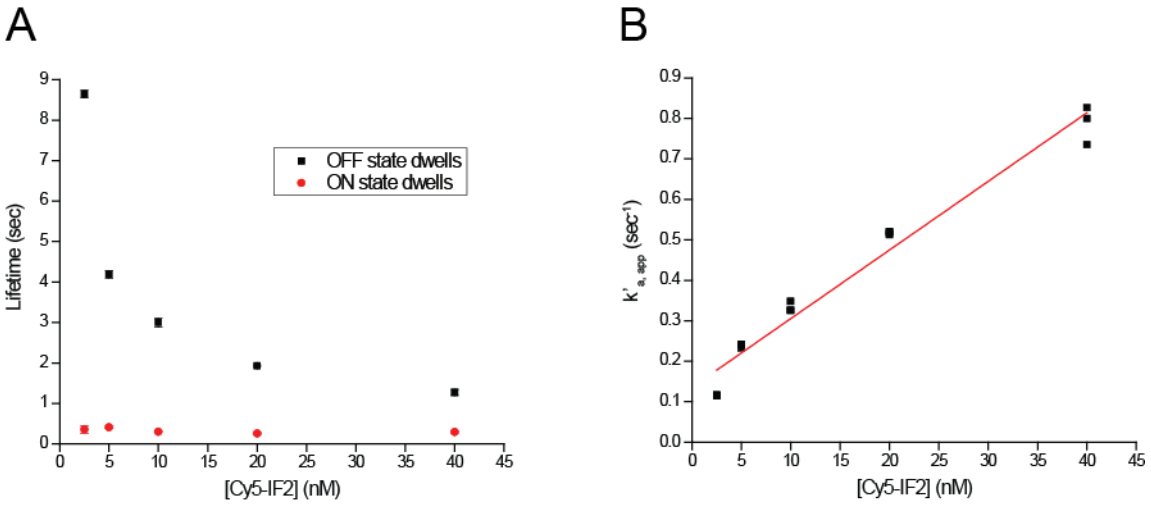

Figure 2.13: Dependence of ON- and OFF-state lifetimes on free (Cy5)-IF2 concentration. (A) A threshold of FRET=0.2 was used to define "OFF state dwells" and "ON state dwells" as corresponding to the amount of time spent in states with FRET≤0.2 and FRET>0.2, respectively. Dwell times in the ON and OFF states were extracted from idealized smFRET versus time trajectories, and population decay histograms were constructed from these dwell times. The curves were fit with single exponential decay of the form $y=A_1*\exp(-x/t_1) + y_0$, and the average lifetime of the decay, t_1, is reported here. Errors were estimated by randomly splitting all smFRET trajectories into three equally sized sets, analyzing each set independently, and taking the average and standard deviation of the resulting lifetimes. Error bars are plotted in the figure but are smaller than the symbol size. (B) Estimate of the second-order rate constant ($k_{a,\ app}$) for (Cy5)-IF2 binding to the pre-formed 70S IC. For each dataset, the pseudo-first order rate constant ($k'_{a,\ app}$) was calculated by taking the inverse of the calculated OFF-state lifetime. The resulting values were plotted as a function of free (Cy5)-IF2 concentration. The data were fit with a linear regression equation of the form y = mx + b, where m = 0.0170 ± 0.0009 $nM^{-1}sec^{-1}$ and b = 0.14 ± 0.02 sec^{-1}. The slope provides an estimate for $k_{a,app}$ of ~17 $\mu M^{-1}sec^{-1}$.

The observation of transient binding interactions between IF2 and the post-initiation ribosome during multiple-turnover GTP hydrolysis (300 msec lifetime) is in stark contrast to the very long-lived binding of IF2 observed following 50S subunit docking to the 30S IC during translation initiation (~100 sec lifetime, Section 2.6.3). This large difference in apparent binding affinities implies two quite different interaction modes between IF2 and the ribosome's GAC. During 50S subunit joining and formation of the 70S IC, IF2 serves as a giant intermolecular bridge that brings the two subunits together, and in so doing, becomes sandwiched between them, positioned within a cavity that extends from the GAC at the ribosome's surface all the way into the P site [12]. I speculate that during the short-lived binding events observed in the context of multiple-turnover GTP hydrolysis, IF2 does not insert into this cavity but instead only interacts with GAC components at the ribosome's surface. In this scenario, contacts between IF2's C-terminal region and fMet-tRNA^{fMet} and the 30S/50S subunit interface would not be formed, and binding would occur primarily through formation of contacts between the N-terminal region of IF2 (i.e. the G domain and domain V) and the GAC. This model would suggest that the C-terminal region of IF2 may not be explicitly required for ribosome binding and GTP hydrolysis. Indeed, the isolated 40 kDa G-domain fragment of IF2, obtained by limited proteolysis, was shown to be, by itself, capable of binding to the ribosome and catalyzing ribosome-dependent GTP hydrolysis [51]. Therefore, I hypothesize that a C-terminal truncation mutant of IF2 would show very similar binding behavior to the post-initiation ribosome as full-length IF2 in my smFRET assay.

These results demonstrating two different binding modes of IF2 with the ribosome raise several important points. First, they suggest that care should be taken in interpreting the results from experiments performed in the presence of IF2 and 70S ribosomes (for example, in references [21, 28]) as reporting on the mechanism of translation initiation per se. Second, they raise the question of whether the short-lived binding events of IF2 with 70S ribosomes occur *in vivo*, and if so, what their function might be. It is entirely possible that there is no *in vivo* function for this binding mode of IF2, and that its occurrence is simply an incidental byproduct of the high degree of structural similarity between domains IV and V of IF2 and the corresponding domains of EF-G and EF-Tu [40], the latter which bind to the intact 70S ribosome as part of their normal catalytic cycle. If this were the case, however, it would imply that either IF2 undergoes energetically costly cycles of ribosome binding and futile GTP hydrolysis, or alternatively, that a cellular mechanism exists to prevent IF2 binding to the elongating 70S ribosome and competing with EF-Tu and EF-G.

2.8 References

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

2. 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.

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

4. 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.

5. 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.

6. 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.

7. 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.

8. 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.

9. Murphy, M.C., I. Rasnik, W. Cheng, T.M. Lohman, and T. Ha, *Probing single-stranded DNA conformational flexibility using fluorescence spectroscopy.* Biophys J, 2004. **86**(4): p. 2530-7.

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

11. 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.

12. 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.

13. 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.

14. Van Dyke, N., W. Xu, and E.J. Murgola, *Limitation of ribosomal protein L11 availability in vivo affects translation termination.* J Mol Biol, 2002. **319**(2): p. 329-39.

15. Bowen, W.S., N. Van Dyke, E.J. Murgola, J.S. Lodmell, and W.E. Hill, *Interaction of thiostrepton and elongation factor-G with the ribosomal protein L11-binding domain.* J Biol Chem, 2005. **280**(4): p. 2934-43.

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

17. 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.

18. 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.

19. 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.

20. 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.

21. Brandi, L., S. Marzi, A. Fabbretti, C. Fleischer, W.E. Hill, C.O. Gualerzi, and J. Stephen Lodmell, *The translation initiation functions of IF2: targets for thiostrepton inhibition.* J Mol Biol, 2004. **335**(4): p. 881-94.

22. 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.

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

24. 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.

25. Traut, R.R. and R.E. Monro, *The Puromycin Reaction and Its Relation to Protein Synthesis.* J Mol Biol, 1964. **10**: p. 63-72.

26. La Teana, A., C.L. Pon, and C.O. Gualerzi, *Late events in translation initiation. Adjustment of fMet-tRNA in the ribosomal P-site.* J Mol Biol, 1996. **256**(4): p. 667-75.

27. Weinger, J.S., K.M. Parnell, S. Dorner, R. Green, and S.A. Strobel, *Substrate-assisted catalysis of peptide bond formation by the ribosome.* Nat Struct Mol Biol, 2004. **11**(11): p. 1101-6.

28. 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.

29. Lakowicz, J.R., *Energy Transfer*, in *Principles of Fluorescence Spectroscopy, Third Edition*. 2006, Springer Science+Business Media, LLC: New York. p. 443-75.

30. Cantor, C.R. and P.R. Schimmel, *Fluorescence Spectroscopy*, in *Biophysical Chemistry*. 1980, W. H. Freeman and Company. p. 433-65.

31. 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.

32. 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.

33. 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.

34. 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.

35. 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.

36. Uemura, S., C.E. Aitken, J. Korlach, B.A. Flusberg, S.W. Turner, and J.D. Puglisi, *Real-time tRNA transit on single translating ribosomes at codon resolution.* Nature, 2010. **464**(7291): p. 1012-7.

37. Wishnia, A., A. Boussert, M. Graffe, P.H. Dessen, and M. Grunberg-Manago, *Kinetics of the reversible association of ribosomal subunits: stopped-flow studies of the rate law and of the effect of Mg2+.* J Mol Biol, 1975. **93**(4): p. 499-15.

38. Wishnia, A. and A.S. Boussert, *The non-specific role of Mg2+ in ribosomal subunit association: kinetics and equilibrium in the presence of other divalent metal ions.* J Mol Biol, 1977. **116**(3): p. 577-91.

39. Zitomer, R.S. and J.G. Flaks, *Magnesium dependence and equilibrium of the Escherichia coli ribosomal subunit association.* J Mol Biol, 1972. **71**(2): p. 263-79.

40. 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.

41. Benne, R., N. Naaktgeboren, J. Gubbens, and H.O. Voorma, *Recycling of initiation factors IF-1, IF-2 and IF-3.* Eur J Biochem, 1973. **32**(2): p. 372-80.

42. Valle, M., A. Zavialov, W. Li, S.M. Stagg, J. Sengupta, R.C. Nielsen, P. Nissen, S.C. Harvey, M. Ehrenberg, and J. Frank, *Incorporation of aminoacyl-tRNA into the ribosome as seen by cryo-electron microscopy.* Nat Struct Biol, 2003. **10**(11): p. 899-906.

43. Qin, H., C. Grigoriadou, and B.S. Cooperman, *Interaction of IF2 with the ribosomal GTPase-associated center during 70S initiation complex formation.* Biochemistry, 2009. **48**(22): p. 4699-706.

44. Kavran, J.M. and T.A. Steitz, *Structure of the base of the L7/L12 stalk of the Haloarcula marismortui large ribosomal subunit: analysis of L11 movements.* J Mol Biol, 2007. **371**(4): p. 1047-59.

45. Korostelev, A., D.N. Ermolenko, and H.F. Noller, *Structural dynamics of the ribosome.* Curr Opin Chem Biol, 2008. **12**(6): p. 674-83.

46. Agrawal, R.K., J. Linde, J. Sengupta, K.H. Nierhaus, and J. Frank, *Localization of L11 protein on the ribosome and elucidation of its involvement in EF-G-dependent translocation.* J Mol Biol, 2001. **311**(4): p. 777-87.

47. Rawat, U., H. Gao, A. Zavialov, R. Gursky, M. Ehrenberg, and J. Frank, *Interactions of the release factor RF1 with the ribosome as revealed by cryo-EM.* J Mol Biol, 2006. **357**(4): p. 1144-53.

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. 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.

50. Mesters, J.R., A.P. Potapov, J.M. de Graaf, and B. Kraal, *Synergism between the GTPase activities of EF-Tu.GTP and EF-G.GTP on empty ribosomes. Elongation*

factors as stimulators of the ribosomal oscillation between two conformations. J Mol Biol, 1994. **242**(5): p. 644-54.

51. Gualerzi, C.O., M. Severini, R. Spurio, A. La Teana, and C.L. Pon, *Molecular dissection of translation initiation factor IF2. Evidence for two structural and functional domains.* J Biol Chem, 1991. **266**(25): p. 16356-62.

Chapter 3
IF3-mediated regulation of 70S IC formation

3.1 Introduction

Having established a single-molecule platform for studying 50S subunit joining and 70S IC formation, I was next interested in using this set-up to explore the regulatory effect exerted by IF3 on these processes. It is known that IF3 plays an important role in ensuring the fidelity of start codon and initiator tRNA selection [1, 2]. Some of these effects appear to be exerted at the level of 30S IC assembly since IF3 has been shown to preferentially destabilize 30S ICs formed with elongator tRNAs or non-canonical start codons [3-5]. In addition, IF3 impacts 70S IC formation by negatively regulating 50S subunit joining to incomplete or incorrectly assembled 30S ICs. For example, light scattering measurements have shown that IF3 blocks subunit joining to 30S ICs formed in the absence of initiator tRNA [6]. IF3 probably exerts its anti-association function by sterically blocking the formation of several key intersubunit bridges at the interface between the 30S and 50S subunits [7].

Several questions remain, however, about the mechanism by which IF3 regulates the 50S subunit joining event. One especially controversial point involves the timing of IF3 dissociation in relation to other events of the initiation pathway, which has important implications as to IF3's mechanism of action. It has been proposed that spontaneous release of IF3 from the 30S IC is a prerequisite for 50S subunit joining, and that the presence of initiator tRNA increases the rate of subunit joining by speeding up IF3

release [8]. In contrast, other research groups have suggested that IF3 dissociation occurs during and not before 70S IC formation [9], and that the presence of IF3 within the 70S complex prevents conversion from a labile to a more stable form of the 70S IC [10].

Recent single-molecule data from our laboratory using a fluorescently labeled IF3 construct has suggested that, under the conditions employed in our microscope experiments, IF3 does not spontaneously dissociate from the 30S IC in response to binding of fMet-tRNA^{fMet} [11]. This implied that IF3 should be present, at least transiently, during the 50S subunit docking event. I reasoned that, if this is the case, the presence of IF3 on the 30S IC could have a detectable effect on the dynamics of 50S subunit joining and 70S IC formation as read out by my IF2-L11 smFRET signal. This chapter reports the results from experiments designed to test this hypothesis.

Addition of IF3 is shown to have dramatic effects on 70S IC formation, converting a largely irreversible process into a dynamic equilibrium in which 50S subunits reversibly dock and undock from the IF2-bound 30S IC. IF3-induced dynamic instability of the 70S IC is shown to be correlated with a change in the distribution of E_{FRET} ratios reporting on the relative conformation of IF2 and the 50S subunit's GAC. The results thus shed light on IF3's mode of action and the timing of IF3 release, as well as the mechanism of subunit joining. Most importantly, they highlight the reversibility of the IF2-catalyzed subunit joining reaction.

3.2 Preparation and Surface Immobilization of 30SIC$_{+IF3}$

30SIC$_{+IF3}$ complexes containing IF1, (Cy3)-IF2-GTP, IF3, biotin-mRNA, and fMet-tRNA^{fMet} were prepared following the procedure described in Section 5.3.1 of the

Materials and Methods. These complexes were then diluted to nanomolar concentrations in buffer containing IF1, IF3, and GTP, and introduced into the microfluidic flowcell to allow surface immobilization for TIRFM imaging (Section 5.4). In order to obtain an appropriate spot density of ~200-400 30S-bound (Cy3)-IF2 molecules per field-of-view, it was necessary to deliver approximately an order of magnitude higher concentration of $30SIC_{+IF3}$ into the flowcell compared with $30SIC_{-IF3}$ (nM versus 100s of pM, respectively). Since observation of Cy3 fluorescence requires that (Cy3)-IF2 be bound to a surface-immobilized 30S subunit at the beginning of the experiment, this suggests that the presence of IF3 destabilizes IF2 binding to the 30S subunit and/or decreases 30S IC stability in general. This is likely related to IF3's propensity to destabilize the binding of tRNAs to the 30S subunit [12]. IF3-induced dissociation of fMet-$tRNA^{fMet}$ could indirectly accelerate (Cy3)-IF2 dissociation, since the interaction between IF2's C-terminal domain VI-2 with the aminoacyl acceptor end of fMet-$tRNA^{fMet}$ constitutes one of the two major points of contact anchoring IF2 to the 30S IC surface [13]. Despite the higher concentrations of $30SIC_{+IF3}$ introduced into the flowcell, non-specific binding of (Cy3)-IF2 to the flowcell surface was negligible (Figure 3.1). The results suggest that ~99% of the (Cy3)-IF2 fluorescence spots observed originate from (Cy3)-IF2 bound to a 30S subunit, which in turn is tethered to the surface via the biotin-mRNA.

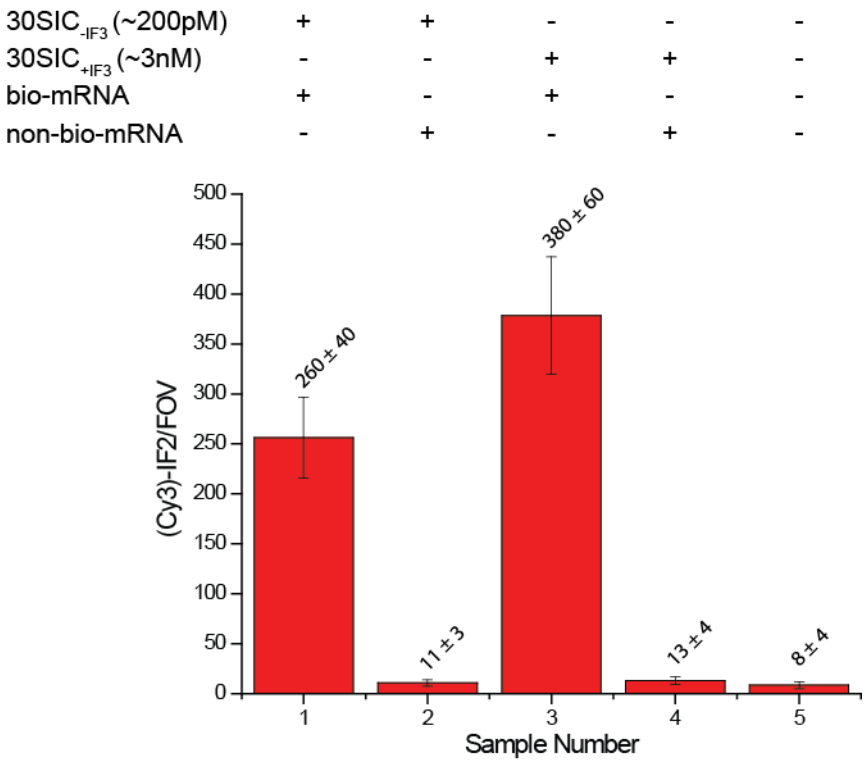

Figure 3.1: Stability of (Cy3)-IF2 binding to 30SIC$_{+IF3}$ and biotin specificity of surface immobilization.
30SIC$_{-IF3}$ contained fMet-tRNA^{fMet}, IF1, (Cy3)-IF2, and GTP, and either biotin- or non-biotin mRNA as indicated. 30SIC$_{-IF3}$ was diluted to ~200 pM and introduced into the flowcell to allow surface immobilization. Buffers used for dilutions, rinsing of the flowcell, and fluorescence imaging contained high concentrations of IF1 (0.9 μM) and GTP (1 mM). 30SIC$_{+IF3}$ contained fMet-tRNA^{fMet}, IF1, (Cy3)-IF2, IF3 and GTP, and either biotin- or non-biotin mRNA as indicated. In order to obtain an appropriate number of Cy3 spots per FOV, 30SIC$_{+IF3}$ was diluted to ~3 nM concentration before being introduced into the flowcell. Buffers used for dilutions, rinsing of the flowcell, and fluorescence imaging contained high concentrations of IF1 (0.9 μM), IF3 (0.9 μM), and GTP (1 mM). Sample number 5 corresponds to a negative control in which imaging buffer containing IF1 (0.9 μM), IF3 (0.9 μM), and GTP (1 mM) was introduced into the flowcell in the absence of (Cy3)-IF2. Five separate FOVs were imaged for each experiment. Fluorescent spots were identified by applying an intensity threshold and selecting regions containing at least two contiguous pixels above the threshold. The number of (Cy3)-IF2 molecules per FOV was quantified and the average and standard deviation for the five FOVs is depicted in the bar graph. Experiments were performed using five individual flowcells from the same microscope slide.

3.3 Direct observation of reversible 50S subunit docking to 30SIC$_{+IF3}$

In order to characterize the effect of IF3 on IF2-catalyzed subunit joining, pre-steady state smFRET experiments were performed in which (Cy5)-L11 labeled 50S subunits were delivered to surface-immobilized 30SIC$_{+IF3}$ complexes containing (Cy3)-IF2-GTP. High concentrations of unlabeled IF1 (0.9 μM), IF3 (0.9 μM), and GTP (1 mM) were kept in solution at all times during the experiment in order to ensure that the majority of 30S ICs on the surface were bound with these ligands. In addition, all observed 30S ICs necessarily contain (Cy3)-IF2, the 30S subunit, and the biotin-mRNA which specifically anchors the complex to the surface (Figure 3.1). fMet-tRNA^{fMet} was assumed to be present on the majority of 30S ICs observed to participate in subunit joining, since fMet-tRNA^{fMet} is stabilized on the 30S subunit by IF2 [8, 12] and formation of the 70S IC is dependent on the presence of both IF2 and fMet-tRNA^{fMet} [14]. This assumption seems justified based on control experiments in which high concentrations of fMet-tRNA^{fMet} (0.9 μM) were included in all dilution, wash, and imaging buffers and shown to have little effect on the observed smFRET signal (see the Appendix, Tables C.1 and C.2). Therefore, it is expected that under these conditions, I am observing 50S subunit joining to "complete" 30S ICs containing all canonical initiation components [15].

As before, the subunit joining event was indicated in the smFRET versus time trajectories by a sharp transition from zero to non-zero E_{FRET} ratios. However, in contrast to the long-lived excursions to non-zero E_{FRET} ratios observed upon subunit docking to 30SIC$_{-IF3}$ (see Chapter 2), transient excursions to non-zero E_{FRET} ratios were observed

with $30SIC_{+IF3}$, and individual $30SIC_{+IF3}$ complexes were observed to undergo multiple transitions between zero and non-zero E_{FRET} ratios during the observation period (Figure 3.2). These transient excursions to non-zero E_{FRET} ratios were initially interpreted as corresponding to individual instances of subunit docking and dissociation from the 30S IC. This interpretation is consistent with the known anti-subunit association properties of IF3. Thus, dwells at an E_{FRET} ratio of zero were initially assigned to an unbound state of the 30S IC and dwells at non-zero E_{FRET} ratios were assigned to a state in which the 50S subunit is bound. This would predict that the lifetime of dwells at an E_{FRET} ratio of zero should decrease with increasing concentrations of free (Cy5)-50S subunits in the flowcell, since subunit association is a bimolecular process whose rate should increase along with (Cy5)-50S concentration. On the other hand, the lifetime of dwells at non-zero E_{FRET} ratios would be expected to be independent of (Cy5)-50S concentration, since subunit dissociation is a unimolecular process whose rate should be unaffected by the concentration of free (Cy5)-50S subunits in solution.

A (Cy5)-50S titration was performed in order to test this prediction. The data reveal that, indeed, as the concentration of (Cy5)-50S subunits is increased, the dwell time between FRET events decreases and the average number of FRET events observed per 30S IC increases (Figure 3.2). Thus, the rate of transition from zero to non-zero E_{FRET} ratios increases with increasing (Cy5)-50S subunit concentrations, verifying that each FRET event corresponds to a separate 50S subunit docking event. Quantitative dwell time analysis was performed to calculate the lifetime of the 50S-subunit bound and free forms of $30SIC_{+IF3}$ (Figure 3.3A). The apparent pseudo-first order rate constant for the

subunit association reaction ($k'_{a,\ app}$) was calculated by taking the inverse of the lifetime in the zero-FRET state. The slope of a linear fit to the plot of $k'_{a,\ app}$ versus (Cy5)-50S concentration then provides an estimate of the apparent second-order rate constant for subunit association ($k_{a,\ app}$) (Figure 3.3B). The value of $k_{a,\ app}$ thus obtained was ~3.5 $\mu M^{-1} sec^{-1}$. This apparent association rate on the order of 10^6 $M^{-1} sec^{-1}$ is considerably slower than the expected diffusion-controlled rate for the encounter of an enzyme and substrate in solution, which is on the order of 10^8 to 10^9 $M^{-1} sec^{-1}$ [16, 17]. This suggests that the rate of association between $30SIC_{+IF3}$ and (Cy5)-50S subunits is slower than the diffusion limit and that not every collision between $30SIC_{+IF3}$ and (Cy5)-50S results in the formation of a 70S IC.

In contrast to the lifetime of dwells at an E_{FRET} ratio of zero, the lifetime of dwells at non-zero E_{FRET} ratios was independent of (Cy5)-50S concentration, consistent with a unimolecular process, in this case dissociation of the transiently formed 70S ICs back into free $30SIC_{+IF3}$ and (Cy5)-50S subunits. The average lifetime of the non-zero FRET state was 0.77 ± 0.08 sec. The inverse of this value yields an estimate for the apparent dissociation rate ($k_{d,\ app}$) of 1.3 ± 0.1 sec^{-1}. The 50S subunit association and dissociation events are indicated in TDPs as peaks centered at transitions between E_{FRET} ratios of zero and ~0.6 and between E_{FRET} ratios of ~0.6 and zero, respectively (Figure 3.2, bottom panel). Correspondingly, the time evolution of population FRET histograms display a distribution of E_{FRET} ratios for the 70S ICs that is centered at ~0.6 (Figure 3.2, middle panel).

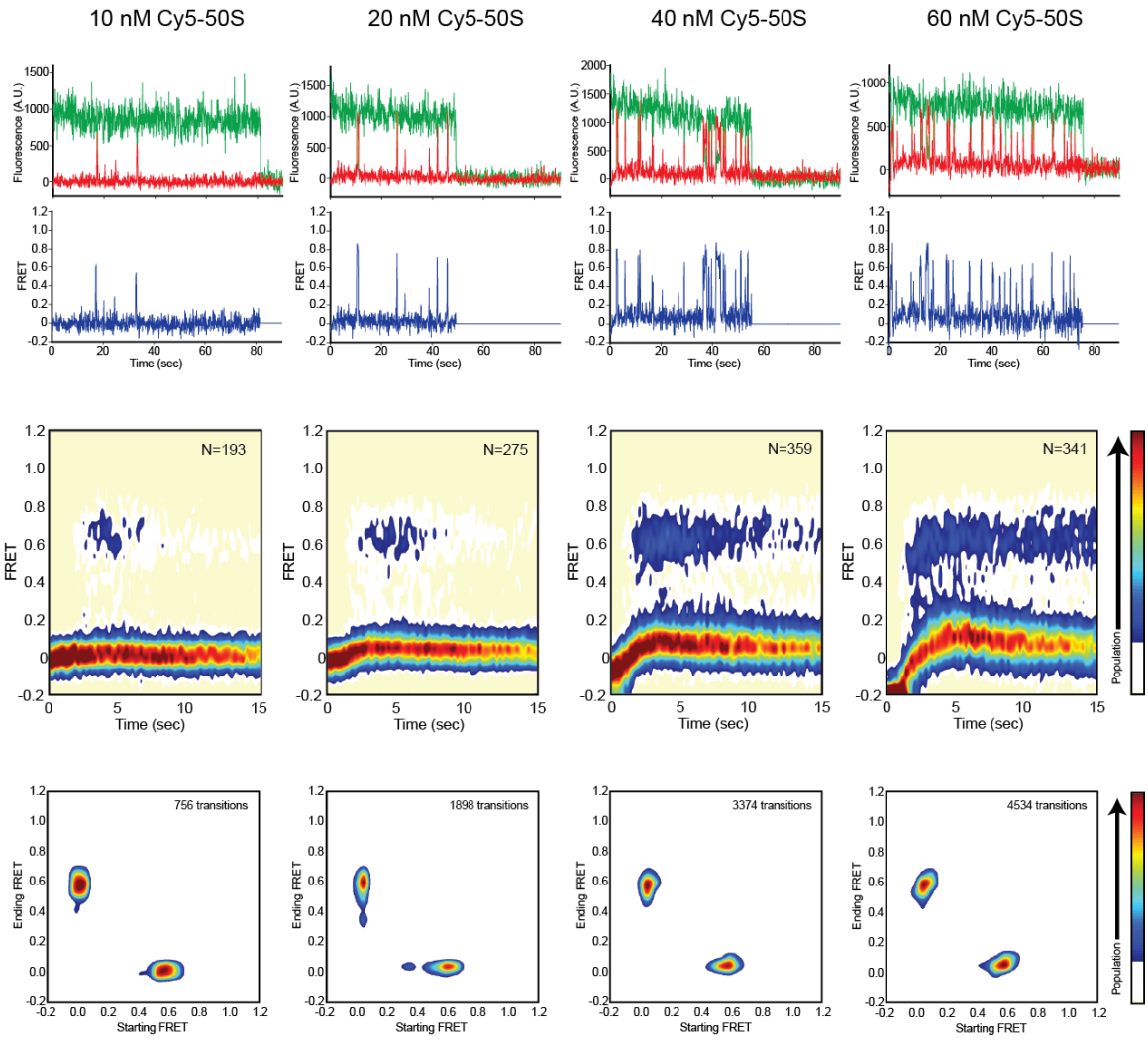

Figure 3.2: Reversible subunit docking to 30SIC$_{+IF3}$
Increasing concentrations of (Cy5)-L11 50S subunits were delivered to surface immobilized, (Cy3)-IF2–bound 30SIC$_{+IF3}$s, and Cy3 and Cy5 fluorescence emission was recorded as a function of time. The free (Cy5)-L11 50S subunit concentrations were 10 nM (A), 20 nM (B), 40 nM (C), and 60 nM (D). First row: Sample Cy3 (green) and Cy5 (red) fluorescence intensity versus time trajectories. Second row: The corresponding smFRET versus time trajectories, where FRET is calculated as $I_{Cy5}/(I_{Cy3} + I_{Cy5})$. Third row: Time evolution of population FRET histograms, made by superimposing individual smFRET trajectories. Unlike the time evolution of population FRET histograms presented in Chapter 2, these are not post-synchronized. The number of trajectories used to construct each plot is indicated by "N." Contours are plotted from tan (lowest population) to red (highest population) as indicated by the color bar. Fourth row: Transition density plots (TDPs) were generated by first idealizing the raw smFRET trajectories with vbFRET software and then plotting the "Starting FRET" versus "Ending FRET" for all transitions within the idealized trajectories as a contour plot representation of a two-dimensional histogram. Data are plotted from white (lowest population) to red (highest population) as indicated by the color bar. The number of transitions used to construct the TDP is indicated in the upper right corner.

These data suggest a different relative orientation of (Cy3)-IF2 with respect to (Cy5)-L11 within 70S ICs formed via subunit docking to $30SIC_{+IF3}$ versus $30SIC_{-IF3}$, since the distribution of E_{FRET} ratios for the latter is weighted towards higher values of ~0.8 (Chapter 2). This shift in the distribution of E_{FRET} ratios is correlated with a much higher stability of the 70S ICs formed in the absence versus the presence of IF3, which may indicate that optimal positioning of IF2 with respect to the GAC is one of the structural determinants of stable 70S IC formation. The IF3-promoted shift in the distribution of E_{FRET} ratios could be explained if IF3 binding induced or stabilized a conformation of the entire 30S IC that affected the positioning of IF2 and thus altered its interactions with the incoming 50S subunit.

Alternatively, IF3 could act by modulating the reaction kinetics underlying a multi-step subunit association process. In this scenario, IF3 would block the transition from an initially formed, unstable intermediate state (characterized by a distribution of E_{FRET} ratios shifted towards ~0.6 FRET) to a more stable state of the 70S IC (characterized by a distribution of E_{FRET} ratios shifted towards ~0.8 FRET). By inhibiting the forward reaction, the presence of IF3 would cause stalling in the intermediate state and thus preferential dissociation of the 50S subunit. This latter model requires that the unstable 70S intermediate state be traversed during both IF3-inhibited and uninhibited 50S subunit joining. Interestingly, a small percentage of the smFRET trajectories acquired for $30SIC_{-IF3}$ exhibit features which suggest that this could be the case. First, some of the trajectories exhibit a short but unambiguous dwell of ~100s of msec in a mid-FRET state (~0.5-0.6 FRET) prior to transitioning to a long-lived high FRET state (~0.8

FRET). Second, on occasion, a long-lived subunit joining event is preceded by a transient excursion to non-zero E_{FRET} ratios similar to those observed in the presence of IF3 (see Figure D.1 in the Appendix). These observations are consistent with the possibility, though they do not prove, that an obligatory subunit-joining intermediate is sampled en route to formation of the 70S IC even in the absence of IF3. If true, in the absence of IF3 the rate of exiting the intermediate state in the forward direction must be much faster than the rate of exiting in the reverse direction, such that subunit dissociation events are rarely observed. Additionally, the rate of the forward reaction must be too rapid to resolve intermediate-state dwells for the majority of ribosomes at the current time resolution. Conversely, in the presence of IF3, the rate of exiting the intermediate state in the forward direction would be slowed to such an extent that the reverse reaction (subunit dissociation) would occur almost exclusively. This model calls to mind the quantitative kinetic scheme for 70S IC formation put forth by Cooperman and coworkers on the basis of bulk fluorescence and light scattering measurements [10, 14]. Their model posits that 50S subunit joining to the IF2-bound 30S IC results in formation of an initially labile 70S IC, which stimulates GTP hydrolysis by IF2 and then either dissociates reversibly into 30S and 50S subunits or is converted into a more stable form.

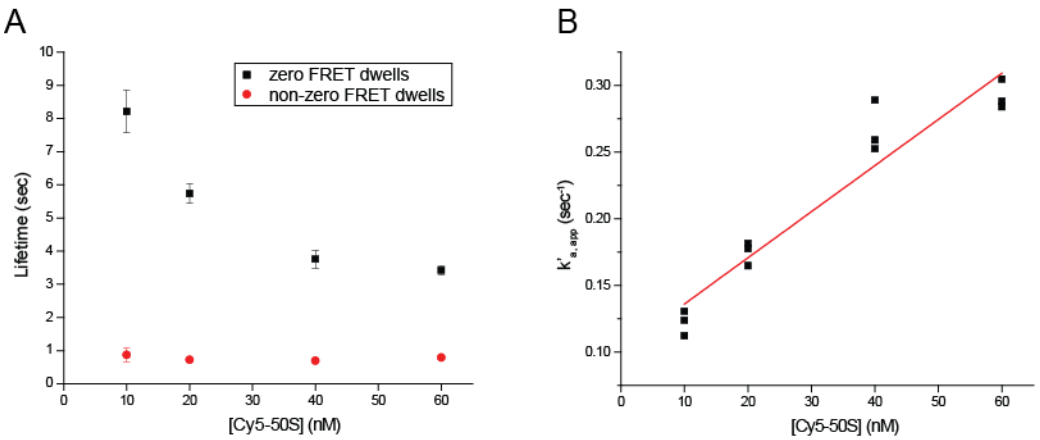

Figure 3.3: Lifetime analysis of dwells in the zero and non-zero FRET states observed upon delivery of varying concentrations of (Cy5)-50S subunits to 30SIC$_{+IF3}$
(A) Dependence of lifetimes in the zero and non-zero FRET states on free (Cy5)-50S concentration. Dwell times were extracted from idealized smFRET versus time trajectories. A threshold of FRET=0.2 was used to define "zero FRET dwells" and "non-zero FRET dwells" as corresponding to the total amount of time spent in states with FRET ≤ 0.2 and FRET > 0.2, respectively. Population decay histograms were constructed from the dwell times spent in the zero-FRET state and the non-zero FRET state. These curves were fit with a single exponential decay of the form $y=A_1*\exp(-x/t_1) + y_0$, and the average lifetime of the decay, t_1, is reported here. Errors were estimated by randomly splitting the traces into three equally sized sets, analyzing each set independently, and taking the average and standard deviation of the resulting lifetimes. (B) Estimate of the second-order rate constant ($k_{a,\ app}$) for (Cy5)-50S subunit association with 30SIC$_{+IF3}$. The pseudo-first order rate constant ($k'_{a,\ app}$) was calculated by taking the inverse of the lifetime of the zero FRET state and plotted as a function of the (Cy5)-50S concentration. The data were fit with a linear regression equation of the form y = mx + b, where m = 0.0035 ± 0.0003 $nM^{-1}sec^{-1}$ and b = 0.10 ± 0.01 sec^{-1}. The slope provides an estimate for $k_{a,app}$ of ~3.5 $\mu M^{-1}sec^{-1}$.

Regardless of whether the transient subunit docking events observed for 30SIC$_{+IF3}$ correspond to an on-pathway intermediate toward 70S IC formation or represent a dead-end complex, these smFRET data provide direct visualization of the effect of IF3 in antagonizing formation of a stable 70S IC. By binding to the 30S subunit at the platform interface, IF3 likely acts by sterically blocking formation of contacts between the 30S and 50S subunits required for stable subunit association. IF3 protects many of the 16S rRNA nucleotides at the 30S interface that are protected upon association of the 30S and 50S

subunits [7, 18]. In particular, the binding site of IF3 would appear to sterically occlude formation of intersubunit bridges B2b, B2c, and B7a [7]. Based on my data, I suggest that IF2 can promote initial docking of the 50S subunit to the 30S IC, but that if IF3 is bound to the 30S subunit, formation of these intersubunit bridges is prevented, the 70S complex is not adequately stabilized, and consequently, dissociates into its constituent parts. In other words, IF3 directly competes with the 50S subunit for stable binding to the 30S IC. In the experiments described above, this competition was artificially biased in favor of IF3 by including unlabeled IF3 in solution at high excess over both 30S ICs and 50S subunits. This was done in order to saturate 30S subunits with IF3 and allow visualization of what happens when 50S subunits encounter an IF3-bound 30S IC.

Under these circumstances, the 50S subunit is not able to bind stably, which suggested that stable subunit association can only occur when the 50S subunit encounters an IF3-free 30S IC (i.e. 30SIC$_{-IF3}$). In other words, IF3 must dissociate from the 30S IC before subunit docking occurs for the latter to result in formation of a stable 70S complex. This proposal is largely similar to that put forth by Ehrenberg and coworkers, who argue that spontaneous release of IF3 from the 30S IC is required to allow 50S subunit docking [8]. My data suggest a modification of this model by specifying that prior release of IF3 is not required for 50S subunit docking per se, but that it is required for stable subunit joining. 50S subunit docking to the 30S IC is in fact unimpeded by the presence of bound IF3, and the effect of IF3 is to instead dramatically reduce the lifetime of the resulting 70S complex. The brief, ~800 msec lifetime 50S subunit association

events with $30SIC_{+IF3}$ are likely uniquely observable using my single-molecule approach due to their transience and non-accumulating nature.

Working within the conceptual framework outlined above, I hypothesized that the reason stable subunit association events to $30SIC_{+IF3}$ were rarely seen in my smFRET assay was due to the ~45-fold molar excess of IF3 over (Cy5)-50S subunits in solution. When one molecule of IF3 dissociates from a surface-immobilized 30S IC, binding of a new IF3 molecule from solution occurs more quickly than, and effectively outcompetes, binding of a (Cy5)-50S subunit. On the other hand, if IF3 rebinding were disallowed by flushing the free IF3 out of the flowcell upon delivery of (Cy5)-50S subunits, one might expect to see the appearance of stable 50S subunit joining events. The results from this experiment will be presented in the next section.

In summary, smFRET interrogation of $30SIC_{+IF3}$ has highlighted the inherent reversibility of the IF2-promoted 50S subunit joining reaction. Ensemble measurements by Cooperman and coworkers had previously suggested that 50S subunits can transiently and reversibly associate with the IF2-bound 30S IC during the early stages of the pathway leading to 70S IC formation [14]. My single-molecule approach has allowed confirmation of this feature of the initiation pathway through direct observation. One noteworthy question that remains, however, is whether GTP is hydrolyzed during the multiple cycles of 50S docking and undocking, or whether stable subunit association is required for GTPase activation of IF2. Answering this question will require the development of a surface-based readout for GTPase activation and/or hydrolysis, perhaps involving the use of a fluorescent GTP analog such as mant-dGTP [19].

While my results demonstrating an IF3-dependent inhibition of stable 50S subunit joining are in line with IF3's known anti-subunit association properties, the effect observed here is considerably more pronounced than that reported from an independent single-molecule study of 70S IC formation [20]. In this study, subunit joining was monitored via FRET between fluorophore-labeled 30S and 50S subunits. 50S subunit delivery to surface-immobilized 30S ICs was performed in the presence of all initiation components under conditions very similar to those reported here. The authors also report the observation of fluctuations between non-zero and zero E_{FRET} ratios within individual smFRET versus time trajectories. However, they observed such fluctuations only rarely (12% of the trajectories), whereas I observed them for the majority of $30SIC_{+IF3}$ complexes (65% of the trajectories). Even more strikingly, the lifetime reported for the dwells at non-zero E_{FRET} ratios within this subset of fluctuating trajectories was 46 sec, which is roughly two orders of magnitude longer than the 0.77 sec lifetime for the dwells at non-zero E_{FRET} ratios measured here. Therefore, it seems that the dynamics of 50S subunit joining, when probed with my single-molecule assay, are much more sensitive to the inclusion of IF3. The origins of this discrepancy between my results and those reported by Marshall et al. are yet to be determined.

3.4 Partitioning of 50S subunit joining to $30SIC_{+IF3}$ into short and long association events

50S subunit delivery to $30SIC_{+IF3}$ in the presence of high concentrations of free IF3, conditions which should allow rapid recycling of IF3 on the 30S subunit, resulted in the repetitive docking and dissociation of (Cy5)-50S, with each docking event

represented by a brief (~100s of msec) burst of non-zero FRET. I hypothesized that stable subunit joining rarely occurs under these conditions because IF3 and the 50S subunit are in direct competition for stable binding to the 30S IC, and that if IF3 were to dissociate from the 30S IC, free IF3 from solution could rapidly rebind, thereby perpetuating the block on stable subunit association for an extended period of time. To test this hypothesis, I designed an experiment in which rebinding of IF3 to the 30S IC would be prevented. In this experiment, $30SIC_{+IF3}$ complexes were prepared identically as before and immobilized on the surface of the flowcell in the presence of high concentrations of IF3 in solution. (Cy5)-50S subunits were then stopped-flow delivered in a buffer from which IF3 had been omitted, such that free IF3 present in the flowcell would be flushed out at the beginning of data collection. This setup ensures that at time zero, the surface-immobilized $30SIC_{+IF3}$ complexes are saturated with IF3 but that at subsequent time points, following buffer-exchange, IF3 dissociation from the 30S IC will be essentially irreversible.

The resulting smFRET versus time trajectories (Figure 3.4) exhibited a number of distinguishing features. First, 50S subunit docking was still readily reversible, as evidenced by multiple fluctuations between zero and non-zero E_{FRET} ratios in the majority (~52%) of the trajectories. It should be noted that this value represents a lower limit to the actual fraction of 30S ICs that reversibly bind 50S subunits, since the observation time is limited by fluorophore photobleaching, which may occur before multiple FRET events can be observed.

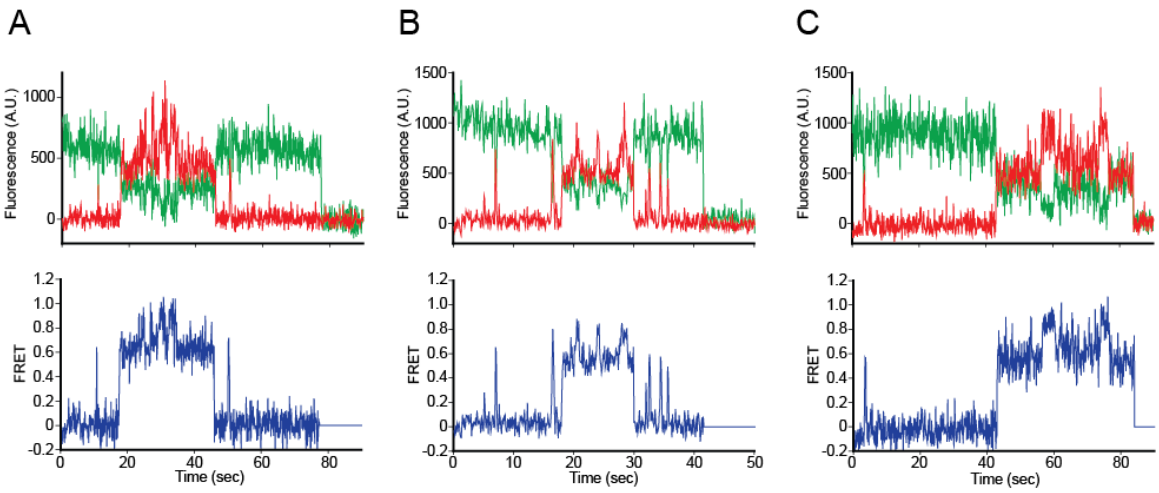

Figure 3.4: Delivery of (Cy5)-50S subunits to 30SIC$_{+IF3}$ in the absence of free IF3.
Top panel: Raw Cy3 (green) and Cy5 (red) fluorescence intensity versus time trajectories. Second row: The corresponding smFRET versus time trajectories, where FRET is calculated as $I_{Cy5}/(I_{Cy3} + I_{Cy5})$.

The interpretation of reversible subunit docking was once again confirmed by performing lifetime analysis for data collected over a range of (Cy5)-50S concentrations (Figure 3.5). The lifetime of the zero-FRET state decreased with increasing concentrations of free (Cy5)-50S subunits, which is equivalent to an increase in the rate of subunit association. The estimated apparent second-order rate of subunit association ($k_{a,\ app} \approx 2.3\ \mu M^{-1}sec^{-1}$) was very similar to the value obtained in the previous section, indicating that the presence or absence of free IF3 in solution does not significantly alter the association kinetics of the 50S subunit with the 30S IC.

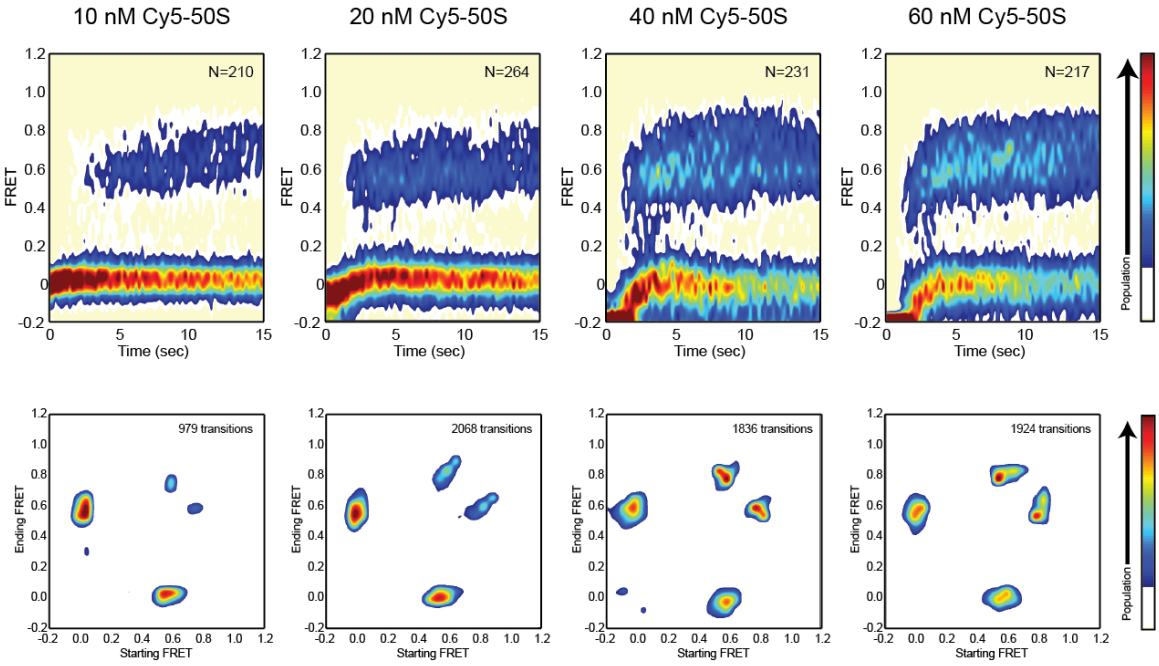

Figure 3.5: Population FRET behavior observed upon delivery of varying concentrations of (Cy5)-50S subunits to 30SIC$_{+IF3}$ in the absence of free IF3.
The (Cy5)-L11 50S subunit concentrations were 10, 20, 40, and 60 nM as indicated. Top panel: Time evolution of population FRET histograms. Bottom panel: Transition Density Plots (TDPs). Plots were generated as described in the caption to Figure 3.2.

The dissociation kinetics, on the other hand, were markedly affected by the absence of IF3 in solution. In this case, two different types of non-zero FRET events were readily apparent in the data, which differed significantly in their duration (Figures 3.4 and 3.6). Thus, both transient and long-lived FRET events were observed, corresponding to short- and long-lived association of the 50S subunit with the 30S IC, respectively. Lifetime analysis was performed for dwells in the non-zero FRET states (Figure 3.6). The population histogram constructed from these dwells was best fit with a double-exponential decay in which the lifetimes for the two components (t_1 = 0.8 sec and t_2 = 11 sec, see Table 3.1) differed by over an order of magnitude. Attempts to model the

data with a single-exponential decay resulted in a much-reduced goodness of fit. These results strengthen the argument that the smFRET data are composed of two distinct classes of non-zero FRET states, namely, one class corresponding to unstable subunit association and a second class corresponding to relatively stable subunit association. Importantly, both short- and long-lived FRET events can be observed within individual smFRET versus time trajectories (Figure 3.4), meaning that single $30SIC_{+IF3}$ complexes are capable of participating in both stable and transient binding to the 50S subunit.

Several features of the longer-lived subunit association events deserve special attention. First, like the shorter-lived subunit association events, they too are apparently reversible within the experimental observation window. This is evidenced by smFRET versus time trajectories in which a long-lived dwell is followed by a transition to an E_{FRET} ratio of zero, and, subsequently, one or more fluctuations back to non-zero E_{FRET} ratios (e.g. Figure 3.4A and B). This scenario is contrasted with smFRET versus time trajectories in which the last dwell at non-zero E_{FRET} ratios of the trajectory is long-lived (e.g. Figure 3.4C), in which case it is necessarily ambiguous whether the transition to an E_{FRET} ratio of zero is caused by (Cy5)-50S subunit dissociation or fluorophore photobleaching. Since this latter behavior is observed in a significant subpopulation of the trajectories, it is likely that observation of some of the longer-lived dwells is prematurely truncated by photobleaching and, consequently, that the estimated lifetime for these dwells is a lower limit on the true lifetime. Nevertheless, it is evident that formation of these relatively stable 70S complexes constitutes a reversible step within the 70S IC assembly pathway.

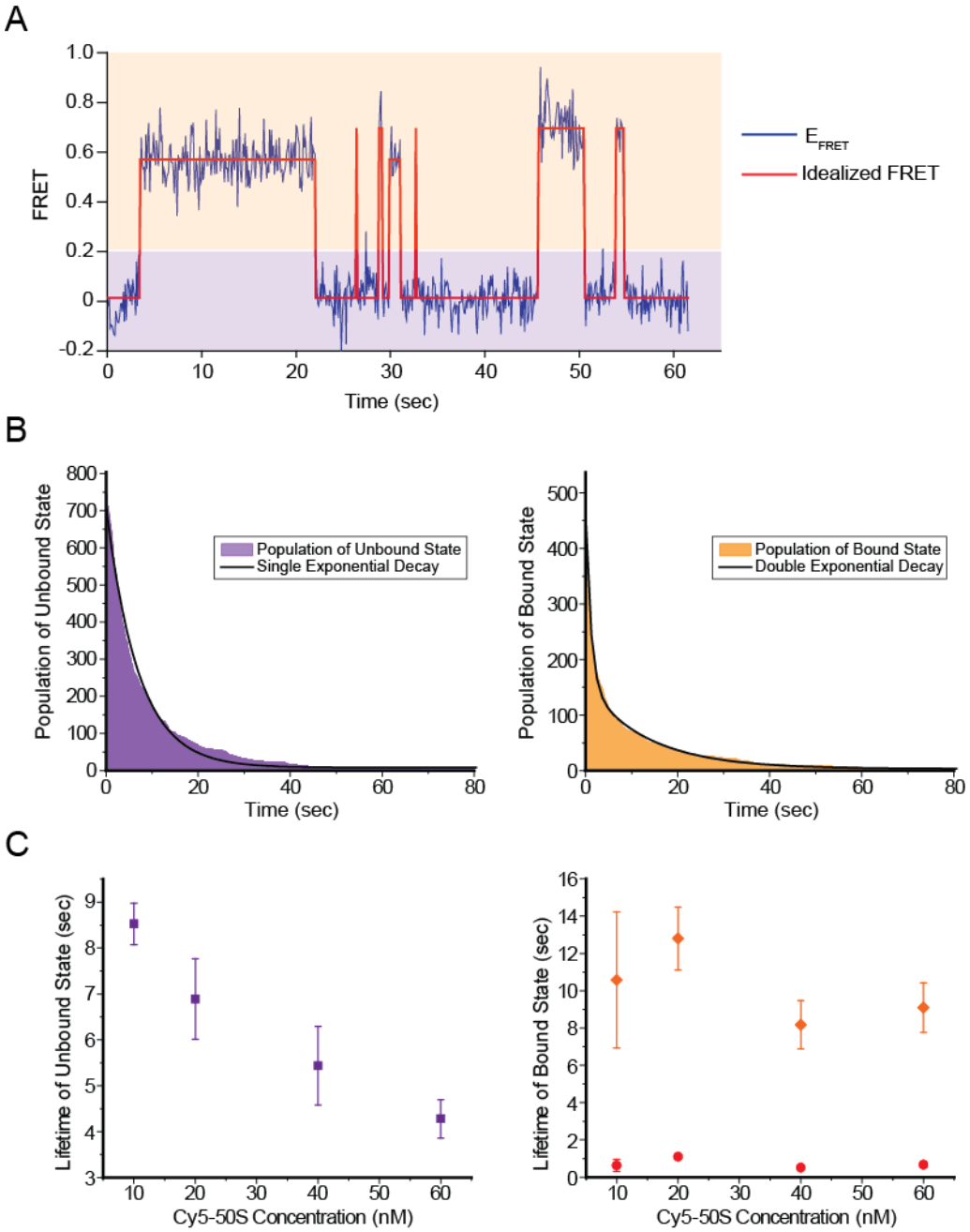

Figure 3.6: Lifetime analysis of the unbound and 50S-subunit bound states of 30SIC$_{+IF3}$ in the absence of free IF3.

(A) Sample smFRET versus time trajectory (blue) overlayed with the corresponding idealized trajectory (red). Dwells in the zero and non-zero FRET states were separated by a threshold of FRET=0.2 and the corresponding dwell times were extracted from the idealized smFRET trajectories. Dwells in the zero-FRET state (purple) correspond to the unbound state of 30SIC$_{+IF3}$ and dwells in the non-zero FRET state (orange) correspond to the 50S-subunit bound state. (B) Population decay histograms were constructed from dwell times spent in the unbound and bound

states. The histogram of dwell times in the unbound state (left) was well fit by a single exponential decay of the form $y=A_1*\exp(-x/t_1) + y_0$, with t_1 providing an estimate of the state's average lifetime. The histogram for the 50S-subunit bound state (right), however, was better fit by a double exponential decay of the form $y=A1*\exp(-x/t_1) + A2*\exp(-x/t_2) + y_0$. The parameters t_1 and t_2 provide an estimate for the average lifetime of 70S complexes with low and intermediate stability, respectively. As described in the text, t_2 is likely limited by fluorophore photobleaching and thus represents a lower limit to the true lifetime. (C) Dependence of the average lifetimes on the concentration of (Cy5)-50S subunits. Left: average lifetime (t_1) of the unbound state. Right: average lifetimes for bound states with low (t_1, red) and intermediate (t_2, orange) stability. Errors were estimated as described in the caption to Figure 3.3.

A second notable feature of the longer-lived subunit association events is the presence of fluctuations between at least two non-zero FRET states, which appear as prominent peaks in the TDPs corresponding to transitions from E_{FRET} ratios of ~0.6 to 0.8 and from E_{FRET} ratios of ~0.8 to 0.6 (Figures 3.4 and 3.5). These FRET fluctuations are indicative of conformational dynamics within the 70S IC, in particular changes in distance between L11 and IF2. This suggests a relative rearrangement of the GAC with respect to IF2 that involves at least two interconvertible structural configurations. These dynamics are reminiscent of the non-zero FRET fluctuations between states centered at E_{FRET} ratios of ~0.6 and ~0.8 observed in a subset of the molecules upon subunit joining to $30SIC_{-IF3}$ complexes (Chapter 2). The distribution of E_{FRET} ratios for $30SIC_{+IF3}$ is weighted towards the ~0.6 FRET state, whereas for $30SIC_{-IF3}$, it is weighted towards the ~0.8 FRET state (see Figure 3.8 below). Subunit association and dissociation appear to occur preferentially from the ~0.6 FRET state for $30SIC_{+IF3}$, as evidenced by the predominance of ~0 to 0.6 and ~0.6 to 0 FRET transitions in the TDPs (Figure 3.5), suggesting that the major pathway for these events involves a specific interaction mode between IF2 and the 50S subunit's GAC.

A simple explanation for the observation of both short- and long-lived subunit association with $30SIC_{+IF3}$ would be that the short-lived events arise from subunit docking to the IF3-bound 30S IC while the long-lived events correspond to subunit docking to IF3-free 30S ICs from which the factor had previously dissociated. However, there is reason to believe that such a simple model does not adequately account for all of the data. If IF3 were released from $30SIC_{+IF3}$, the resulting complex would be compositionally identical to $30SIC_{-IF3}$. It would therefore be expected that the long-lived dwells at non-zero E_{FRET} ratios should resemble the analogous dwells observed upon subunit delivery to $30SIC_{-IF3}$ (Chapter 2). In reality, however, they are significantly different, both in terms of their distributions of E_{FRET} ratios and their apparent lifetimes. When the long-lived dwells within the $30SIC_{+IF3}$ dataset were isolated and analyzed independently, their distribution of E_{FRET} ratios was characterized by two roughly equally sized peaks centered at E_{FRET} ratios of ~0.6 and 0.8 (Figure 3.7). (For the sake of this analysis, “long-lived” non-zero FRET events were defined as those lasting longer than four seconds; see caption to Figure 3.7 and in-text discussion below.) The distribution of E_{FRET} ratios for long-lived dwells within the $30SIC_{-IF3}$ dataset, on the other hand, was shifted towards E_{FRET} ratios of ~0.8. The long-lived dwells for $30SIC_{+IF3}$ also appear to have a shorter characteristic lifetime than the long-lived dwells for $30SIC_{-IF3}$, as is evidenced by occurrences of 50S subunit dissociation from these dwells in the former, but not the latter, dataset. Taken together, these results suggest that the long-lived subunit association events observed with $30SIC_{+IF3}$ and $30SIC_{-IF3}$ correspond to the formation of structurally and dynamically distinct 70S complexes. This, in turn, leads me to suspect

that IF3 is not released from the 30S IC prior to subunit docking and that it is, in fact, physically present on the 70S IC influencing ribosome conformation and stability. Assuming this to be the case, the partitioning between short- and long-lived 50S subunit association events to $30SIC_{+IF3}$ could be explained by an underlying dynamic equilibrium between two conformations of the IF3-bound 30S IC, one which binds to the 50S subunit stably and one that does not. I expound upon this mechanistic model for IF3-dependent regulation of 50S subunit joining in Section 3.5.

The degree of partitioning between relatively stable and unstable subunit association events was quantitatively estimated from the amplitudes of the slow and fast components, respectively, of the double-exponential fit to the dwell time histograms. The fast component accounts for roughly two-thirds of the decay and the slow component accounts for the remaining one-third, for all concentrations of free (Cy5)-50S tested (Table 3.1). This result suggests that approximately two out of every three 50S subunit docking events are short-lived, or in other words, that each attempt at 50S subunit joining to the 30S IC has a ~33% chance of resulting in the formation of a relatively stable 70S IC. Similar results were obtained when this ratio was estimated by defining "long-lived" events as those lasting more than four seconds, and counting the number of short- and long-lived events observed across the entire dataset. The four second cut-off was chosen based on the fact that, for datasets in which the dwell time histogram is well-described by a single-exponential decay containing only a fast time constant (i.e. $30SIC_{+IF3}$ in the presence of saturating concentrations of free IF3, see Section 3.3), this cut-off results in nearly all ($\geq$95%) of the dwells at non-zero E_{FRET} ratios being categorized as short-lived.

Using this approach, roughly three-quarters of the 50S subunit docking events were classified as short-lived and one-quarter as long-lived, again independent of the free (Cy5)-50S concentration (see the Appendix, Table C.3). In summary, the data suggest that, on average, approximately two to three transient 50S docking events occur for every instance of stable 70S IC formation.

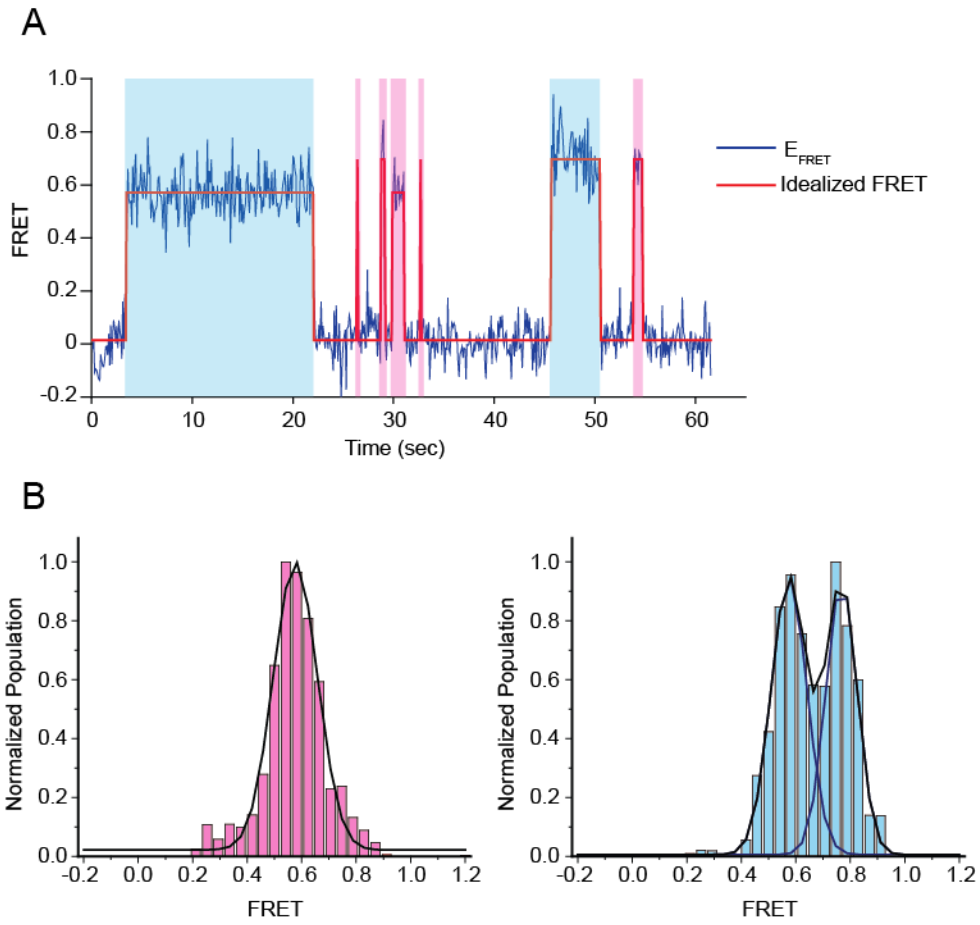

Figure 3.7: Short- and long-lived subunit association events are characterized by different FRET distributions.
(A) Raw smFRET versus time trajectories (blue) were idealized to a hidden Markov model (red), and subunit association events were identified as dwells with FRET > 0.2. Individual subunit association events were sorted into two categories, short- and long-lived, based on their dwelltimes. Short-lived subunit association events were defined as those with dwelltime < 4 sec (highlighted in pink) and long-lived events were defined as those with dwelltime ≥4 sec (highlighted in light blue). (B) FRET distributions for the short- (left) and long-lived (right) subunit association events. FRET data from the idealized smFRET trajectories for all (Cy5)-50S concentrations was plotted and normalized to the most populated bin. The distributions were fit with a single Gaussian (left) or a sum of two Gaussians (right).

I next tested whether the identity of the mRNA's start codon affects the partitioning between long and short subunit association events by repeating the experiment with $30SIC_{+IF3}$ complexes assembled on an mRNA bearing a non-canonical AUU start codon ($30S_{+IF3}^{AUU}$). The gene encoding IF3 (*infC*) begins with an AUU start codon, and its expression is autoregulated at the level of translation initiation by IF3 *in vivo* [21]. IF3-dependent repression of *infC* translation is specifically dependent on the identity of the start codon, since its mutation from AUU to AUG results in a loss of regulation [22]. Recent biochemical experiments have suggested that IF3's discrimination of the start codon is achieved during the conversion of the 30S IC to the 70S IC, through a mechanism in which transition of the 70S IC from a labile to a stable form is inhibited in the presence of IF3 and an AUU-bearing mRNA [10]. I thus sought to investigate whether subunit joining, as probed by my smFRET assay, would show a similar sensitivity to the substitution of AUU for AUG.

However, no significant differences were detected between the smFRET versus time trajectories generated upon delivery of (Cy5)-50S subunits to $30S_{+IF3}^{AUU}$ versus $30S_{+IF3}^{AUG}$. Indeed, the apparent subunit association rates were very similar, as were the average lifetimes and relative occupancies of the short- and long-lived association events (Table 3.1). This may suggest that discrimination of the AUU start codon happens at a step along the initiation pathway that is not being probed by my smFRET assay. One possibility is that IF3-mediated discrimination takes place at the level of 30S IC assembly. Biochemical experiments have demonstrated that IF3 preferentially dissociates 30S ICs containing codons other than AUG, GUG, or UUG and/or tRNA other than

$tRNA^{fMet}$ at the P site [3, 23, 24]. Additionally, IF2's binding affinity to the 30S IC is decreased when AUG is replaced with AUU, according to steady-state smFRET experiments conducted in our laboratory [25]. In either of these cases, an overall decrease in translation initiation could be accounted for by a reduction in the fraction of 30S ICs that are fully assembled and primed for 50S subunit joining. My smFRET assay, however, would not report on this level of regulation, since data is only collected from the subpopulation of 30S ICs that are still intact at the beginning of the experiment.

It is also possible that differences in subunit joining behavior for $30S_{+IF3}^{AUU}$ and $30S_{+IF3}^{AUG}$ exist, but do not become manifest during the short observation window (~20 sec on average) to which my measurements are confined. Ensemble measurements are typically conducted at 50S subunit concentrations one to two orders of magnitude higher than the small, ~10s of nM concentrations of (Cy5)-50S subunits used here in order to minimize background fluorescence in the single-molecule experiments. Therefore, the subunit association rate will be much faster, and the number of subunit association events per unit time will be much higher, in the context of an ensemble experiment. This could cause differences between AUU and AUG to become apparent at earlier time points in the ensemble version of the experiment than in the single-molecule version. One way to investigate this possibility, which is currently being pursued, is to use a laser shuttering strategy in which, rather than imaging during the first ~20 sec following delivery of (Cy5)-50S subunits, the image acquisition is delayed so as to capture the subunit association behavior at later periods of time.

Table 3.1: Lifetime analysis of the 50S-subunit bound state of $30SIC_{+IF3}$ under various conditions.
When IF3 was not included in solution, the dwell time histograms were best fit by a double exponential decay with fast and slow components t_1 and t_2. When 0.9 μM IF3 was included in solution, the dwell time histograms were well described by a single exponential decay with average lifetime t_1.

IF3 in solution[a]	Start codon	[Cy5-50S], nM	t_1, sec (A1, %)[b,c]	t_2, sec (A2, %)[b,c]
-	AUG	10	0.7 ± 0.2 (64 ± 4%)	11 ± 2 (36 ± 4%)
-	AUG	20	1.1 ± 0.3 (68 ± 4%)	13 ± 3 (32 ± 4%)
-	AUG	40	0.7 ± 0.3 (60 ± 4%)	10 ± 2 (40 ± 4%)
-	AUG	60	0.7 ± 0.2 (59 ± 4%)	9 ± 1 (41 ± 4%)
-	AUU	20	1.3 ± 0.4 (67 ± 5%)	15 ± 4 (33 ± 5%)
+	AUG	10	0.9 ± 0.1	N.A.
+	AUG	20	0.8 ± 0.1	N.A.
+	AUG	40	0.7 ± 0.1	N.A.
+	AUG	60	0.8 ± 0.1	N.A.
+	AUU	20	0.6 ± 0.1	N.A.

(a) In all experiments, IF3 was included in the buffers used to dilute $30SIC_{+IF3}$s and to rinse the flowcell following surface immobilization. The stopped-flow buffer delivered into the flowcell contained (Cy5)-50S subunits, IF1, and GTP, either in the presence or absence of IF3 as indicated. fMet-$tRNA^{fMet}$ was not included in wash, dilution, or stopped-flow buffers except for the experiment in the last row. Control experiments in which fMet-$tRNA^{fMet}$ was included in all buffers demonstrated little impact on the measured lifetimes (see Table C.2 in the Appendix).

(b) When IF3 was not included in solution, dwell time histograms were fit with a double exponential decay of the form $y=A1*\exp(-x/t_1) + A2*\exp(-x/t_2) + y_0$. The percent contribution of the fast and slow components to the decay were estimated as A1/(A1+A2)*100 and A2/(A1+A2)*100, respectively.

(c) Errors were estimated by bootstrapping 1000 random samples of the experimental dwell times and determining the mean and standard deviation of the resultant values. Bootstrapping procedures were implemented using code written in R [26].

Finally, I am looking into the possibility that the specific step at which IF3 discriminates the codon-anticodon interaction, as well as the extent of discrimination, is dependent on additional features of the mRNA's translation initiation region (TIR) besides just the start codon, such as the strength of the Shine-Dalgarno (SD) sequence and the length of the spacer between SD sequence and start codon, which have been shown to have a large impact on translation initiation efficiency [27]. The T4gp32 mRNAs used for the smFRET experiments reported here have a five-nucleotide SD sequence and a seven-nucleotide spacer (Table A.1). The sequence context of the start codon is thus different from that of the mRNAs used in the *in vitro* studies where IF3-mediated repression of translation starting from AUU was traced to the subunit joining step of the initiation pathway [10, 27]. In the future, I plan to vary these components of the mRNA's TIR in order to seek an explanation for the apparent discrepancies between my single-molecule data with others' ensemble results, and to gain an understanding of whether and how they influence the subunit joining reaction.

3.5 Mechanistic model for IF3-dependent regulation of 50S subunit joining

As demonstrated above, 50S subunit joining is highly sensitive to the presence or absence of IF3 on the 30S IC as well as the presence or absence of free IF3 in solution. Depending on the conditions, 50S subunit docking to the 30S IC can be transient and reversible or stable and largely irreversible. Figure 3.8 summarizes the smFRET data collected under the various conditions. When (Cy5)-50S subunits were delivered to $30SIC_{+IF3}$ in the presence of saturating concentrations of free IF3, transient subunit

association and dissociation events were observed, with a lifetime for the bound complex of ~800 msec. These short-lived interactions were characterized by E_{FRET} ratios of ~0.6. When, instead, free IF3 was not kept in solution—conditions under which recycling of IF3 was prevented—delivery of (Cy5)-50S subunits to $30SIC_{+IF3}$ led to the observation of two types of subunit association events. The first type was characterized by a short, ~800 msec lifetime and an E_{FRET} ratio of ~0.6; these events were thus indistinguishable from the transient association events observed previously. The second type was characterized by an intermediate lifetime (~11 sec, though likely underestimated due to fluorophore photobleaching, see above) and a distribution of E_{FRET} ratios with two peaks centered at ~0.6 and 0.8. Finally, when Cy5-50S subunits were delivered to 30S ICs devoid of IF3 (i.e. $30SIC_{-IF3}$), subunit association resulted in the formation of very stable 70S ICs (estimated lifetime of ~100 sec, see Section 2.6.3) whose distribution of E_{FRET} ratios was shifted towards ~0.8. Taken together, the smFRET data indicate the presence of at least three distinct classes of 70S IC, possessing either low, intermediate, or high stability.

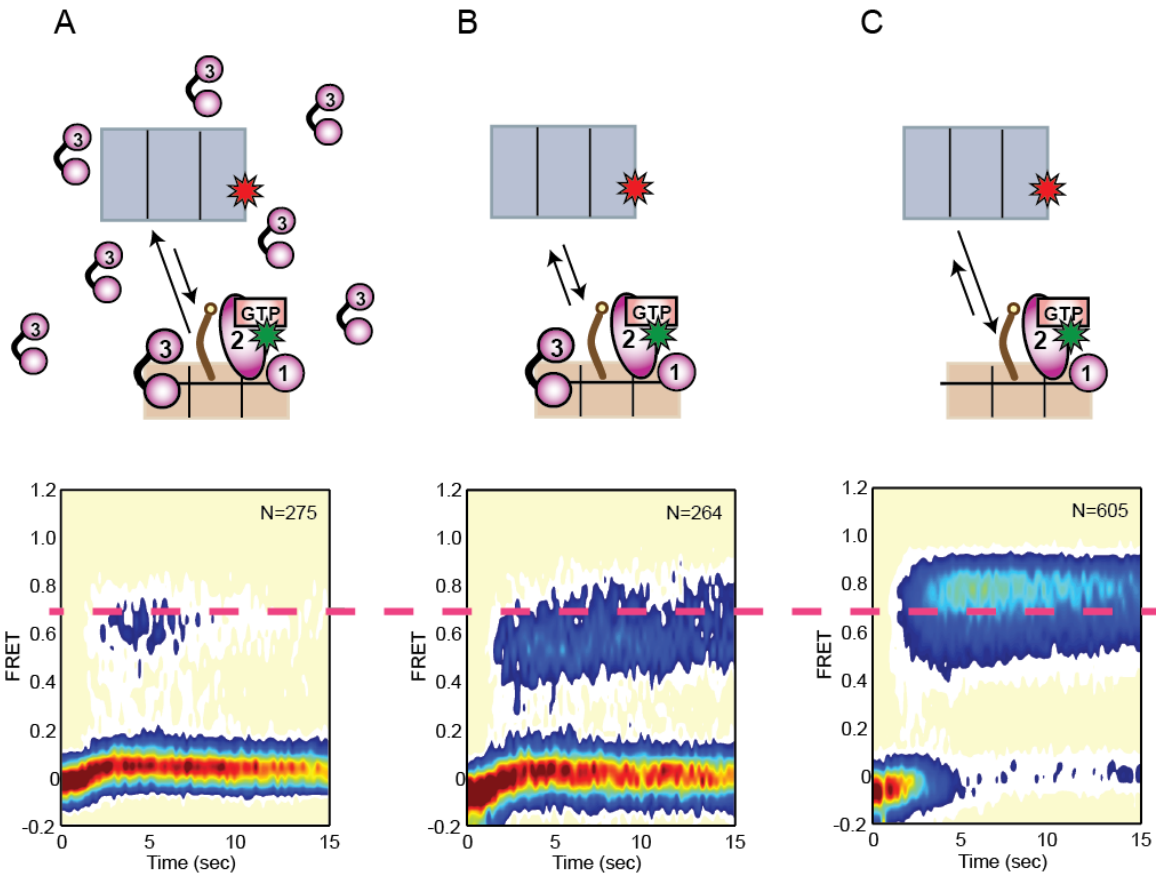

Figure 3.8: The stability and FRET distribution of 70S complexes formed upon 50S subunit joining depend on the presence and concentration of IF3.
Top row: Cartoons of the various conditions tested. (Cy5)-50S subunits were delivered to (A) $30SIC_{+IF3}$ in the presence of saturating concentrations of free IF3, (B) $30SIC_{+IF3}$ in the absence of free IF3, or (C) $30SIC_{-IF3}$. Bottom row: Time evolution of population FRET histograms, generated and plotted as described in the caption to Figure 3.2. The pink dashed line at FRET=0.7 is drawn to illustrate the shift of the distribution from ~0.6 to ~0.8 FRET. All data were collected at a (Cy5)-50S subunit concentration of 20 nM.

These observations could be explained by a model in which the 30S IC can adopt three separate compositionally and/or conformationally distinct states, each of which interacts with the incoming 50S subunit in a unique way (Figure 3.9). According to this model, a transient association event would occur when the 50S subunit encounters a 30S IC in an "inhibitory" state, such that relatively few and/or weak intersubunit contacts are

formed. Conversely, stable subunit association would occur when the 50S subunit binds to a “productive” state of the 30S IC that can participate in rapid formation of an optimal pattern of intersubunit interactions. In the absence of IF3, the vast majority of 30S ICs appear to occupy a productive conformation that can efficiently bind the 50S subunit and, presumably, quickly form the full complement of intersubunit bridges along the subunit interface. Binding of IF3 switches the 30S IC from the productive state to less productive ones. IF3’s binding site at the platform region of the 30S subunit interface overlaps with a subset of intersubunit bridges, and the presence of IF3 could sterically block their formation during subunit association [7]. In addition, IF3-induced structural changes of the 30S subunit [28, 29] could contribute to its assuming an inhibitory conformation.

An important finding from my data is that binding of IF3 does not impart uniform destabilization of the 70S IC. Instead, the data suggest that the IF3-bound 30S IC can assume at least two distinct conformations that inhibit subunit association to different degrees. The idea that the IF3-bound 30S IC can adopt multiple conformations has strong support from smFRET data collected in our laboratory using a doubly labeled IF3 construct containing donor and acceptor fluorophores attached to its N- and C-terminal domains [11]. Steady-state smFRET experiments revealed the existence of multiple discreet, slowly interconverting conformations of the IF3-bound 30S IC, characterized by different distances between IF3’s two domains. These large-scale interdomain rearrangements of IF3 are likely facilitated by the long and flexible linker region that connects IF3’s globular C- and N-terminal domains [30, 31]. The interdomain rearrangements of IF3 might be coupled to structural changes of the 30S subunit, or

alternatively, might allow IF3's globular domains to sample different binding sites on the 30S platform. The latter explanation could help account for apparent discrepancies in the literature regarding the precise location of IF3's binding site, as well as the fact that IF3 has been crosslinked to r-proteins and rRNA elements spread over such a wide area of the 30S subunit interface [32-35]. Based on these considerations, I speculate that transitioning of the IF3-30S complex between different conformations could be accompanied by steric occlusion of more or less intersubunit bridges by IF3, which would result in the formation of 70S ICs with low or intermediate stability, respectively (Figure 3.9).

Another notable feature of the smFRET data is that, in the presence of high concentrations of free IF3, nearly all observed 50S subunit docking events result in formation of 70S ICs with low stability. In the context of the mechanistic model outlined above, this indicates that the dynamic equilibrium of the IF3-30S complex is shifted heavily towards the inhibitory conformation, which in turn implies that the presence of free IF3 in solution somehow increases the occupancy of this conformation over the alternate one. This is a curious finding, since, as discussed above, both conformations are expected to correspond to IF3-bound 30S ICs. In other words, in order for free IF3 to influence the conformational equilibrium, it apparently must interact with a 30S IC that already contains a bound molecule of IF3. This suggests that the 30S IC may be able to simultaneously accommodate more than one IF3 molecule at a time, a hypothesis for which there are, in fact, glimmers of support in the literature.

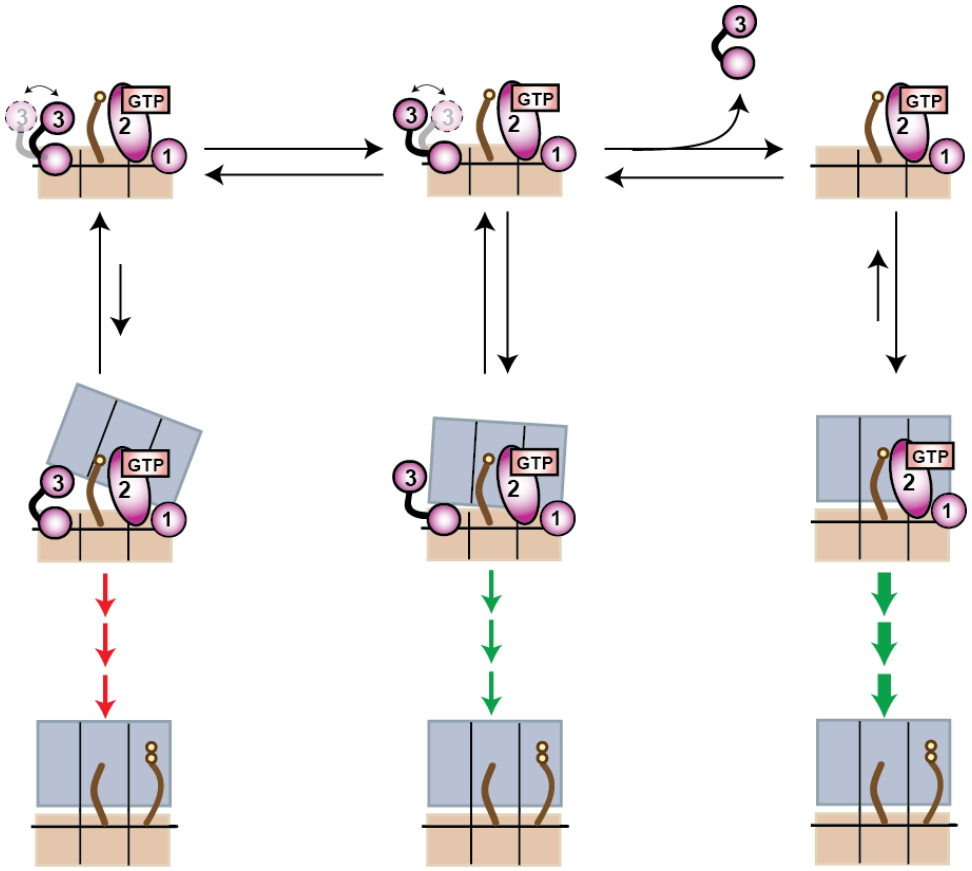

Figure 3.9: Mechanistic model for IF3-mediated regulation of 70S IC formation.
In this model, the 30S IC exists in a dynamic equilibrium between at least three compositionally and conformationally distinct states: two conformations of the IF3-bound 30S IC and the IF3-free 30S IC. Each state of the 30S IC interacts with the incoming 50S subunit in a unique way, resulting in the formation of 70S complexes with low, intermediate, or high stability and a corresponding shift in the equilibrium between 70S complexes and free subunits. I propose that more stable 70S complexes have a higher probability of binding the first ternary complex and entering into the elongation stage of protein synthesis. Transitions between the two conformations of the IF3-bound 30S IC are depicted in the cartoon as an interdomain reconfiguration of IF3's N- and C-domains that results in the occlusion of more or less intersubunit bridges.

One line of evidence is the X-ray crystallographic study in which IF3's isolated C-domain was soaked into 30S subunit crystals and found to bind to the solvent side, rather than the interface side, of the 30S platform, between h23, h26, and the 3' end of h45 [36]. It has been noted that binding of IF3 to the interface side of the platform—

where it has been localized by multiple chemical probing [7, 37] and cryo-EM [15, 29] studies—would not be possible within the context of the crystal lattice due to crystal contacts between neighboring 30S subunits [38]. Despite this, the localization of IF3's C-domain to the solvent side of the platform as seen in the crystal structure is consistent with crosslinks identified between IF3 and certain rRNA nucleotides and r-proteins [32, 39], and could correspond to a second, low-affinity binding site that only becomes substantially occupied at high concentrations of IF3. In this scenario, free IF3 would bind to the secondary site and cause stabilization of the inhibitory conformation of $30SIC_{+IF3}$, thereby resulting in a near-complete block of stable subunit association.

A second possibility arises from the finding that IF3's two domains bind independently to the 30S subunit and with different affinities [40]. Titration of isotopically labeled IF3 with 30S subunits was followed by two-dimensional heteronuclear NMR spectroscopy, with the results suggesting the existence of an equilibrium between two species of IF3 on the 30S subunit: a "partly bound" species in which only IF3's C-domain interacts with the 30S surface, and a "fully bound" species in which the 30S binding sites for both the C-domain and N-domain are occupied. According to this model, the C-domain is tightly anchored to the 30S subunit, while the lower affinity N-domain reversibly interacts with its binding site in order to modulate the thermodynamic stability of IF3 on the ribosome [4, 40]. It seems reasonable to speculate that 50S subunit association would be more strongly opposed when both the C-domain and N-domain are bound to the 30S surface; if so, the fully bound species of IF3 may correspond to the inhibitory conformation of $30SIC_{+IF3}$. In this scenario, occupancy of the

inhibitory conformation could be maintained by saturating the binding sites for both the C-domain and N-domain. This would be facilitated by the presence of free IF3 in solution if the N-domain from one molecule of IF3 were able to bind to its site even when the C-domain binding site is occupied by a separate IF3 molecule. This idea could be tested in the future by including the isolated N-domain in solution rather than full-length IF3, to see if the same effect on subunit joining is observed.

In summary, my data show that stable subunit joining is inhibited by IF3 and that this inhibition is strengthened at high IF3:ribosome ratios. The presence of IF3 on the 30S IC converts IF2-catalyzed subunit joining from a largely irreversible reaction into a dynamic equilibrium between 70S ICs and free subunits, thereby providing an opportunity for regulation of translation initiation at the subunit joining step. Since only 70S ICs can enter into the elongation phase of protein synthesis, the efficiency with which a given mRNA is translated could be modulated by shifting the equilibrium in one direction or the other (Figure 3.9).

My data further suggest a model in which the productivity of 50S subunit docking is controlled through IF3-dependent changes in the compositional and conformational state of the 30S IC, a proposal which is very similar to that put forth by Rodnina and coworkers [27]. The IF3-free 30S IC forms stable 70S ICs, while two interconverting conformations of the IF3-bound 30S IC form 70S ICs of either low or intermediate stability. By controlling the fractional occupancy of the different 30S IC conformations, the efficiency of 70S IC formation and thus translation initiation could be manipulated. Future experiments will be geared toward testing this hypothesis by making use of

mRNA constructs that are translated with different efficiencies *in vitro* [27] and by extending my single molecule assay to monitor incorporation of the first EF-Tu:GTP:aa-tRNA ternary complex into the 70S IC.

Previous single-molecule studies have suggested that the formation of a 70S complex per se may not be sufficient for entry into elongation, since the 70S IC must acquire a particular conformation to efficiently accommodate the first ternary complex into the ribosomal A site [20, 41]. In this regard, it seems worth revisiting the observation that my IF2-L11 smFRET data indicates the presence of at least two interconvertible conformations of IF2 with respect to the GAC within the 70S IC, characterized by E_{FRET} ratios of ~0.6 and ~0.8. In addition, the progression from strong inhibition of subunit association to stable subunit joining coincides with a gradual shift in the distribution of E_{FRET} ratios to higher occupancy of the ~0.8 FRET state (Figure 3.8). Therefore, it could be that this conformational equilibrium of the 70S IC plays an additional role in regulating the efficiency of entry into elongation, and this possibility will serve to motivate future experiments.

3.6 Open questions and future directions

3.6.1 Timing of IF3 release from the ribosome

The presence of IF3 on the 30S IC alters the stability and distribution of E_{FRET} ratios for the 70S ICs formed upon 50S subunit docking, as outlined above. Both the short- and intermediate-lifetime 70S ICs observed with $30SIC_{+IF3}$ are unique compared with the stable, long-lifetime 70S ICs formed with $30SIC_{-IF3}$. The simplest explanation

for this observation is that, when included, IF3 is physically present on the 70S ICs so as to directly affect their dynamics. Thus, my data strongly suggest that IF3 does not dissociate from the 30S IC prior to 50S subunit docking. This has not been directly demonstrated, however, and alternative mechanisms to account for IF3's regulatory effects cannot be entirely ruled out, such as the possibility that the presence of IF3 causes an irreversible conformational change of the 30S IC that affects its subunit joining capacity and persists even after IF3 has been released.

The answer to the question of when IF3 dissociates from the ribosome is of importance for understanding the factor's mechanism of action, and it is, furthermore, a matter of considerable controversy in the literature. Ehrenberg and coworkers have argued that release of IF3 from the 30S IC is a prerequisite for 50S subunit docking, and that the presence of fMet-tRNA^{fMet} favors the latter reaction by causing an increase in the rate of IF3 dissociation from the 30S subunit [8]. On the other hand, results from Rodnina and coworkers have suggested that IF3 is tightly associated with the 30S IC and that binding of fMet-tRNA^{fMet} does not cause any significant dissociation of the factor. In their model, 50S subunits bind to the IF3-bound 30S IC and actually slow the rate of IF3 release [27]. The presence of IF3 within the 70S IC was also suggested by bulk biochemical data from Cooperman and coworkers [10], and the cryo-EM reconstruction of the *E. coli* 70S IC contains electron density that may be attributable to one of IF3's globular domains [42].

I plan to directly test for the presence of IF3 on the 70S IC through a modification of my (Cy3)-IF2/(Cy5)-L11 smFRET subunit joining assay that makes use of an IF3

construct labeled with Atto488, whose fluorescence emission (λ_{max} = 523 nm) is spectrally separable from that of Cy3 and Cy5. In this experiment, $30SIC_{+IF3}$ complexes will be assembled using both (Atto488)-IF3 and (Cy3)-IF2, and (Cy5)-L11 50S subunits will be stopped-flow delivered to surface-immobilized complexes as before. In addition to the green, 532 nm laser used to excite Cy3 for a typical smFRET experiment, a blue, 488 nm laser will be used to directly excite Atto488. Fluorescence from all three dyes can be spectrally separated and imaged onto three independent quadrants of the EMCCD detector. Given an appropriately low spot density within the field-of-view, co-localization of (Cy3)-IF2 and (Atto488)-IF3 should indicate the presence of both initiation factors on the same 30S IC. By directly probing the presence of (Atto488)-IF3, it should be possible to determine conclusively whether IF3 remains bound during the FRET events corresponding to formation of short- and intermediate-lifetime 70S ICs. It will probably be necessary to shutter the 488 nm laser at regular intervals in order to compensate for the relatively fast photobleaching rates of green-fluorescing dyes such as Atto488, so that IF3's presence or absence can be reliably determined over the course of multiple subunit docking events. If, as is suspected, IF3 remains bound to the 70S IC, downstream events, such as translocation of $tRNA^{fMet}$ from the P to the E site during the first elongation cycle, may be required for its ejection. As a first step toward implementing these proposed three-color experiments, an (Atto488)-IF3 construct has been generated and purified by Dr. Margaret Elvekrog in our laboratory.

3.6.2 Regulation of subunit joining dynamics by TIR elements, 30S IC composition, and antibiotics

The notion of IF3 being bound within the 70S IC may seem at odds with the factor's known anti-subunit association properties. My data, however, offer an explanation for this apparent discrepancy by suggesting a model in which the IF3-bound 30S IC can adopt two conformations, one that is compatible with simultaneous binding of the 50S subunit and one that is not. By switching between these two conformations, IF3 can either allow stable subunit association to occur or strongly impede 70S IC formation by permitting only transient sampling of the 30S IC by the 50S subunit. This behavior is manifested in the data as the partitioning of 50S docking events into two separate categories, namely short-lived and relatively long-lived FRET events. I hypothesize that the efficiency of translation initiation can be up-regulated or down-regulated at the subunit joining step by shifting the conformational equilibrium of the IF3-bound 30S IC to favor either shorter- or longer-lived subunit association events, respectively.

The efficiency of 70S IC formation on a given mRNA has been shown to depend on features within the mRNA's TIR, such as the strength of the SD–anti-SD interaction, the length of the spacer between SD and start codon, and the start codon's identity [27]. Thus, one way to test my hypothesis is to introduce alterations into the mRNA's TIR which have been shown to either promote or impede 70S IC formation. For example, an mRNA with a strong SD–anti-SD interaction was found to strongly inhibit 70S IC formation in the presence of IF3 [27]. From this, it might be predicted that by increasing the strength of the SD–anti-SD interaction, a decrease in the percentage of longer-lived subunit association events would be observed in my smFRET assay.

In addition to the mRNA's TIR elements, one could test the effect on 50S subunit joining dynamics of other components of the 30S IC. For example, since IF3 has been shown to prevent the formation of aberrant 70S ICs by antagonizing 50S subunit joining to 30S ICs lacking fMet-tRNA^{fMet} [6], one could conduct experiments using 30S ICs formed in the presence of elongator tRNAs or in the absence of tRNA altogether. Similarly, one could test the regulatory effects of IF1 by excluding it from the 30S IC. IF1 cooperates with IF3 to enhance its biochemical activities, and it might therefore be predicted that omission of IF1 would result in relaxation of the IF3-dependent inhibition of 50S subunit joining. If IF3 does in fact regulate 70S IC formation based on variations in mRNA TIR elements and tRNA identity in a way that is detectable by my smFRET assay, one could then assess to what extent IF1 contributes to this regulation.

Whereas IF1 enhances IF3 activities, the antibiotic streptomycin has been shown to suppress IF3-dependent regulation of translation initiation efficiency [27]. The streptomycin binding site encompasses interactions with 16S rRNA helices h27, h18 and h44, and r-protein S12 [43]. This antibiotic does not significantly impair 30S IC assembly, and was suggested to instead exert its effects on initiation by stabilizing a particular conformation of the 30S subunit [27]. Streptomycin would thus be predicted to counteract IF3-dependent regulatory effects in my smFRET assay, and could serve as a particularly useful tool for testing the hypothesis that partitioning between short- and long-lived 50S subunit docking events is controlled by an underlying conformational equilibrium of the IF3-bound 30S IC.

One set of conditions under which the fraction of short- and long-lived subunit docking events is altered has already been identified, namely saturating concentrations of IF3. The observed shift towards short-lived events is consistent with the results from ensemble light scattering measurements demonstrating stronger inhibition of subunit association as a function of increasing IF3 concentrations [8]. I have so far only collected smFRET data on $30SIC_{+IF3}$ at two concentrations of IF3 in solution, corresponding to the extremes of very high and very low molar excess over 30S and 50S subunits. In the future, one might consider performing an IF3 titration to better characterize the concentration dependence of subunit joining inhibition.

Finally, as discussed in Section 3.5, my results have led to the suggestion that at high molar excess of IF3, more than one IF3 molecule may bind to the 30S IC at a time. This idea could be tested in the future through fluorescence co-localization experiments in which IF3 is labeled with two different fluorophores. For example, $30SIC_{+IF3}$ complexes could be prepared and surface-immobilized in the presence of (Cy3)-labeled IF3. (Cy5)-labeled IF3 could then be delivered into the flowcell, and the co-existence of both IF3 species on the same 30S IC could be probed through simultaneous excitation with green and red lasers.

3.6.3 Incorporation of aa-tRNA into the A site and entry into elongation

To convincingly demonstrate that the changes in 50S subunit joining dynamics observed using my smFRET assay correlate with changes in the overall efficiency of translation initiation, it will be necessary to probe the final step of the initiation pathway:

incorporation of the first elongator aa-tRNA into the ribosomal A site. Thus, surface-based methods to monitor binding of fluorescently labeled EF-Tu:GTP:aa-tRNA ternary complex to the ribosome are currently being developed; preliminary work toward this end is described in Chapter 4. In one approach, the EF-Tu–catalyzed incorporation of Cy5-labeled Phe-tRNA^{Phe} into the A site will be followed through the appearance of Cy5 spots under direct red laser illumination. One might predict that accumulation of Cy5 spots within the field-of-view will occur more quickly under conditions in which subunit docking events have a higher probability of resulting in formation of stable 70S ICs. These experiments might also shed light on whether the unstable 70S ICs formed almost exclusively at high IF3 concentrations are competent to bind and incorporate the ternary complex, or alternatively, whether their metastable nature indicates that they have failed to undergo a conformational change necessary for entry into elongation. In a second approach, a three-color experiment will be implemented in which (Atto488)-labeled ternary complex is co-delivered with (Cy5)-L11 50S subunits to 30S ICs bearing (Cy3)-IF2, under green and blue laser excitation. In this set-up, subunit joining should be signaled by the onset of FRET and subsequent incorporation of aa-tRNA should be indicated by the appearance of co-localized spots of Atto488 fluorescence. The three-color experiment should eventually allow us to answer such questions as: Are the short-, intermediate-, and long-lifetime 70S ICs all capable of incorporating the ternary complex? Does the ternary complex exhibit a preference toward binding to 70S ICs in the ~0.6 versus the 0.8 FRET conformational state?

3.7 References

1. 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.

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

3. 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.

4. 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.

5. 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.

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

7. 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.

8. 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.

9. 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.

10. 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.

11. Elvekrog, M.M., *The Role of Initiation Factor Dynamics in Translation Initiation.* 2011, Columbia University.

12. 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.

13. 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.

14. 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.

15. 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.

16. Berg, J.M., J.L. Tymoczko, and L. Stryer, *Enzymes: Basic Concepts and Kinetics*, in *Biochemistry*. 2002, W. H. Freeman and Company. p. 189-225.

17. Cantor, C.R. and P.R. Schimmel, *Kinetics of Ligand Interactions*, in *Biophysical Chemistry*. 1980, W. H. Freeman and Company. p. 887-938.

18. Merryman, C., D. Moazed, J. McWhirter, and H.F. Noller, *Nucleotides in 16S rRNA protected by the association of 30S and 50S ribosomal subunits.* J Mol Biol, 1999. **285**(1): p. 97-105.

19. Rodnina, M.V., R. Fricke, L. Kuhn, and W. Wintermeyer, *Codon-dependent conformational change of elongation factor Tu preceding GTP hydrolysis on the ribosome.* Embo J, 1995. **14**(11): p. 2613-9.

20. 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.

21. 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.

22. Butler, J.S., M. Springer, and M. Grunberg-Manago, *AUU-to-AUG mutation in the initiator codon of the translation initiation factor IF3 abolishes translational autocontrol of its own gene (infC) in vivo.* Proc Natl Acad Sci U S A, 1987. **84**(12): p. 4022-5.

23. 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.

24. 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.

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

26. *R: A language and environment for statistical computing*. 2011, R Development Core Team.

27. 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.

28. 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.

29. 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.

30. 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.

31. 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.

32. 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.

33. 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.

34. Cooperman, B.S., J. Dondon, J. Finelli, M. Grunberg-Manago, and A.M. Michelson, *Photosensitized cross-linking of IF-3 to Escherichia coli 30 S subunits.* FEBS Lett, 1977. **76**(1): p. 59-63.

35. 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.

36. 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.

37. 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.

38. Myasnikov, A.G., A. Simonetti, S. Marzi, and B.P. Klaholz, *Structure-function insights into prokaryotic and eukaryotic translation initiation.* Curr Opin Struct Biol, 2009. **19**(3): p. 300-9.

39. 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.

40. 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.

41. 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.

42. 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.

43. Carter, A.P., W.M. Clemons, D.E. Brodersen, R.J. Morgan-Warren, B.T. Wimberly, and V. Ramakrishnan, *Functional insights from the structure of the 30S ribosomal subunit and its interactions with antibiotics.* Nature, 2000. **407**(6802): p. 340-8.

Chapter 4

Transition from initiation to elongation: timing of ternary complex binding to the 70S IC

4.1 Introduction

Following initiation-factor mediated assembly of a 70S IC, entry into the elongation phase of protein synthesis is signaled by accommodation of an aa-tRNA into the ribosomal A site and formation of the first peptide bond. The first elongator aa-tRNA is delivered to the ribosome by EF-Tu in ternary complex with GTP. The EF-Tu:GTP:aa-tRNA ternary complex binds to the ribosome's GAC at a location that is at least partially overlapping with the binding site for the other translation factors, including IF2 [1-3]. A question of key mechanistic importance for understanding how the ribosome coordinates the transition from initiation into elongation thus concerns the relative timing of ternary complex binding with respect to IF2 dissociation from the 70S IC. This is associated with the more general problem of how the ribosome coordinates the sequential binding of translation factors to the GAC throughout protein synthesis in order to efficiently incorporate the appropriate translation factor at the correct time and to avoid a molecular "traffic jam" on the ribosome's surface [4].

Due to the overlapping nature of the binding sites for IF2 and ternary complex, it is generally assumed that binding of the ternary complex will be inhibited until IF2 is released from the ribosome [5, 6]. Dissociation of IF2 would remove the steric block towards ternary complex binding and thus allow unimpeded interactions between ternary complex and components of the GAC, and accommodation of aa-tRNA into the A site.

Indeed, formation of the initiation dipeptide is slowed dramatically when IF2 is bound to the ribosome in the presence of the non-hydrolyzable GTP analog GDPNP, and under these conditions, the rate of IF2 dissociation was found to be rate-limiting for dipeptide formation [7]. This observation implies that full accommodation of aa-tRNA into the A site and subsequent peptide bond formation can only occur upon IF2 release, but does not necessarily mean that IF2 and ternary complex cannot simultaneously be bound to the ribosome. In fact, my smFRET results indicated that IF2 dissociation from the 70S IC is modestly accelerated in the presence of ternary complex, which suggested that the ternary complex may be able to interact productively with an IF2-bound 70S IC (see Section 2.6.3). Incorporation of ternary complex into the ribosome proceeds in a series of steps [8, 9], and it is possible that one or more early steps in this pathway are permitted even in the presence of bound IF2.

I was therefore interested in exploring in greater detail the possibility that the ribosome might be capable of simultaneously accommodating both IF2 and the ternary complex. In this chapter, I report the development of surface-based single molecule approaches to directly probe the relative timing of ternary complex binding with respect to IF2 release during the transition from initiation into elongation. One approach involves fluorescence co-localization of IF2 and ternary complex, each labeled with a different colored fluorophore (Section 4.2). In a second approach, my standard 50S subunit joining assay, based on FRET between donor-labeled IF2 and acceptor-labeled L11, is expanded to include direct detection of ternary complex labeled with a third fluorophore (Section 4.3). This three-color approach potentially provides a means to monitor IF2-catalyzed

50S subunit joining to the 30S IC, IF2 release from the 70S IC, and ternary complex binding, all within the same experiment. Importantly, the single-molecule sensitivity of these experiments should permit detection of even transient co-occupancy of IF2 and ternary complex on the same ribosome, which would likely be difficult or impossible to achieve using traditional bulk biochemical techniques.

I therefore aimed to directly test whether ternary complex can bind to the IF2-bound 70S IC, and if so, to characterize the duration of its binding events as well as the timing of its binding with respect to IF2 release. The preliminary results from these experiments suggest that ternary complex can indeed bind to a 70S IC containing IF2, and that these binding events correlate with the timing of IF2 dissociation. Thus, the data are beginning to provide important insights into the sequence of events that occurs during the final stages of translation initiation and the entry into elongation. Additionally, they highlight the capacity of the ribosome to simultaneously interact with more than one translation factor at a time, which I postulate may have important implications for understanding how the ribosome rapidly and efficiently coordinates the sequential shuttling of translation factors into and out of the GAC throughout the entire protein synthesis cycle. I discuss in structural terms how the ribosome might accommodate multiple translation factors simultaneously, and how dynamics of structural components within the ribosome's GAC might be involved in regulating translation factor binding and dissociation events.

4.2 Two-color fluorescence co-localization experiments to investigate the timing of ternary complex binding to the 70S IC with respect to IF2 release

Based on the observation that the presence of ternary complex causes a modest acceleration of IF2 release from the 70S IC following 50S subunit joining, I reasoned that binding of ternary complex might occur prior to IF2 dissociation. The idea that ternary complex may be able to bind ribosomes still containing IF2 has been proposed previously by Ehrenberg and coworkers [7]. I sought to test this idea directly by labeling IF2 and the ternary complex with different color fluorophores and monitoring their presence or absence on single, surface-immobilized ribosomes during the final stages of translation initiation using multiwavelength fluorescence microscopy. The efficacy of single-molecule, multiwavelength fluorescence co-localization for monitoring the order and timing of macromolecular binding and dissociation events has recently been demonstrated by Hoskins, et al., who used this approach to study the assembly of spliceosomal subcomplexes onto a surface-immobilized pre-mRNA substrate [10].

The experimental set-up employed to study the timing of ternary complex binding and IF2 release was similar to that used to monitor IF2-catalyzed subunit joining (Chapters 2 and 3; Figure 4.1). 30S ICs containing IF1, (Cy3)-labeled IF2-GTP, fMet-tRNA^{fMet}, and 5'-biotinylated mRNA were prepared and immobilized on the surface of the microfluidic flowcell. This was followed by stopped-flow delivery of a mixture containing unlabeled 50S subunits, IF1, GTP, and pre-formed, Cy5-labeled Phe-tRNA^{Phe} ternary complex (referred to hereafter as (Cy5)-T3), which is cognate to the UUC codon at the second position on the mRNA. The ternary complex was site-specifically labeled

with Cy5 within the elbow region of the tRNA, at position acp^{3}U47 [11]. Fluorescence from (Cy3)-IF2 and (Cy5)-T3 was directly excited by simultaneous illumination with 532 nm and 635 nm lasers, respectively. Fluorescence emission from the two fluorophores was separated using dual-view optics and imaged onto separate halves of the EMCCD camera detector, and fluorescence images were collected as a function of time at a frame rate of 10 sec^{-1}. When a fluorophore-labeled factor binds to a surface-immobilized ribosome, it will be confined within the evanescent field and generate a spatially localized fluorescent spot. Conversely, dissociation of the fluorophore-labeled factor from the ribosome will lead to its diffusion out of the evanescent field and disappearance of the fluorescence signal.

Since (Cy3)-IF2 is bound to the surface-immobilized 30S ICs, fluorescent spots are present in the Cy3 channel from the onset of data acquisition. Following stopped-flow delivery of unlabeled 50S subunits and (Cy5)-T3, fluorescent spots begin to accumulate in the Cy5 channel as well (Figure 4.2). In order to clearly visualize discrete Cy5 fluorescent spots at the surface, it was necessary to reduce the background fluorescence within the Cy5 channel. This was accomplished by both attenuating the excitation power of the 635 nm laser and by delivering only low concentrations (~500 pM) of (Cy5)-T3 into the flowcell. Higher concentrations, in addition to decreasing the signal-to-noise ratio of the observed (Cy5)-T3 binding events, resulted in a decrease in their specificity, as judged based on the criteria described below.

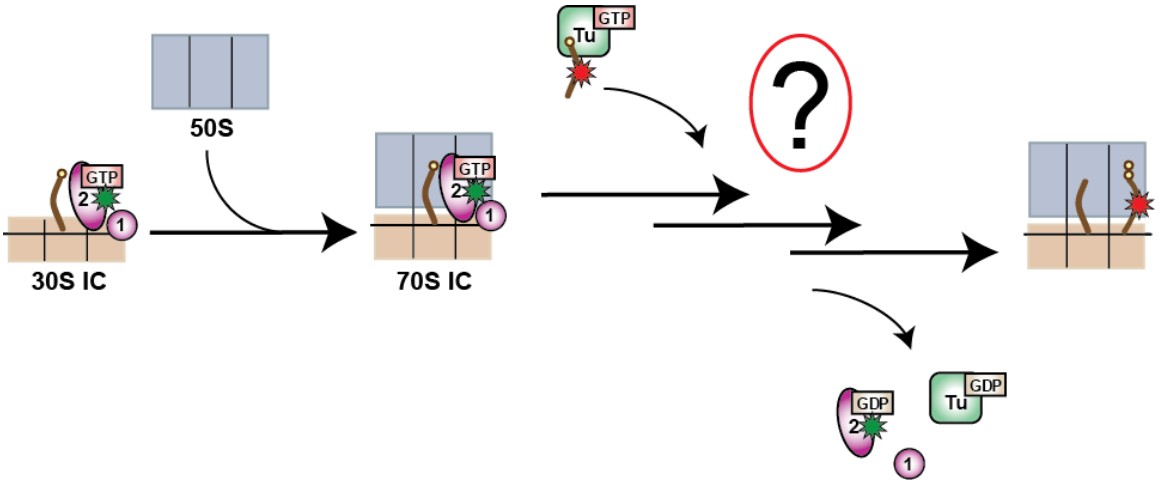

Figure 4.1: Timing of ternary complex binding to the 70S IC with respect to IF2 release. Cartoon schematic of the two-color fluorescence co-localization approach used to monitor the binding and dissociation of IF2 and ternary complex from single ribosomes. Following IF2-catalyzed formation of the 70S IC, aa-tRNA is delivered to the ribosome in ternary complex with EF-Tu and GTP. The aa-tRNA is accommodated into the ribosomal A site in a process that involves GTP hydrolysis, conformational change, and dissociation of EF-Tu. Successful accommodation of aa-tRNA into the A site results in formation of the first peptide bond catalyzed by the ribosome's peptidyl transferase center. The precise sequence of events leading from 70S IC formation to peptide bond formation is not well defined. Here, the relative timing of IF2 dissociation and ternary complex binding was investigated by labeling IF2 with Cy3 (green star) and Phe-tRNA^{Phe} with Cy5 (red star) and co-localizing Cy3 and Cy5 fluorescence to single, surface-immobilized ribosomes.

Image analysis was performed using SFTracer software, implemented in Java, which is currently being developed by Victor Naumov in our laboratory for fully automated analysis of smFRET data (Section 5.5.2). In the first step, the alignment parameters that result in the best overlay of the Cy3 and Cy5 images were determined in a process termed "mapping". This was achieved by imaging a control sample of surface-tethered, 5'-end labeled (Cy3)-DNA oligonucleotide at high laser power, conditions under which fluorescent molecules are visible in both images due to significant bleedthrough of Cy3 emission into the Cy5 channel. The alignment parameters defining the relationship between the Cy3 and Cy5 channels (translation, rotation, and skew) are

then varied to maximize the degree of overlap between the two images. Fluorescent spots in the images were identified by an algorithm that searches for 2x2 pixel regions whose mean intensity exceeds a background threshold.

The specificity of (Cy5)-T3 binding to the surface was assessed by performing several control experiments (Figure 4.2). When (Cy5)-T3 was stopped-flow delivered into a surface-passivated flowcell in the absence of surface-immobilized 30S ICs and 50S subunits, there was minimal accumulation of fluorescent Cy5 spots within the field-of-view. This demonstrates that (Cy5)-T3 does not non-specifically bind to the surface to an appreciable extent. Furthermore, significant accumulation of Cy5 spots was shown to require the presence of both 30S ICs and 50S subunits; relatively few binding events were observed when (Cy5)-T3 was delivered to surface-tethered (Cy3)-IF2–bound 30S ICs in the absence of 50S subunits, or to 30S IC-free surfaces in the presence of 50S subunits. Taken together, these results suggest that the observed binding events correspond to (Cy5)-T3 association with the 70S IC following IF2-catalyzed 50S subunit joining.

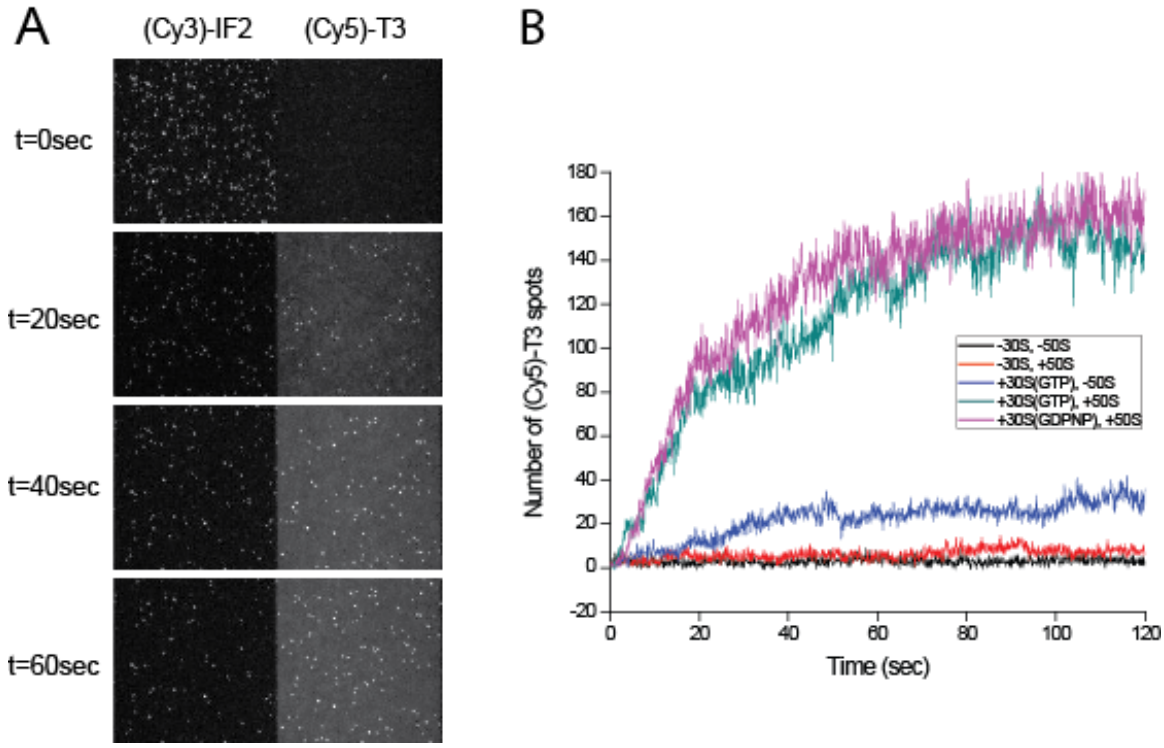

Figure 4.2: Specificity of (Cy5)-T3 binding to surface-immobilized ribosomes.
(A) Stopped-flow delivery of (Cy5)-labeled Phe-tRNA^{Phe} ternary complex and 50S subunits to surface-immobilized 30S ICs containing IF1, (Cy3)-IF2-GTP, fMet-tRNA^{fMet}, and biotin-mRNA. The flowcell was simultaneously illuminated with 532 nm and 635 nm lasers to directly excite fluorescence from Cy3 and Cy5, respectively. Sample dual-view images are shown at t=0, 20, 40, and 60 sec, demonstrating the disappearance of fluorescent spots from the Cy3 channel and appearance of fluorescent spots in the Cy5 channel as a function of time. The sample images were obtained by calculating the average intensity of each pixel over ten consecutive data frames in order to increase the image contrast for ease of viewing. Each image corresponds to half of the actual experimental field-of-view (256 x 128 pixels with 2 x 2 binning). (B) (Cy5)-T3 was delivered into flowcells with or without immobilized 30S ICs, and in the presence or absence of 50S subunits as indicated. When present, surface-immobilized 30S ICs contained IF2 in either the GTP- or GDPNP-bound form as indicated. The number of fluorescent spots within the Cy5 channel was counted for each frame of the movie using SFTracer software and plotted as a function of time. High levels of (Cy5)-T3 binding were observed only in the presence of both 30S ICs and 50S subunits.

Under these conditions, co-localized Cy3 and Cy5 fluorescent spots are interpreted to report on association of (Cy5)-T3 and (Cy3)-IF2 with the same ribosomal complex. As a precaution to ensure that this was the case, experiments were conducted at

a low surface density of (Cy3)-IF2–containing 30S ICs, which minimizes the probability of multiple 30S ICs being located within the region corresponding to a diffraction-limited (Cy3)-IF2 fluorescent spot. To confirm the low probability that co-localization of Cy3 and Cy5 fluorescence would be observed by random chance, the movies corresponding to delivery of (Cy5)-T3 to (Cy3)-IF2–containing 30S ICs in the presence and absence of 50S subunits were analyzed as follows. It was expected that a high degree of fluorescence co-localization would be observed only in the presence of 50S subunits, when 70S ICs are able to form. For the sake of this analysis, regions of interest (ROIs) were identified based on the presence of (Cy3)-IF2 fluorescent spots at the beginning of the experiment, and a (Cy5)-T3 binding event was defined as a burst of Cy5 fluorescence within the ROI in which five or more consecutive data frames exceed an intensity threshold. Using these criteria, only 5% of the ROIs in the absence of 50S subunits exhibited a (Cy5)-T3 binding event over the course of the 1200 frame movie, compared with 54% of the ROIs in the presence of 50S subunits. These results suggest a greater than 10-fold specificity and provide further evidence that the majority of (Cy5)-T3 binding events occur subsequent to 50S subunit joining and formation of the 70S IC.

While the majority (57%) of fluorescence versus time trajectories exhibited only one burst of Cy5 fluorescence over the course of the two minute long observation window, a large sub-population (43%) of trajectories exhibited multiple bursts (Figure 4.3). There are two possible explanations for the latter behavior. The first is that the disappearance and reappearance of Cy5 fluorescence represents multiple cycles of (Cy5)-T3 binding and dissociation from the 70S IC. If this were the case, it would imply that the

observed binding events correspond to relatively unstable and reversible association of (Cy5)-T3 with the ribosome, perhaps during an early stage of the multi-step ternary complex accommodation pathway [8, 9]. The estimated average lifetime of the fluorescence bursts is ~15 sec, which would correspond to a dissociation rate constant of $k_{d,\,app} \approx 0.068\ sec^{-1}$, though this value may represent an upper limit due to photobleaching of the Cy5 fluorophore. The second possible origin of the observed Cy5 intensity fluctuations is fluorophore photoblinking. Under direct 635 nm laser excitation, Cy5 is known to undergo reversible transitions into long-lived dark states with a duration of seconds to tens of seconds [12-15], and there is evidence that simultaneous excitation with 532 nm light may exacerbate these Cy5 blinking dynamics [14, 15]. Future experiments will be required to distinguish between these two possibilities. Fluorophore blinking dynamics, but not biochemical association/dissociation kinetics, are expected to be sensitive to laser excitation power. Thus, conducting experiments using different laser powers and/or alternative laser shuttering strategies may provide information on the relative contribution of fluorophore blinking to the observed Cy5 intensity fluctuations.

Regardless, inspection of the individual fluorescence versus time trajectories revealed numerous instances in which a burst of Cy5 fluorescence occurs prior to loss of the Cy3 signal (Figure 4.3). This strongly suggested that, in agreement with my hypothesis, ternary complex binding to the 70S IC can precede IF2 release, which implies that the ribosome is able to accommodate more than one translation factor at a time. For the subset of ribosomes that exhibited (Cy5)-T3 binding, the binding event occurred before (Cy3)-IF2 signal loss in 21% of the trajectories, after (Cy3)-IF2 signal loss in 78%

of the trajectories, and rarely, within the same data frame (~1% of the trajectories). In order to quantify the timing of (Cy5)-T3 binding with respect to (Cy3)-IF2 signal loss, the arrival time of the Cy5 signal was subtracted from the departure time of the Cy3 signal on a molecule-by-molecule basis. The values thus obtained were plotted as a histogram, where negative and positive values represent (Cy5)-T3 binding to the ribosome before and after (Cy3)-IF2 signal loss, respectively (Figure 4.4). Interestingly, ternary complex binding appears to occur with the highest probability shortly before or shortly after (Cy3)-IF2 signal loss. This may be indicative of a temporal correlation between ternary complex binding and IF2 release from the 70S IC. In other words, binding of ternary complex could enhance the rate of IF2 dissociation, and conversely, IF2 dissociation might facilitate more rapid binding of the ternary complex.

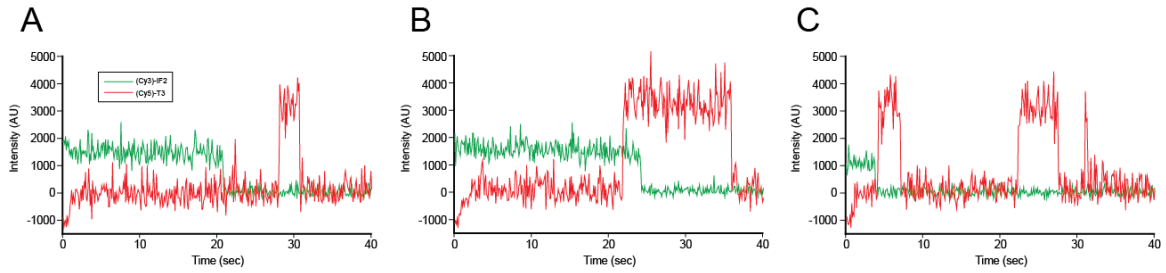

Figure 4.3: Sample fluorescence intensity versus time trajectories from co-localized (Cy3)-IF2 and (Cy5)-T3.
Fluorescence intensities from co-localized (Cy3)-IF2 (green) and (Cy5)-T3 (red) are plotted as a function of time. Loss of the (Cy3)-IF2 signal results from either Cy3 photobleaching or dissociation of (Cy3)-IF2 from the ribosome, while bursts of Cy5 fluorescence correspond to (Cy5)-T3 binding. Various types of fluorescence trajectories were observed. (Cy5)-T3 binding was observed to occur either after or before (Cy3)-IF2 signal loss (Panels A and B, respectively). Additionally, a significant subset of the trajectories (43%) exhibited multiple bursts of (Cy5)-T3 fluorescence (Panel C).

The most important observation from the data is that the 70S IC can apparently interact productively with both IF2 and the ternary complex simultaneously. Unambiguous co-occupancy of IF2 and ternary complex on the same ribosome based on temporal overlap of Cy3 and Cy5 fluorescence was only observed in roughly one-fifth of the trajectories, but this likely represents a lower limit for the fraction of ribosomes that actually bind ternary complex prior to IF2 release and/or that are capable of doing so. This is due to certain limitations inherent to the experimental set-up. Specifically, visualization of ternary complex binding to an IF2-bound ribosome would only be expected to occur with a high probability under conditions where the rate of (Cy5)-T3 binding is significantly faster than the composite rate of (Cy3)-IF2 dissociation plus Cy3 photobleaching, both of which contribute to loss of the Cy3 fluorescence signal. I therefore speculated that the reason co-residency of ternary complex and IF2 was not observed on a larger fraction of ribosomes was not because bound IF2 significantly impedes ternary complex binding, but rather that (Cy3)-IF2 dissociated or the Cy3 fluorophore photobleached faster than ternary complex could associate with the 70S IC. In support of this notion, the average lifetime of the (Cy3)-IF2 signal for the subpopulation of molecules exhibiting (Cy5)-T3 binding before (Cy3)-IF2 signal loss (39.9 sec) was ~3.2-fold longer than the corresponding lifetime for the subpopulation where (Cy5)-T3 binding was observed only after (Cy3)-IF2 signal loss (12.6 sec). This implies that the probability of observing co-localization increases as the lifetime of the (Cy3)-IF2 fluorescence signal increases.

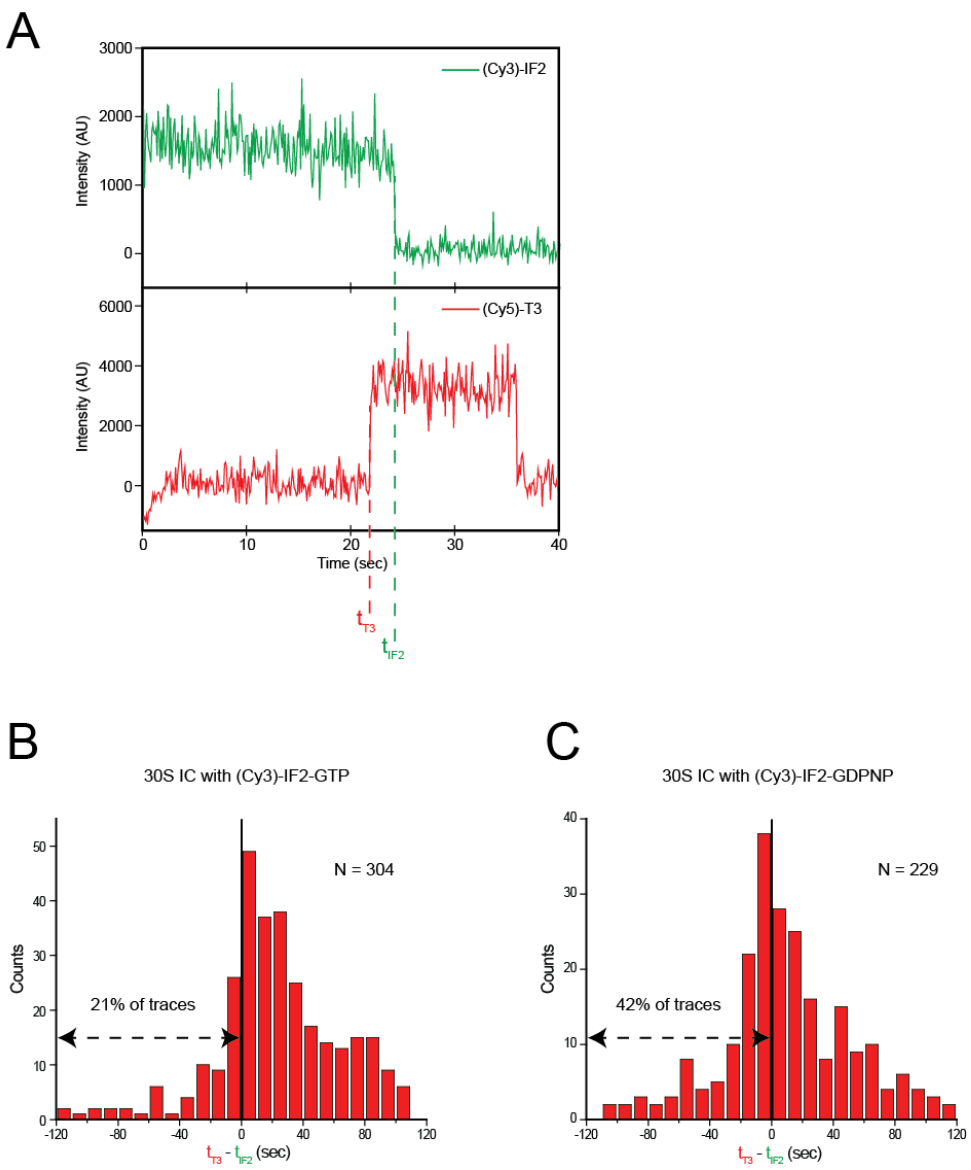

Figure 4.4: Analysis of the time difference between (Cy5)-T3 binding and (Cy3)-IF2 signal loss on single ribosomes.
(A) For each co-localized (Cy3)-IF2/(Cy5)-T3 pair, the timepoints corresponding to (Cy3)-IF2 signal loss (t_{IF2}, green dashed line) and the first instance of (Cy5)-T3 binding (t_3, red dashed lined) were identified. The time difference between these events (t_{T3} - t_{IF2}) was calculated on a trace-by-trace basis. Negative values correspond to occurrences of (Cy5)-T3 binding to the 70S IC prior to (Cy3)-IF2 release. (B, C) Histogram of the time difference (t_{T3} - t_{IF2}) for experiments performed in the presence of 30S ICs containing either (Cy3)-IF2-GTP or (Cy3)-IF2-GDPNP, respectively. The total number of single-molecule trajectories used to generate the histograms is indicated by "N".

Based on these results, I predicted that simultaneous binding of IF2 and ternary complex could be observed for a higher percentage of ribosomes under conditions in which the association rate of (Cy5)-T3 is increased, the dissociation rate of (Cy3)-IF2 is decreased, and/or the photobleaching rate of Cy3 is decreased. In theory, this prediction could be tested by delivering a higher concentration of (Cy5)-T3 into the flowcell and thereby increasing its rate of association with the ribosome. This was not possible in the context of the current experimental setup, however, as delivery of (Cy5)-T3 concentrations higher than ~500 pM led to a decrease in the binding specificity.

Instead, I sought to test my prediction by slowing the dissociation rate of (Cy3)-IF2 by substituting GTP with the non-hydrolyzable GTP analog GDPNP. The GDPNP-bound form of IF2 promotes rapid 50S subunit joining, but is subsequently stabilized on the 70S IC (see Chapter 2 and references [7, 16]). When (Cy5)-T3 and 50S subunits were co-delivered to surface-immobilized 30S ICs containing (Cy3)-IF2-GDPNP, co-localized ternary complex binding events were again observed for roughly half (48%) of the (Cy3)-IF2–bound 30S ICs. Among this subset of ribosomes, (Cy5)-T3 binding occurred before (Cy3)-IF2 signal loss in 42% of the fluorescence versus time trajectories, compared with 21% of the trajectories in the experiments with (Cy3)-IF2-GTP (Figure 4.4). This two-fold increase paralleled the increase in lifetime of the (Cy3)-IF2 signal in the presence of GDPNP versus GTP (30 sec versus 18 sec, respectively). Thus, in agreement with my prediction, slowing the rate of IF2 dissociation from the 70S IC leads to an increase in the fraction of ribosomes on which co-residency of ternary complex and IF2 is observed.

I hypothesize that if the rate of (Cy3)-IF2 signal loss were decreased even further, by, for example, shuttering the 532 nm laser at regular intervals to extend the survival time of the Cy3 fluorophore, the frequency with which co-residency of ternary complex and IF2 is observed would continue to increase. Higher concentrations of ternary complex would be expected to have the same effect by speeding up its rate of association with the ribosome. In this regard, it is interesting to note that EF-Tu is the most abundant protein in *E. coli*, with an estimated *in vivo* concentration of ~100 μM [17], and tRNA concentrations are likely on the order of ~10s of μM [18, 19]. I speculate that, at such high concentrations, formation of the 70S IC would be followed almost immediately by binding of ternary complex. In this case, binding of ternary complex prior to IF2 release may represent the predominant sequence of events occurring during the late stages of translation initiation *in vivo*.

4.3 Three-color experiments

Results from the two-color fluorescence co-localization experiments described above strongly suggested that, following 50S subunit joining to the 30S IC, ternary complex can bind to the 70S IC before IF2 is released. However, since the fluorescent probes were attached to IF2 and ternary complex, the 50S subunit joining event was not directly visualized and its occurrence could only be inferred. A three-color approach was designed in order to allow direct visualization of both IF2-catalyzed 50S subunit joining and ternary complex binding to the 70S IC within the same experiment. In this approach, subunit joining was detected based on FRET between (Cy3)-IF2 and (Cy5)-L11

reconstituted 50S subunits, while ternary complex binding was monitored by co-localization of a third fluorophore, Atto488, attached to the tRNA. Atto488 can be directly excited with a blue, 488 nm laser, and its fluorescence emission (λ_{max} = 523 nm) has sufficient spectral separation from that of Cy3 and Cy5 (λ_{max} = 570 and 670 nm, respectively) to allow for three-color imaging. Thus, $tRNA^{Phe}$ was labeled with Atto488 at position acp^3U47, charged with phenylalanine, and used to form ternary complexes. This triple-labeled translation system was shown to retain wild-type levels of activity in IF2-dependent 70S IC formation as judged by the initiation dipeptide assay (Figure 4.5).

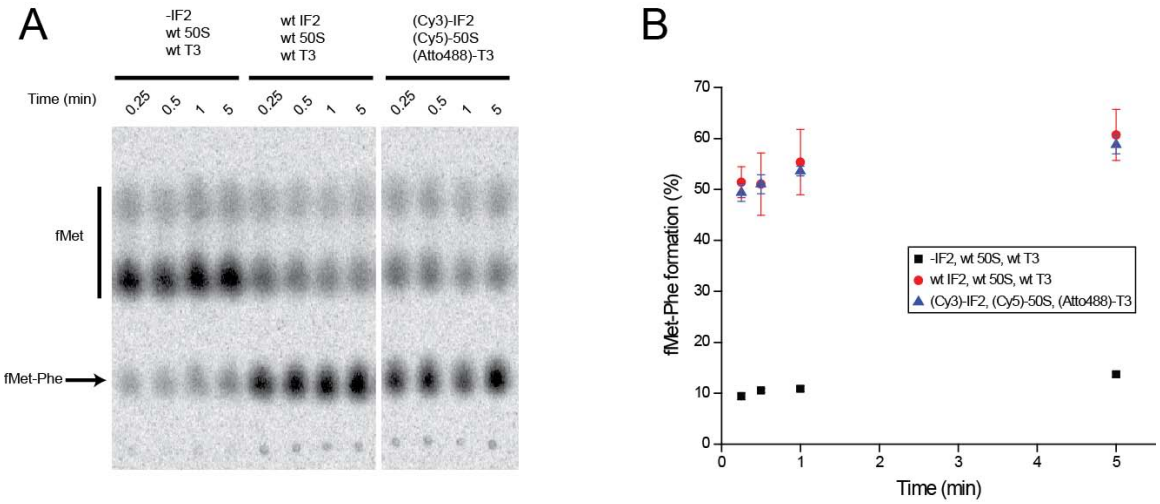

Figure 4.5: The triple-labeled translation initiation system exhibits wild-type levels of 70S IC formation and initiation dipeptide formation.
30S ICs were formed by incubating 30S subunits with IF1, mRNA, ^{35}S-fMet-$tRNA^{fMet}$, with or without IF2, in Low-Salt Tris-polymix buffer supplemented with GTP. These 30S ICs (1.5 pmol) were then mixed with 50S subunits (2.25 pmol) and preformed EF-Tu:GTP:Phe-$tRNA^{Phe}$ ternary complex (6 pmol). The reaction was incubated at room temperature and quenched with base at 15 sec, 30 sec, 1 min and 5 min time points, and formation of ^{35}S-fMet-Phe was monitored by eTLC (Panel A). Lanes 1-4: 30S ICs formed in the absence of IF2 were mixed with unlabeled, wild-type 50S subunits and ternary complex. Lanes 5-8: 30S ICs formed in the presence of IF2 were mixed with unlabeled, wild-type 50S subunits. Lanes 9-12: Triple-labeled system. 30S ICs containing (Cy3)-IF2 were mixed with (Cy5)-L11 reconstituted 50S subunits and (Atto488)-labeled ternary complex. Panel B: Results from quantification of the eTLC phosphor images. The percent of fMet converted to dipeptide was calculated by dividing intensity of the fMet-Phe spot by the sum of fMet-Phe and unreacted fMet spots, and multiplying by 100. The resulting values were plotted as a time course. Data points and error bars represent the mean and standard deviation from three independent experiments.

Single-molecule experiments were performed by stopped-flow delivering a mixture of (Cy5)-L11 50S subunits and (Atto488)-T3 to surface-immobilized 30S ICs containing IF1, (Cy3)-IF2-GTP, biotin-mRNA, and fMet-tRNA^{fMet} in the P site. Constant, simultaneous illumination with 488 nm and 532 nm lasers was used to excite fluorescence from (Atto488)-T3 and (Cy3)-IF2, respectively. Fluorescence emission from the three fluorophores was separated using a series of dichroic filters and imaged onto three separate quadrants of the EMCCD detector. The procedure for mapping the three fields-of-view was analogous to that described above for the two-color co-localization experiments. In this case, however, two different control samples were imaged: a surface-tethered (Cy3)-labeled DNA oligonucleotide excited with the 532 nm laser and an (Atto488)-labeled oligo excited with the 488 nm laser. SFTracer software was then used to calculate the alignment parameters that maximize overlap of the three images based on bleedthrough of Cy3 emission into the Cy5 channel and bleedthrough of Atto488 emission into the Cy3 channel.

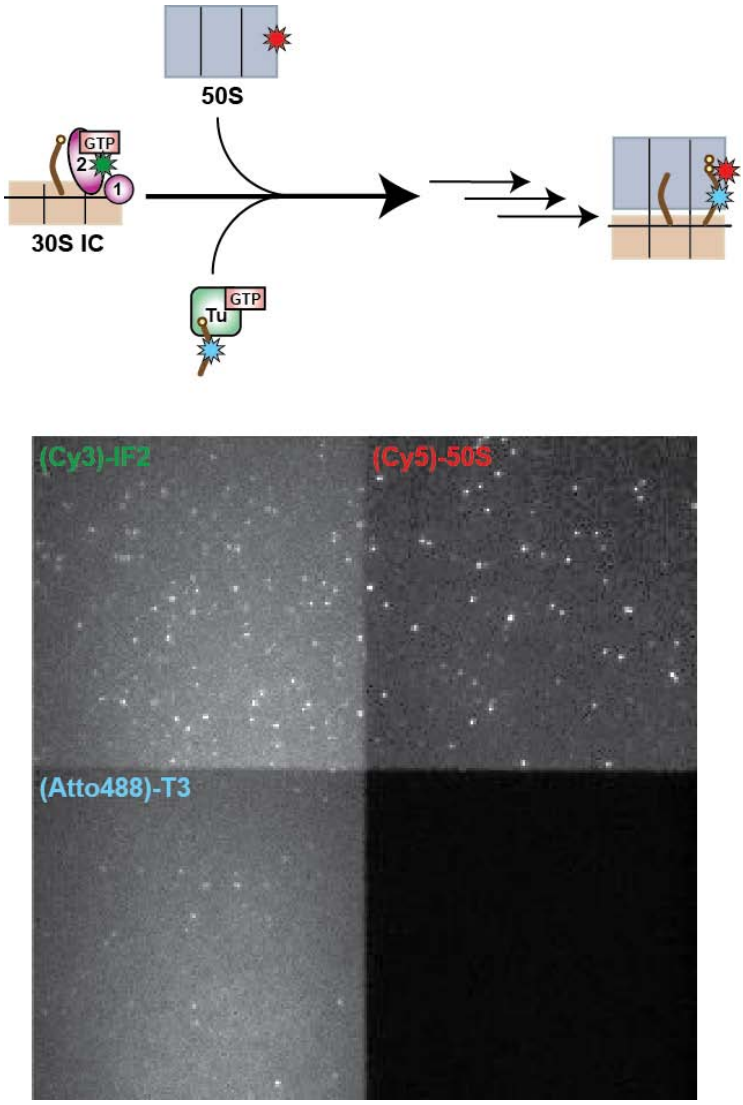

Figure 4.6: Three-color observation of IF2-catalyzed 50S subunit joining and ternary complex binding to the 70S IC.
Top row: Cartoon schematic of the experimental set-up. 30S ICs containing biotin-mRNA, fMet-tRNA^{fMet}, IF1, and (Cy3)-IF2-GTP were tethered to the surface of the flowcell. A mixture of (Cy5)-L11 reconstituted 50S subunits and (Atto488)-labeled Phe-tRNA^{Phe} ternary complex was stopped-flow delivered into the flowcell, and data was collected under dual-illumination with 488 nm and 532 nm lasers. Cy3 and Atto488 fluorescence results from direct laser excitation, whereas Cy5 fluorescence arises via FRET between Cy3 and Cy5. Fluorescence emission from the three dyes was separated using quad-view optics and imaged onto three separate quadrants of the CCD detector. Bottom row: Sample quad-view image showing discrete fluorescent spots from (Cy3)-IF2 (top left quadrant), (Cy5)-50S subunits (top right quadrant), and (Atto488)-T3 (bottom left quadrant). The image was obtained by calculating the average intensity of each pixel over ten consecutive data frames (frames 101-110, t=10-10.1 sec) of the movie. This was done in order to increase the image contrast such that fluorescent spots could be easily discerned in all three channels.

Fluorescence versus time trajectories that exhibited a decrease in Cy3 intensity and concomitant increase in Cy5 intensity—the signature for IF2-catalyzed subunit joining—were selected for further analysis. Many of these trajectories exhibited discrete bursts of co-localized Atto488 fluorescence, indicative of Phe-tRNA^{Phe} ternary complex binding to the ribosome. The bursts of Atto488 fluorescence almost always occurred after the 50S subunit joining event (~99% of the trajectories), implying that ternary complex is indeed binding to the 70S IC, rather than the 30S IC. Notably, many of the same features characterizing (Cy5)-T3 fluorescence bursts in the two-color experiment were observed for the (Atto488)-T3 fluorescence bursts in the three-color experiment. Specifically, the majority of the trajectories (~80%) exhibited (Atto488)-T3 binding subsequent to (Cy3)-IF2 signal loss, while in a smaller but significant subpopulation (~20%), ternary complex binding occurs after subunit joining but before (Cy3)-IF2 signal loss. The latter finding provides direct evidence that the ribosome is capable of recruiting ternary complex prior to IF2 release from the 70S IC. Another parallel between the two- and three-color experiments is the presence of multiple (Atto488)-T3 fluorescence bursts within individual fluorescence versus time trajectories (Figure 4.7). As before, this could theoretically arise from either a series of (Atto488)-T3 association/dissociation events or, alternatively, fluorophore photoblinking. However, the fact that very similar fluorescence behavior was observed, and in a comparable percentage of the trajectories, for two quite different fluorophores seems to suggest that its origin is more biochemical than photophysical. The fluorescence fluctuations may therefore be reporting on reversible

binding of ternary complex to the 70S IC, perhaps at an early stage of the aa-tRNA accommodation pathway.

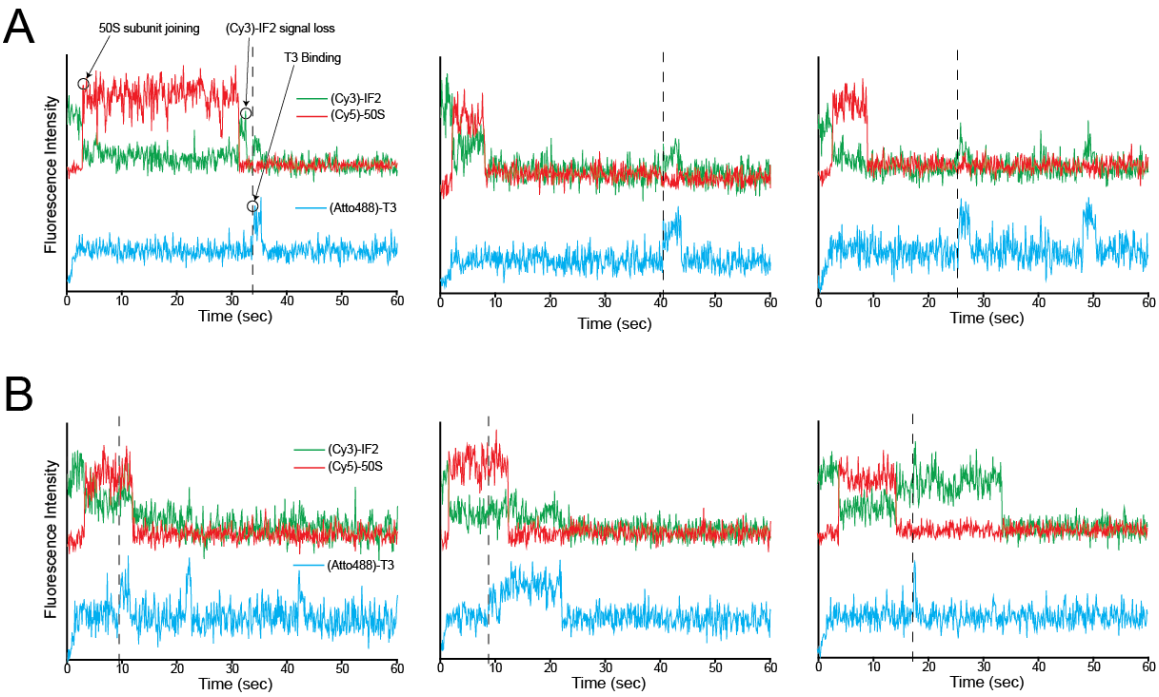

Figure 4.7: Sample three-color fluorescence intensity versus time trajectories.
Fluorescence intensities from co-localized (Cy3)-IF2 (green), (Cy5)-50S subunits (red), and (Atto488)-T3 (light blue) spots are plotted as a function of time. Cy3 and Cy5 traces are overlayed, while the Atto488 trace is shifted down on the y-axis for clearer visualization. The anticorrelated drop in Cy3 intensity and jump in Cy5 intensity at the beginning of the trace indicates the IF2-catalyzed 50S docking event. Bursts of Atto488 fluorescence indicate binding of ternary complex to the surface-tethered 70S IC. The time corresponding to the first (Atto488)-T3 binding event is indicated by a vertical dashed line. (A) Three example traces in which (Atto488)-T3 binding occurs after (Cy3)-IF2 signal loss. (B) Three example traces in which (Atto488)-T3 binding occurs before (Cy3)-IF2 signal loss, indicating co-occupancy of IF2 and ternary complex on the 70S IC.

Although the three-color approach allows direct observation of 50S subunit joining and ternary complex binding within the same experiment, it has a number of disadvantages compared with the analogous two-color approach. First, since the size of the field-of-view is reduced by half, it is more difficult to obtain a statistically significant

number of single-molecule fluorescence trajectories. Second, three-color imaging suffers from lower signal-to-noise ratios on account of the greater number of optical components placed in the emission path for wavelength separation. Third, probing ternary complex binding with Atto488 versus Cy5 is less robust, since Atto488 has significantly lower fluorophore brightness and photostability. Finally, shorter fluorophore survival times were observed for both Cy3 and Cy5 under the dual 488 nm and 532 nm illumination conditions used for the three-color experiments. This limits the effective experimental observation window and decreases the probability of detecting the overlap of IF2 and ternary complex binding to the ribosome when it occurs. In the future, it may be possible to address this problem by using an alternating laser excitation (ALEX) strategy for fluorophore excitation [20].

4.4 Mechanistic implications and future directions

The two- and three-color co-localization experiments described above have provided direct evidence that, following IF2-catalyzed 50S subunit joining, the first EF-Tu:GTP:aa-tRNA ternary complex can bind to the 70S IC prior to IF2 release. This finding adds a new level of detail to our understanding of the sequence of events occurring during the late stages of the translation initiation pathway. It suggests that models in which IF2 dissociation precedes ternary complex binding [5, 6] may amount to an oversimplification of how the transition from initiation to elongation actually proceeds on the ribosome. Taken one step further, my results might suggest that ribosome-

translation factor interactions during protein synthesis are more complex than a stepwise binding of factors to the ribosome one at a time.

Structurally, the observation of IF2 and ternary complex co-residency on the ribosome means that the GAC is able to accommodate at least two translation factors simultaneously. In other words, the ribosome's translation factor binding region must contain at least two separate binding sites. In support of this notion, the accommodation of aa-tRNA into the ribosomal A site is believed to proceed in a series of steps that involve interaction of ternary complex with different structural components of the GAC [8, 9]. In the first step, ternary complex binds to the ribosome's L7/L12 stalk, a protein-rich protuberance that extends from the body of the 50S subunit out into solution, via protein-protein interactions between L7/L12 and EF-Tu. This is followed by formation of the codon-anticodon interaction within the 30S subunit's decoding center, and subsequently, GTPase activation of EF-Tu. In the GTPase-activated state, the aa-tRNA adopts the so-called A/T hybrid configuration, and EF-Tu interacts with 50S GAC components closer to the 50S subunit core, such as L11, 23S rRNA helices H42-44, and the sarcin-ricin loop [21-23].

I propose that the L7/L12 stalk plays a major role in facilitating ternary complex binding to the ribosome prior to IF2 release. In *E. coli*, the L7/L12 stalk is composed of four copies of the highly acidic L12 protein (L7 is an N-terminally acetylated form of L12), which assemble as two dimers onto an α-helical extension of r-protein L10 [24, 25]. This L10-$(L7/L12)_4$ pentameric complex, in turn, binds to the surface of the 50S

subunit via interactions between L10 and 23S rRNA nucleotides 1030-1124, at the base of the stalk proximal to L11 (Figure 4.8) [26-28].

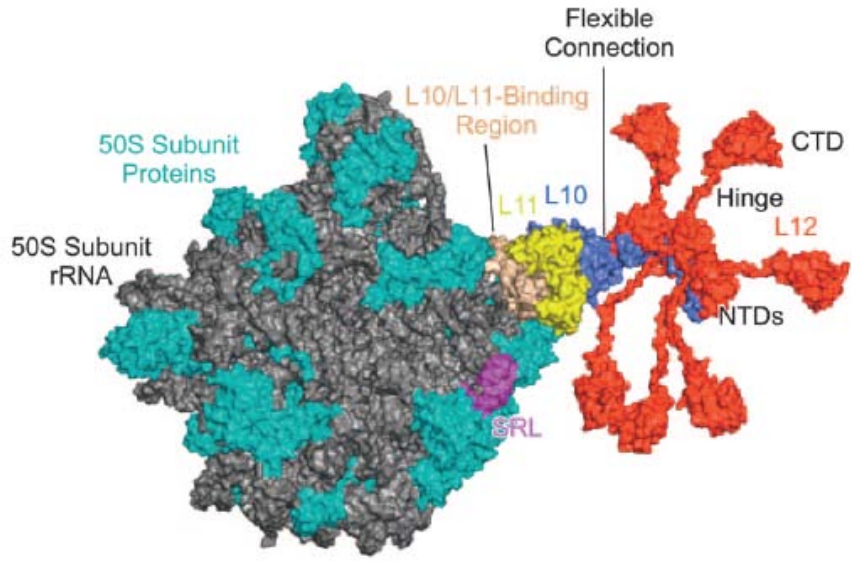

Figure 4.8: Structural model of the L7/L12 protein stalk.
A model of the complete L7/L12 protein stalk on the 50S ribosomal subunit, built from independently determined structures of the 50S ribosomal subunit and isolated stalk components. These include a refined crystal structure of the *Haloarcula marismortui* 50S subunit, crystal structures of the L10-$(L7/L12_{NTD})_6$ complex and the L11-rRNA complex from *Thermotoga maritima*, and the NMR solution structure of L12 from *E. coli*. The individual structures contain overlapping structural features, which were superimposed to build the model. The stalk is composed of three major regions: 1) the stalk base comprising the L10/L11 binding region of 23S rRNA, L11, and the L10 NTD, 2) the L10 CTD helix α8 bound to $L12_{NTD}$ dimers, and 3) the L12 CTDs. The L12 CTDs are connected to the NTDs via a flexible hinge region. They are highly mobile with respect to the ribosome and are depicted here in a random orientation. This model contains six copies of L12, as found in *T. maritima* ribosomes, though *E. coli* ribosomes contain only four copies. Figure reproduced from [28].

Each copy of L12 is comprised of an N-terminal dimerization domain connected to a globular C-terminal domain via a flexible hinge region. The highly mobile CTDs extend away from the ribosome into solution, and are thought to "catch" diffusing translation factors and deliver them to the ribosome's factor binding site [28]. The role of the L12 CTDs in promoting rapid translation factor recruitment was demonstrated by

biochemical experiments in which ribosomes reconstituted with L12 CTD-truncation mutants were found to bind EF-Tu:GTP:Phe-tRNA^{Phe} over ten times slower than wild-type ribosomes [28]. Additionally, point mutations in both EF-Tu and the L12 CTD were identified that cause a slower rate of ternary complex association with the ribosome [29]. Heteronuclear NMR spectroscopy demonstrated that isolated L12 in solution binds to EF-Tu and EF-G with sub-millimolar affinity, and the binding site was mapped to a conserved region of the L12 CTD [30].

The L12 CTDs exhibit a high degree of rotational diffusion and move rather independently of the rest of the ribosome, which suggested that in the extended conformation, they are far away from, and do not interact significantly with, the ribosomal core [31, 32]. It therefore seems possible that, by extending out into solution, one or more copies of the L12 CTD could associate with the ternary complex without generating a prohibitive steric clash between ternary complex and an IF2 molecule positioned within the ribosome's interior [16, 33]. Based on these considerations, I hypothesize that the occurrences of (Cy5)-T3 association with (Cy3)-IF2–bound 70S ICs observed in my experiments correspond to interactions of ternary complex with L12 during the first step of the aa-tRNA accommodation pathway. This hypothesis could be tested by depleting the 50S subunit of L7/L12 by NH_4Cl/ethanol treatment [34, 35], reconstituting the ribosome with L12 CTD-truncation mutants [28], or by introducing mutations into L12 or EF-Tu that affect the association rate of ternary complex [29]. One might predict that removal or truncation of L7/L12 would result in a loss of the ribosome's ability to bind ternary complex before IF2 dissociation, while the use of L12

and EF-Tu point mutants might alter the kinetics and stability of ternary complex binding to the IF2-bound 70S IC.

Less than half of the ribosomes in my single-molecule assay exhibited clear co-residency of (Cy5)-T3 and (Cy3)-IF2, but this could be due in large part to a competition between the rates of (Cy5)-T3 association and (Cy3)-IF2 dissociation/photobleaching that disfavors observation of temporally overlapping fluorescence signals in the current experimental set-up. Higher concentrations of (Cy5)-T3 would increase the association rate, which, I predict, would lead to the observation of a greater number of ternary complex association events that occur prior to (Cy3)-IF2 signal loss. The implementation of zero-mode waveguide technologies, which allow for single-molecule detection at higher, micromolar concentrations of fluorophore-labeled molecules [36, 37], should allow testing of this hypothesis.

Assuming that all ribosomes within the population are indeed capable of binding ternary complex and IF2 simultaneously, and given the very high *in vivo* concentrations of EF-Tu and tRNA, it seems likely that ternary complex binding to the 70S IC occurs immediately after 50S subunit joining, and thus precedes IF2 release, during initiation of protein synthesis in the cell. Such a sequence of events could be mechanistically important in facilitating a seamless transition from initiation into elongation. If ternary complex is already pre-bound to the ribosome, accommodation of aa-tRNA into the A site could proceed immediately after it is vacated by IF2, without requiring a “wait time” for ternary complex recruitment. Under these circumstances, the rate-limiting step for aa-tRNA accommodation and peptide bond formation would likely be dissociation of IF2.

Binding of ternary complex to L7/L12 may also play the role of accelerating IF2 release, as suggested by results from us and others (see Chapter 2 and reference [7]). While the exact mechanism underlying this acceleration is not known, the interconnectivity of r-protein and rRNA components within the 50S subunit's GAC suggests that allosteric mechanisms could play a role, perhaps by communicating ternary complex binding at L7/L12 to the stalk base via coupled conformational changes that disrupt IF2-ribosome interactions.

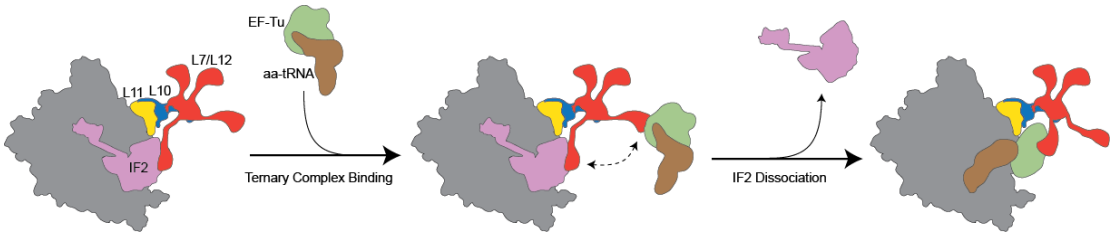

Figure 4.9: Possible mechanistic role for L7/L12 in the transition from initiation to elongation.
Cartoon schematic highlighting a potential role for the L7/L12 protein stalk in recruiting the first ternary complex to the IF2-bound 70S IC. The ribosome is shown in gray, IF2 in light violet, L11 in yellow, L10 in blue, L7/L12 in red, EF-Tu in light green, and aa-tRNA in brown. During initiation, 50S subunit joining to the 30S IC is promoted by the formation of interactions between L7/L12 and IF2-GTP [38]. Subsequent recruitment of the first ternary complex could occur prior to IF2 dissociation, through formation of an $L12_{CTD}$–EF-Tu interaction, facilitated by an arm of L12 that extends out into solution. Ternary complex pre-bound to L7/L12 would be at a high local concentration relative to the ribosomal A site and could rapidly proceed along the aa-tRNA accommodation pathway following IF2 release, thereby promoting efficient factor exchange and a seamless transition from initiation into elongation. The rightmost cartoon depicts an early state of the aa-tRNA accommodation pathway corresponding to tRNA binding in the A/T hybrid state.

Since translation factors bind to the GAC during all phases of translation, I speculate that similar mechanisms might help to regulate the efficiency of translation factor turnover throughout all of protein synthesis. For example, during each round of the elongation cycle, the correct aa-tRNA substrate must be selected based on basepairing of

the mRNA codon and tRNA anticodon at the 30S subunit's decoding center. Selection of the correct substrate from the cellular pool of non- and near-cognate aa-tRNAs likely requires many A-site sampling events for each amino acid incoporated into the nascent polypeptide. Rapid sampling of different aa-tRNA species would be facilitated if the next ternary complex in queue were pre-bound to L7/L12 and could probe the A-site codon immediately after the previous ternary complex has been ejected. In this context, it seems noteworthy that, given its sub-millimolar affinity and high *in vivo* concentrations (~10s of μM), ternary complex would likely saturate its putative L7/L12 binding sites in the cell.

After aa-tRNA accommodation and peptide bond formation, EF-G–GTP binds to the ribosome and catalyzes translocation of the mRNA-tRNA complex by precisely one codon, which is followed by another round of aa-tRNA selection. During the elongation cycle, therefore, ternary complex and EF-G repetitively and successively associate and dissociate from overlapping sites at the GAC. The efficiency of peptide chain elongation might therefore be enhanced through a mechanism in which the ability of the ribosome to bind multiple translation factors at a time allows ternary complex and EF-G to increase the rate of their own recycling. For example, the overall rate of translation could be sped up if binding of ternary complex to the post-translocation ribosome accelerated the dissociation of EF-G–GDP. Such a mechanism could provide a rationale for the observed cooperativity between the GTPase activities of EF-Tu and EF-G on empty ribosomes, whereby inclusion of EF-G decreases the apparent K_M for EF-Tu–dependent GTP hydrolysis and vice versa [39]. Single-molecule co-localization experiments utilizing fluorescently labeled EF-G and EF-Tu would provide a means to test this idea.

In summary, I propose that the observation of simultaneous binding of IF2 and ternary complex to the 70S IC is indicative of a more general capability of the ribosome to accommodate multiple translation factors at a time, and that this could have important implications for understanding how the ribosome efficiently coordinates shuttling of translation factors into and out of the A site during *in vivo* protein synthesis. I predict that the L7/L12 stalk represents the key structural component of the ribosome that enables this functionality, in effect acting as a standby binding site for translation factors prior to their interactions with ribosomal elements at the stalk base. Future efforts will thus be geared toward characterizing the role of the L7/L12 stalk in translation factor recruitment, and investigating putative L7/L12 conformational changes whose functional significance is not yet known.

4.5 References

1. Moazed, D., J.M. Robertson, and H.F. Noller, *Interaction of elongation factors EF-G and EF-Tu with a conserved loop in 23S RNA.* Nature, 1988. **334**(6180): p. 362-4.

2. La Teana, A., C.O. Gualerzi, and A.E. Dahlberg, *Initiation factor IF 2 binds to the alpha-sarcin loop and helix 89 of Escherichia coli 23S ribosomal RNA.* Rna, 2001. **7**(8): p. 1173-9.

3. Marzi, S., W. Knight, L. Brandi, E. Caserta, N. Soboleva, W.E. Hill, C.O. Gualerzi, and J.S. Lodmell, *Ribosomal localization of translation initiation factor IF2.* Rna, 2003. **9**(8): p. 958-69.

4. Brandi, L., S. Marzi, A. Fabbretti, C. Fleischer, W.E. Hill, C.O. Gualerzi, and J. Stephen Lodmell, *The translation initiation functions of IF2: targets for thiostrepton inhibition.* J Mol Biol, 2004. **335**(4): p. 881-94.

5. Luchin, S., H. Putzer, J.W. Hershey, Y. Cenatiempo, M. Grunberg-Manago, and S. Laalami, *In vitro study of two dominant inhibitory GTPase mutants of Escherichia coli translation initiation factor IF2. Direct evidence that GTP*

hydrolysis is necessary for factor recycling. J Biol Chem, 1999. **274**(10): p. 6074-9.

6. 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.

7. 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.

8. Rodnina, M.V. and W. Wintermeyer, *Ribosome fidelity: tRNA discrimination, proofreading and induced fit.* Trends Biochem Sci, 2001. **26**(2): p. 124-30.

9. Rodnina, M.V., K.B. Gromadski, U. Kothe, and H.J. Wieden, *Recognition and selection of tRNA in translation.* FEBS Lett, 2005. **579**(4): p. 938-42.

10. 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.

11. 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.

12. Rasnik, I., S.A. McKinney, and T. Ha, *Nonblinking and long-lasting single-molecule fluorescence imaging.* Nat Methods, 2006. **3**(11): p. 891-3.

13. 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.

14. 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.

15. Sabanayagam, C.R., J.S. Eid, and A. Meller, *Long time scale blinking kinetics of cyanine fluorophores conjugated to DNA and its effect on Forster resonance energy transfer.* J Chem Phys, 2005. **123**(22): p. 224708.

16. 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.

17. Pedersen, S., P.L. Bloch, S. Reeh, and F.C. Neidhardt, *Patterns of protein synthesis in E. coli: a catalog of the amount of 140 individual proteins at different growth rates.* Cell, 1978. **14**(1): p. 179-90.

18. Soutourina, O., J. Soutourina, S. Blanquet, and P. Plateau, *Formation of D-tyrosyl-tRNATyr accounts for the toxicity of D-tyrosine toward Escherichia coli.* J Biol Chem, 2004. **279**(41): p. 42560-5.

19. Ikemura, T., *Correlation between the abundance of Escherichia coli transfer RNAs and the occurrence of the respective codons in its protein genes.* J Mol Biol, 1981. **146**(1): p. 1-21.

20. Kapanidis, A.N., N.K. Lee, T.A. Laurence, S. Doose, E. Margeat, and S. Weiss, *Fluorescence-aided molecule sorting: analysis of structure and interactions by alternating-laser excitation of single molecules.* Proc Natl Acad Sci U S A, 2004. **101**(24): p. 8936-41.

21. Moazed, D. and H.F. Noller, *Interaction of tRNA with 23S rRNA in the ribosomal A, P, and E sites.* Cell, 1989. **57**(4): p. 585-97.

22. Valle, M., A. Zavialov, W. Li, S.M. Stagg, J. Sengupta, R.C. Nielsen, P. Nissen, S.C. Harvey, M. Ehrenberg, and J. Frank, *Incorporation of aminoacyl-tRNA into the ribosome as seen by cryo-electron microscopy.* Nat Struct Biol, 2003. **10**(11): p. 899-906.

23. Voorhees, R.M., T.M. Schmeing, A.C. Kelley, and V. Ramakrishnan, *The mechanism for activation of GTP hydrolysis on the ribosome.* Science, 2010. **330**(6005): p. 835-8.

24. Wahl, M.C. and W. Moller, *Structure and function of the acidic ribosomal stalk proteins.* Curr Protein Pept Sci, 2002. **3**(1): p. 93-106.

25. Chandra Sanyal, S. and A. Liljas, *The end of the beginning: structural studies of ribosomal proteins.* Curr Opin Struct Biol, 2000. **10**(6): p. 633-6.

26. Egebjerg, J., S.R. Douthwaite, A. Liljas, and R.A. Garrett, *Characterization of the binding sites of protein L11 and the L10.(L12)4 pentameric complex in the GTPase domain of 23 S ribosomal RNA from Escherichia coli.* J Mol Biol, 1990. **213**(2): p. 275-88.

27. Klein, D.J., P.B. Moore, and T.A. Steitz, *The roles of ribosomal proteins in the structure assembly, and evolution of the large ribosomal subunit.* J Mol Biol, 2004. **340**(1): p. 141-77.

28. Diaconu, M., U. Kothe, F. Schlunzen, N. Fischer, J.M. Harms, A.G. Tonevitsky, H. Stark, M.V. Rodnina, and M.C. Wahl, *Structural basis for the function of the ribosomal L7/12 stalk in factor binding and GTPase activation.* Cell, 2005. **121**(7): p. 991-1004.

29. Kothe, U., H.J. Wieden, D. Mohr, and M.V. Rodnina, *Interaction of helix D of elongation factor Tu with helices 4 and 5 of protein L7/12 on the ribosome.* J Mol Biol, 2004. **336**(5): p. 1011-21.

30. Helgstrand, M., C.S. Mandava, F.A. Mulder, A. Liljas, S. Sanyal, and M. Akke, *The ribosomal stalk binds to translation factors IF2, EF-Tu, EF-G and RF3 via a conserved region of the L12 C-terminal domain.* J Mol Biol, 2007. **365**(2): p. 468-79.

31. Christodoulou, J., G. Larsson, P. Fucini, S.R. Connell, T.A. Pertinhez, C.L. Hanson, C. Redfield, K.H. Nierhaus, C.V. Robinson, J. Schleucher, and C.M. Dobson, *Heteronuclear NMR investigations of dynamic regions of intact Escherichia coli ribosomes.* Proc Natl Acad Sci U S A, 2004. **101**(30): p. 10949-54.

32. Mulder, F.A., L. Bouakaz, A. Lundell, M. Venkataramana, A. Liljas, M. Akke, and S. Sanyal, *Conformation and dynamics of ribosomal stalk protein L12 in solution and on the ribosome.* Biochemistry, 2004. **43**(20): p. 5930-6.

33. 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.

34. Mohr, D., W. Wintermeyer, and M.V. Rodnina, *GTPase activation of elongation factors Tu and G on the ribosome.* Biochemistry, 2002. **41**(41): p. 12520-8.

35. Tokimatsu, H., W.A. Strycharz, and A.E. Dahlberg, *Gel electrophoretic studies on ribosomal proteins L7/L12 and the Escherichia coli 50 S subunit.* J Mol Biol, 1981. **152**(2): p. 397-412.

36. Levene, M.J., J. Korlach, S.W. Turner, M. Foquet, H.G. Craighead, and W.W. Webb, *Zero-mode waveguides for single-molecule analysis at high concentrations.* Science, 2003. **299**(5607): p. 682-6.

37. Uemura, S., C.E. Aitken, J. Korlach, B.A. Flusberg, S.W. Turner, and J.D. Puglisi, *Real-time tRNA transit on single translating ribosomes at codon resolution.* Nature, 2010. **464**(7291): p. 1012-7.

38. Huang, C., C.S. Mandava, and S. Sanyal, *The ribosomal stalk plays a key role in IF2-mediated association of the ribosomal subunits.* J Mol Biol, 2010. **399**(1): p. 145-53.

39. Mesters, J.R., A.P. Potapov, J.M. de Graaf, and B. Kraal, *Synergism between the GTPase activities of EF-Tu.GTP and EF-G.GTP on empty ribosomes. Elongation factors as stimulators of the ribosomal oscillation between two conformations.* J Mol Biol, 1994. **242**(5): p. 644-54.

Chapter 5
Materials and Methods

5.1 Reagent preparation

5.1.1 tRNAs

$tRNA^{fMet}$ was purchased from MP Biomedicals and $tRNA^{Phe}$ was purchased from Sigma or Chemical Block. tRNAs were dissolved in nanopure water, aliquoted, and stored at -20°C until use in labeling or charging reactions, which are described below.

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

$tRNA^{fMet}$ was aminoacylated and formylated following procedures described previously [1, 2]. Aminoacylation and formylation efficiency of fMet-$tRNA^{fMet}$ was assessed using hydrophobic interaction chromatography (HIC) as described [1], and found to be >95%. Radiolabeled ^{35}S-fMet-$tRNA^{fMet}$ used in the eTLC dipeptide formation assays (Sections 2.4.3 and 5.2.3) was kindly provided by Dr. Michael Englander.

5.1.1.2 Fluorescent labeling and aminoacylation of $tRNA^{Phe}$

Both labeled and unlabeled $tRNA^{Phe}$ species were used in the work described in this thesis. The labeled species were generated by first fluorescent labeling and then charging with phenylalanine. $tRNA^{Phe}$ was labeled at the primary amine of the 3-(3-amino-3-carboxypropyl)-uridine residue at position 47 within the tRNA's elbow region (acp^3U47) using NHS-ester linked Cy3, Cy5, and Atto488 fluorophores. The procedure for fluorescent labeling with Cy3/Cy5 has been described previously [1], and the procedure for labeling with Atto488 is described below. Labeling efficiencies were typically ~30-40% (Figure 5.1). Aminoacylation of $tRNA^{Phe}$ was performed according to the procedure described in reference [1]. Charging efficiencies for unlabeled $tRNA^{Phe}$ were >90%, while charging efficiencies were typically slightly lower for fluorescently labeled $tRNA^{Phe}$ (~70%) (Figure 5.1).

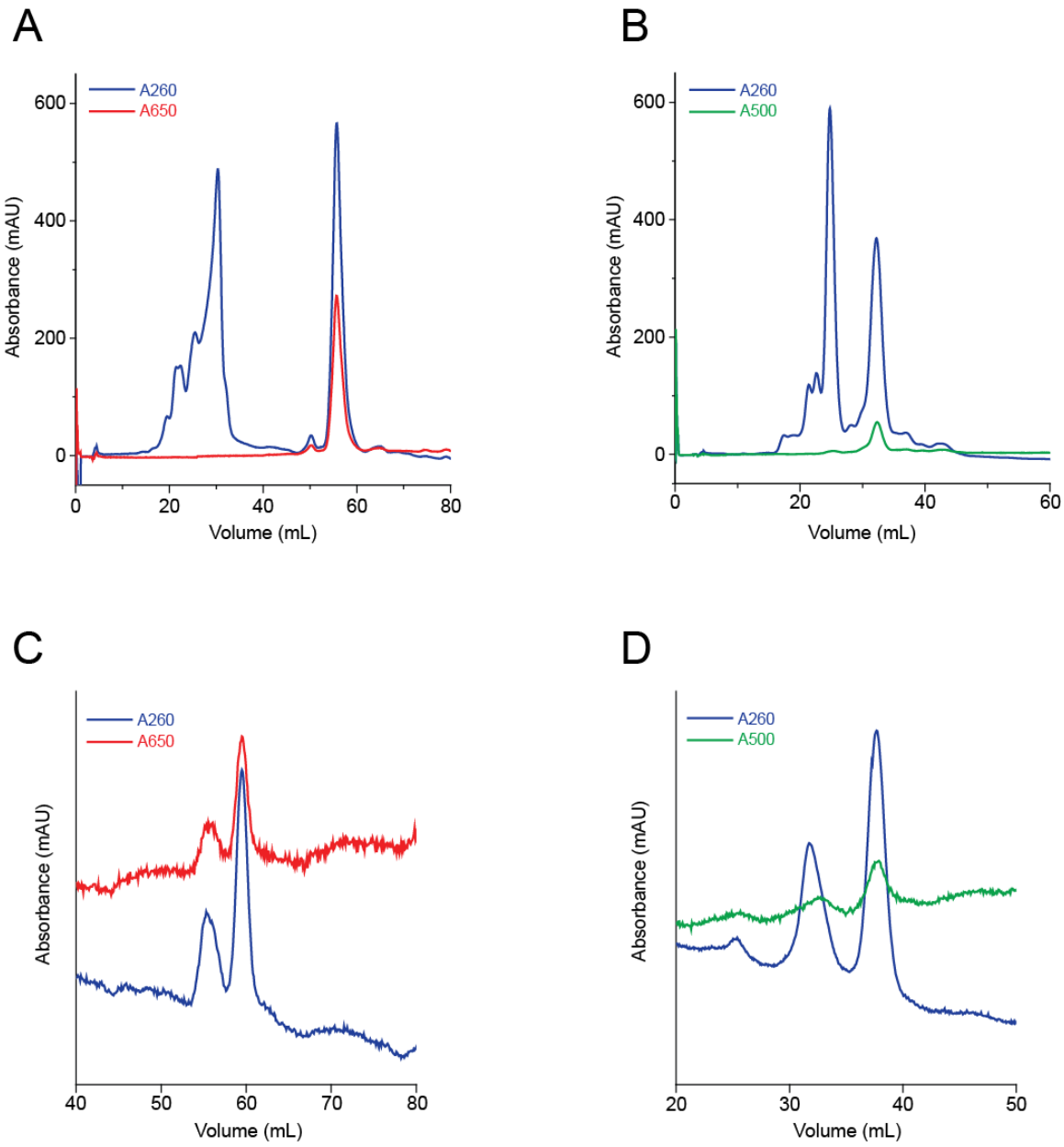

Figure 5.1: Preparation of fluorescently labeled Phe-tRNA^{Phe}
(A, B): HIC purification of tRNA^{Phe} following fluorescent labeling with Cy5 or Atto488 NHS ester, respectively. (Cy5)-tRNA^{Phe} elutes at 56 mL/ 58% Buffer B and (Atto488)-tRNA^{Phe} elutes at 32 mL/ 34% Buffer B. (C, D): Analytical HIC chromatograms used to assess charging efficiency for (Cy5)-Phe-tRNA^{Phe} and (Atto488)-Phe-tRNA^{Phe}, respectively. (Cy5)-Phe-tRNA^{Phe} elutes at 60 mL/ 62% Buffer B and (Atto488)-Phe-tRNA^{Phe} elutes at 38 mL/ 39% Buffer B.

Labeling $tRNA^{Phe}$ with Atto488:

Buffers:

HIC Buffer A: 10 mM NH_4OAc (pH = 6.3*), 1.7 M $(NH_4)_2SO_4$

HIC Buffer B: 10 mM NH_4OAc (pH = 6.3*), 10% methanol

*Note: the 1 M NH_4OAc stock solution used to make HIC Buffers A and B, rather than the final buffers, was adjusted to pH = 6.3.

Procedure:

$tRNA^{Phe}$ was labeled with Atto488 NHS ester (Sigma) using a procedure similar to that recommended by the manufacturer.

1. Dissolve 6 nmol of lyophilized $tRNA^{Phe}$ in 60 μL of Na_2CO_3 buffer (pH = 8.4).
2. Prepare dye solution by dissolving 0.2 mg dried dye pellet in 36 μL of anhydrous DMSO.
3. Combine tRNA and dye solutions and mix by pipetting up and down.
4. Incubate reaction at room temperature for two hours, with additional mixing every half hour.
5. Quench reaction with 1x volume (~10 μL) 3M NaOAc (pH = 5.2).
6. Extract six times with 1x volume (110 μL) Tris-buffered phenol. Save both aqueous and phenol phases.
7. Back-extract phenol phases with 30 μL of 0.4 M NaOAc. Combine the aqueous phase with the original aqueous phase.
8. Perform two chloroform extractions with 1x volume (140 μL).
9. Ethanol precipitate by adding 3x volumes (420 μL) -20°C ethanol, mixing thoroughly, and incubating at -80°C for at least one hour.
10. Pellet the tRNA by centrifuging for 20 min at 18,000 x g and 4°C. Carefully remove and discard supernatant.
11. Resuspend pellet in 50 μL of HIC Buffer A and inject onto the TSKgel Phenyl-5PW column (Tosoh Bioscience) pre-equilibrated with HIC Buffer A.

12. Elute the unlabeled and labeled species over a linear gradient of 0-100% HIC Buffer B. Monitor absorbance at 260 nm (tRNA) and 500 nm (Atto488). Due to the added hydrophobicity, Atto488-labeled $tRNA^{Phe}$ elutes after unlabeled $tRNA^{Phe}$.
13. Collect and pool fractions corresponding to (Atto488)-$tRNA^{Phe}$. Concentrate and buffer-exchange into nanopure water using an Amicon Ultra Centrifugal Filter device (MWCO 10,000; Millipore).
14. Measure concentration using UV-Vis ($\varepsilon_{260} \approx 760{,}000\ M^{-1}cm^{-1}$), and store at -20°C.

5.1.2 mRNAs

All mRNAs were derived from the sequence of the mRNA encoding gene product 32 of T4 bacteriophage. They were either chemically synthesized and purchased from Dharmacon, or in vitro transcribed. The sequences of all mRNAs used are provided in Appendix A.

5.1.3 Translation factors

Initiation factors IF1 and IF3 were provided by Dr. Margaret Elvekrog, and elongation factors EF-Tu and EF-Ts were provided by Dr. Michael Englander. The procedures used to prepare IF2 constructs, which were initially developed by Dr. Jiangning Wang [3], are described below.

5.1.3.1 Generation of IF2 mutants

The pProEX-HTb expression vector harboring the cloned gene for wild-type IF2 (γ-isoform) was obtained from Dr. Jiangning Wang. The pProEX-HTb vector introduces a six-histidine (6xHis) affinity tag at the N-terminus of the cloned protein, followed by a tobacco etch virus (TEV) protease cleavage site. These features allow for efficient affinity purification on Ni^{2+}-nitrilotriacetic acid (Ni^{2+}-NTA) resin followed by subsequent cleavage and removal of the affinity tag. Expression of the cloned protein is placed under control of the Trc promoter, allowing overexpression to be induced with β-D-1-thiogalactopyranoside (IPTG). Additionally, the pProEX-HTb vector harbors the

ampicillin resistance gene, which allows selection of transformants based on antibiotic resistance. The cloned IF2-γ construct contains five extra non-wild-type amino acids at the N-terminus, with an N-terminal sequence of G-A-Q-D-D-M1, where M1 is the first methionine of the wild-type sequence.

Beginning with this construct, three IF2 point mutants were generated: IF2 R561C, IF2 S566C, and IF2 S672C. These residues were chosen due to their low level of conservation among bacterial sequences, their expected surface accessibility, and, in the case of the S566C and S672C mutants, the conservative nature of a Ser-to-Cys mutation. Mutations were introduced into the vector using the QuikChange II-E Site-Directed Mutagenesis Kit (Stratagene) according to the manufacturer's protocol. The primers used for mutagenesis are shown in Table 5.1. The full sequence of all IF2 constructs was verified by DNA sequencing.

Table 5.1: Primers used for the generation of IF2 point mutants by PCR.

Primer	Sequence
R561C p1	5'-GGAGCTGAAAGCGGTATGTAAAGGTATGGCGAGCG-3'
R561C p2	5'-CGCTCGCCATACCTTTACATACCGCTTTCAGCTCC-3'
S566C p1	5'-CGTAAAGGTATGGCGTGCGGTGCGGTTATCGAATCC-3'
S566C p2	5'-GGATTCGATAACCGCACCGCACGCCATACCTTTACG-3'
S672C p1	5'-GCGCGTCAGCAGAAATGTAAACTCGAGAACATGTTCG-3'
S672C p2	5'-CGAACATGTTCTCGAGTTTACATTTCTGCTGACGCGC-3'

5.1.3.2 IF2 purification

Wild-type and mutant IF2 constructs were purified according to the same procedure, which involves a combination of Ni^{2+}-NTA affinity chromatography and cation-exchange chromatography, and is described in detail in reference [1]. Purified IF2 constructs are stored as a 50% glycerol stock in Initiation Factor Buffer (10 mM Tris-OAc ($pH_{4°C}$ = 7.5), 50 mM KCl, 10 mM $Mg(OAc)_2$, 5 mM BME) at -20°C.

5.1.3.3 Labeling of IF2 constructs with Cy3/Cy5

IF2 constructs are labeled at the mutagenized Cys residue with Cy3 or Cy5 maleimide. Wild-type IF2 contains three Cys residues, at positions 599 (domain V), 815 (domain VI-2), and 861 (domain VI-2), which are present in the IF2 mutants as well. However, these cysteines have been shown to be relatively inaccessible to fluorophore labeling, and reaction conditions have been developed in which the surface-exposed, mutagenized Cys residue is selectively labeled. These reactions conditions should be followed closely so as to avoid labeling of the wild-type cysteines, and a side-by-side labeling reaction should always be performed with wild-type IF2 to confirm that it does not get labeled.

Buffers:

Factor Labeling Buffer: 100 mM Tris-OAc (pH_{RT} = 6.9), 50 mM KCl, 300 μM *tris*(2-carboxyethyl)phosphine (TCEP)

IF2 Gel Filtration Buffer (from reference [4]) : 40 mM Tris-Cl (pH_{RT} = 6.9), 80 mM NaCl, 40 mM NH_4Cl, 5 mM $MgCl_2$, 2 mM BME

2x Initiation Factor Buffer: 20 mM Tris-OAc ($pH_{4°C}$ = 7.5), 100 mM KCl, 20 mM $Mg(OAc)_2$, 10 mM BME

Procedure:

1. Make aliquots of Cy3/Cy5 maleimide: Dissolve ~0.2 mg dye pellet in anhydrous DMSO. Take a small amount (0.5 μL) of the dissolved dye, and serially dilute 10,000-fold in nanopure water. Measure the dye concentration by UV-Vis (Cy3: ε(550) = 150,000 $M^{-1}cm^{-1}$, Cy5: ε(650) = 250,000 $M^{-1}cm^{-1}$). Make 60 nmol aliquots of the original dye solution and lyophilize. Work quickly to avoid degradation/hydrolysis of the dye. Store wrapped in foil at 4°C prior to use.
2. Equilibrate several Micro Bio-Spin P6 gel filtration columns (BioRad) with Factor Labeling Buffer. Buffer-exchange ~10 nmol IF2 into Factor Labeling Buffer. Measure IF2 concentration by UV-Vis (ε(280) = 27,450 $M^{-1}cm^{-1}$).
3. Take 6 nmol of buffer-exchanged IF2 and dilute to a final volume of 200 μL with Factor Labeling Buffer (IF2 concentration = 30 μM; TCEP concentration = 300 μM).
4. Incubate for 30 min at room temperature to reduce any disulfide bonds.

5. Dissolve 60 nmol Cy3/Cy5 dye pellet in 5 µL anhydrous DMSO. Add to the IF2 solution and mix by pipetting up and down.
6. Incubate reaction overnight (~12 hr) at 4°C.
7. Inject labeling reaction onto FPLC. Use the HiLoad Superdex 200 prep grade column (GE Healthcare) pre-equilibrated with IF2 Gel Filtration Buffer to separate IF2 from free dye. Monitor absorbance at 280 nm (IF2) and 550 or 650 nm (for Cy3 or Cy5, respectively). Labeled and unlabeled IF2 will elute as one peak well before the dye peak (Figure 5.2).

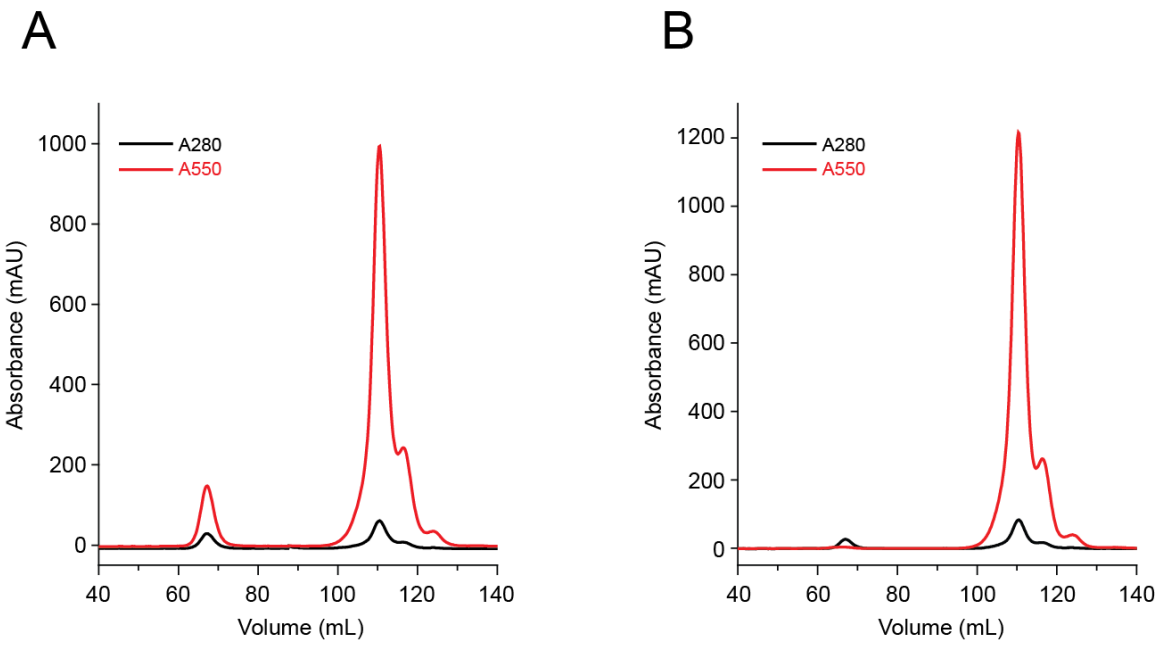

Figure 5.2: Gel filtration purification of fluorescently labeled IF2.
IF2 S672C (A) and wild-type IF2 (B) were incubated with Cy3 maleimide side-by-side under identical reaction conditions, followed by purification using gel filtration chromatography. IF2 elutes at 67 mL and free Cy3 elutes at 110 mL. Labeling was highly specific for the engineered Cys residue, as indicated by the absence of an A_{550} peak co-migrating with wild-type IF2. A zoomed-in view of the IF2 peaks is shown in Figure 2.2.

8. Collect and combine the fractions corresponding to IF2. Concentrate and buffer-exchange into 2x Initiation Factor Buffer through an Amicon Ultra Centrifugal Filter device (MWCO 10,000).

9. Dilute two-fold by adding sterile, 100% glycerol. Measure concentration of protein and dye with UV-Vis, and use these values to calculate labeling efficiency. Store (Cy3/Cy5)-labeled IF2 glycerol stocks in the dark at -20°C.

5.1.4 L11

The gene encoding r-protein L11 (*rplK*) was PCR-amplified from *E. coli* genomic DNA and cloned into the pProEX-HTb vector. The N-terminal amino acid sequence of the recombinant L11 construct is G-A-M1, where M1 is the first amino acid of the wild-type sequence. L11 contains a single wild-type Cys residue within its NTD (Cys38), which can be specifically labeled with CyDye maleimides. The protocols for cloning, purification, and fluorescent labeling of L11 are described below.

5.1.4.1 Cloning of L11

1. The *rplK* gene was PCR-amplified from *E. coli* K12 genomic DNA using the following primers:

   L11 p1: 5'-AAAAGGCGCCATGGCTAAGAAAGTACAAGCCTAT-3'

   L11 p2: 5'-AAAATCTAGATTAGTCCTCCACTACCAGGCC-3'

   L11 p1 and L11 p2 contain KasI and XBaI restriction sites, respectively, at their 5' ends, which facilitates insertion into the pProEX-HTb vector's multiple cloning site.
2. The PCR product was run on a 1% agarose gel and purified using the QIAquick Gel Extraction Kit (Qiagen).
3. The PCR product and pProEX vector were restriction-digested with KasI and XBaI (New England BioLabs, 2 Units per μg DNA) at 37°C overnight. Restriction enzymes were subsequently inactivated by heating at 65°C for 20 min.
4. Restriction-digested pProEX was treated with calf intestinal alkaline phosphatase (CIP, New England BioLabs, 1 Unit per μg DNA) in order to remove DNA 5'-phosphates and thus help prevent recircularization of the cloning vector during the subsequent ligation step.

5. Restriction-digested and CIP-treated vector and restriction-digested PCR product were purified using the QIAquick PCR purification kit (Qiagen).
6. The PCR-amplified *rplK* gene was inserted into the pProEX vector and ligated. 20 μL ligation reactions were prepared containing 50 ng of restriction-digested and CIP-treated pProEX, restriction-digested *rplK* insert at 1:3 or 1:5 vector:insert molar ratios, 1 μL T4 DNA ligase (New England BioLabs), and T4 DNA ligase reaction buffer at 1x final concentration. Ligation reactions were incubated overnight at 16°C, followed by heat-inactivation of the DNA ligase by incubation at 65°C for 10 min. The ligation reaction mixture was buffer-exchanged into nanopure water through a Micro Bio-Spin P6 column. Desalting can help to improve transformation efficiency in the subsequent step.
7. The ligation reaction products were transformed into Zapper electrocompetent cells (Novagen) by electroporation. The cells were then mixed with 1 mL SOC media and incubated at 37°C for ~1 hr with shaking, followed by dilution and plating on agarose plates supplemented with 100 μg/mL carbenicillin. Plates were incubated overnight at 37°C.
8. Single colonies were picked and used to inoculate 5 mL LB starter cultures supplemented with carbenicillin. The cultures were grown overnight at 37°C with shaking, and subsequently used to prepare 20% glycerol stocks that were flash-frozen and stored at -80°C.
9. Plasmid DNA was isolated using the QIAprep Miniprep kit (Qiagen). Presence of the insert was verified by analytical restriction digest of the plasmid with KasI and XbaI. The correct sequence of the insert was verified by DNA sequencing.

5.1.4.2 Purification of recombinant L11 under denaturing conditions

Buffers:

r-protein Buffer A: 50 mM Tris-Cl ($pH_{4°C}$ = 8.0), 5 mM $MgCl_2$, 0.1 mM phenylmethyl sulfonyl fluoride (PMSF, from a 100 mM stock solution in ethanol), 5 mM BME

r-protein Buffer B: 10 mM Tris-Cl (variable pH; $pH_{4°C}$ = 8.0, 6.8, or 5.5), 100 mM NaH_2PO_4, 6 M urea, 5 mM BME

r-protein Buffer C: 50 mM Na_2HPO_4 (pH_{RT} = 7.0), 100 mM NaCl, 2 mM BME

Procedure:

1. Transform BL21(DE3) cells with the plasmid containing cloned *rplK* from above. Plate on agarose plates supplemented with 100 μg/mL carbenicillin. Incubate overnight at 37°C.
2. Pick single colonies and use to inoculate 5 mL starter cultures in TB media supplemented with carbenicillin. Grow overnight at 37°C with shaking.
3. Use starter cultures to inoculate a 1L TB culture supplemented with carbenicillin.
4. Grow at 37°C with shaking. When cells reach OD_{600} = 0.8, add 1 mL of 1 M IPTG dissolved in water to induce overexpression. Grow cells for an additional four hours.
5. Harvest cells by centrifuging for 20 min at 4,000 rpm and 4°C. Freeze in liquid nitrogen and store at -20°C overnight.
6. Thaw cell pellet and resuspend in ~30-40 mL r-protein Buffer A. Note: This and all subsequent steps should be performed either on ice or in the cold room at 4°C.
7. Lyse cells by passing through French press 3x at a gauge pressure of 1,100 psi.
8. Clear lysate by centrifugation at 7,500 rpm for 45 min at 4°C in JA17 rotor.
9. Discard supernatant, since overexpressed L11 is found predominately in the cell pellet. Scrape pellet out of the tube and mix with 35 mL of r-protein Buffer B ($pH_{4°C}$ = 8.0, supplemented with 0.1 mM PMSF) in a beaker. Break apart the pellet a little bit with a spatula. Stir overnight at 4°C to allow the pellet to dissolve completely.
10. Clear mixture again by centrifugation at 7,500 rpm for 30 min at 4°C in JA17 rotor.
11. Equilibrate Ni^{2+}-NTA resin (~2-3 mL, Qiagen) with r-protein Buffer B ($pH_{4°C}$ = 8.0).
12. Mix resin with the resuspended and cleared cell pellet, transfer to a Falcon tube, and place on rocker for 1 hr to allow binding of 6xHis-tagged L11.
13. Transfer mixture to a disposable polypropylene column. Wash resin with 5 column volumes Buffer B ($pH_{4°C}$ = 8.0), then with 8 column volumes of Buffer B ($pH_{4°C}$ = 6.8). Elute protein with 5 column volumes of Buffer B ($pH_{4°C}$ = 5.5).

14. Pool elution fractions and measure protein concentration with the Bradford assay [5]. Dilute to ~0.1-0.2 mg/mL with Buffer B ($pH_{4°C}$ = 5.5).
15. Place in dialysis tubing (MWCO 8,000) and dialyze extensively against r-protein Buffer C to remove urea and renature the protein. Some protein precipitation will occur.
16. Following dialysis, concentrate to ~30 mL using an Amicon Ultra Centrifugal Filter device and place in new dialysis tubing. Save a small aliquot for gel analysis of the uncleaved L11 construct. Then, add TEV protease (.05 mg per 1 mg of L11) to affect cleavage of the 6xHis tag. Dialyze overnight against r-protein Buffer C.
17. Assess the extent of cleavage by running samples collected before and after TEV cleavage on a Tris-Tricine gel. If cleavage is <90% complete, add more TEV protease and continue dialysis for another overnight period.
18. Pre-equilibrate fresh Ni^{2+}-NTA resin (~2-3 mL) with r-protein Buffer C. Mix with the dialyzed protein sample in a Falcon tube and place on rocker for ~2 hr to allow binding of cleaved 6xHis tags and 6xHis-tagged TEV protease to the resin.
19. Transfer mixture to a disposable polypropylene column. Collect the flow-through and two washes with two column volumes Buffer C each. This solution contains cleaved L11.
20. Concentrate solution using an Amicon Ultra Centrifugal Filter device. Dilute two-fold with sterile, 100% glycerol. Measure concentration with the Bradford assay, and store glycerol stocks at -20°C until use.

5.1.4.3 Labeling of L11 with Cy3/Cy5

Buffers:

L11 Labeling Buffer: 50 mM Na_2HPO_4 (pH_{RT} = 7.0), 100 mM NaCl, 300 μM TCEP

L11 Gel Filtration Buffer: 50 mM Tris-Cl ($pH_{4°C}$ = 7.5), 100 mM KCl, 2 mM BME

Procedure:

1. Buffer-exchange ~15 nmol L11 into Labeling Buffer using an Amicon Ultra Centrifugal Filter device. Replace buffer at least 3x for complete buffer exchange.

This is necessary to remove BME from the buffer, which, if present, will quench the labeling reaction.

2. Measure concentration of buffer-exchanged L11 with the Bradford assay. Typically, ~30% of the protein is lost on the filter.
3. Dilute to ~30 μM in labeling buffer. Incubate at room temperature for 30 min to allow reduction of any disulfide bonds.
4. Dissolve 0.1 mg of Cy3/Cy5 maleimide in 5 μL anhydrous DMSO, add to L11 solution and mix by pipetting up and down. Incubate for one hour at room temperature.
5. Dissolve a second 0.1 mg aliquot of Cy3/Cy5 maleimide in 5 μL anhydrous DMSO and add to the reaction. Incubate for an additional hour at room temperature, then transfer to 4°C overnight.
6. A considerable amount of protein was found to precipitate during the labeling reaction. Centrifuge the sample at 18,000 x g for 1 min to pellet precipitate. Carefully remove supernatant, transfer to a new tube, and centrifuge again.
7. Inject the supernatant onto the FPLC. Use the HiLoad Superdex 75 prep grade column (GE Healthcare) pre-equilibrated with L11 Gel Filtration Buffer. Monitor absorbance at 280 and/or 230 nm (L11) and 550 or 650 nm (for Cy3 or Cy5, respectively). Three peaks are observed to elute from the column prior to the free dye. From left to right, these peaks correspond to aggregated (Cy3/Cy5)-L11, which elutes in the column void volume, soluble (Cy3/Cy5)-L11, and unlabeled L11 (Figure 5.3). The peak identities were confirmed by concentrating the corresponding fractions and analyzing them on a Tris-tricine gel with fluorescence scanning and Coomassie staining.
8. Collect and pool the fractions corresponding to soluble (Cy3/Cy5)-L11. Concentrate to ~200 μL using an Amicon Ultra Centrifugal Filter device. Measure protein concentration with the Bradford assay. Make 2 nmol aliquots, freeze in liquid nitrogen, and store at -80°C.

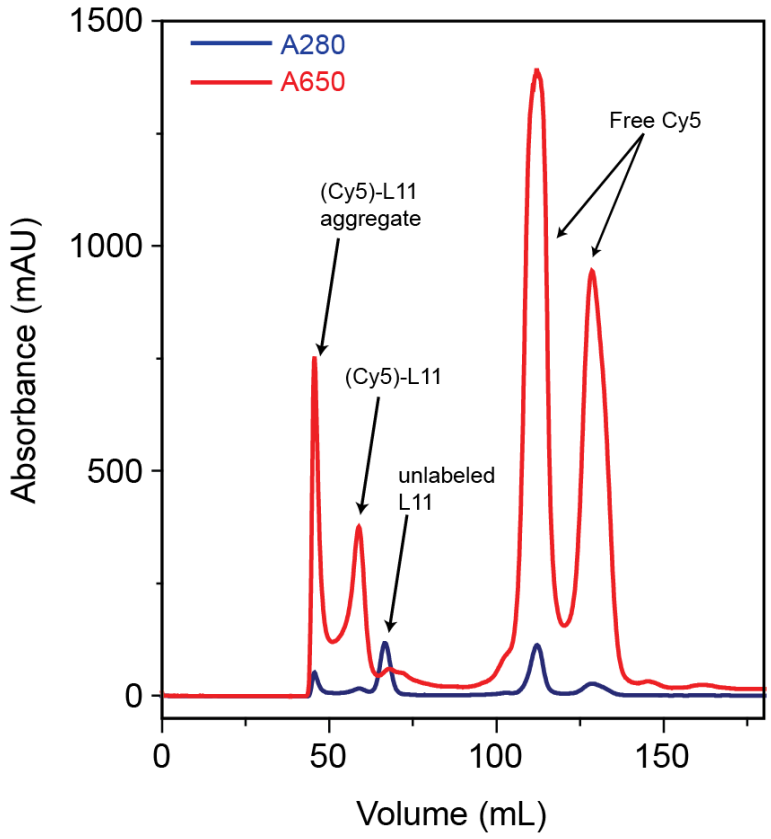

Figure 5.3: Gel filtration purification of fluorescently labeled L11.
Wild-type L11 was reacted with Cy5 maleimide followed by purification with gel filtration chromatography. Three different L11 species were separated, which are labeled in the figure. Unincorporated Cy5 elutes as two peaks.

5.1.5 Ribosomes and ribosomal subunits

Ribosomes and ribosomal subunits were purified from the wild-type strain NVD001 or the L11-knockout strain NVD005 [6, 7], which were kindly provided by the laboratory of Professor Walter Hill (University of Montana). Both strains were derived from *E. coli* K-12. NVD005 was generated by an in-frame deletion within the L11 gene (*rplK*) that removes a 249 nucleotide fragment comprising codon positions 40-122. Cells were grown in TB media supplemented with 7 μg/mL tetracycline (for NVD001) or 100 μg/mL carbenicillin (for NVD005). Ribosomes were purified using sucrose density gradient ultracentrifugation according to the protocol described in the thesis of Dr. Jingyi Fei [8]. In this procedure, tight-coupled 70S ribosomes are first purified, then split into

30S and 50S ribosomal subunits by resuspension in low, 1 mM Mg^{2+} buffer. The isolated subunits are then purified by sucrose density gradient ultracentrifugation, flash-frozen, and stored at -80°C. 50S subunits purified from strain NVD005 were reconstituted with recombinant wild-type L11 or (Cy3/Cy5)-labeled L11 according to a procedure similar to that used by Seo et al [9], as described below.

Reconstitution of 50S subunits with (Cy3/Cy5)-L11:

Buffers:

Reconstitution Buffer ("W Buffer" from reference [9]): 50 mM Tris-Cl (variable pH; pH_{RT} = 6.9 or 7.6), 30 mM NH_4Cl, 70 mM KCl, 7 mM $MgCl_2$, 1 mM DTT

Ribosome Storage Buffer: 10 mM Tris-OAc ($pH_{4°C}$ = 7.5), 60 mM NH_4Cl, 7.5 mM $Mg(OAc)_2$, 0.5 mM EDTA, 6 mM BME

Procedure:

1. Mix 1 nmol of purified NVD005 50S subunits, 2 nmol of (Cy3/Cy5)-L11, and Reconstitution Buffer (pH_{RT} = 7.6) to a final volume of 400 μL.
2. Incubate reaction for 15 min at 37°C on a heat block.
3. Load entire sample onto the top of a chilled SW28 sucrose gradient (10-40% w/v sucrose in Reconstitution Buffer (pH_{RT} = 6.9)).
4. Centrifuge in SW28 rotor for 17 hr at 22,000 rpm and 4°C. Set acceleration to "slow" and deceleration to "no brake".
5. Analyze the gradient with the following gradient analyzer settings:

   wavelength= 260 nm
   pathlength = 5 mm
   sensitivity = 2.0
   flow rate = 1.5 mL/min
   chart speed = 15 cm/h
   reference cell = air
   slit 1/8-1/4 open

6. Collect 50S peak. Transfer to Ti70.1 tube and fill with Reconstitution Buffer (pH_{RT} = 6.9, -sucrose).
7. Centrifuge in Ti70.1 rotor for 24 hr at 50,000 rpm and 4°C to pellet the reconstituted subunits. Use maximum acceleration and deceleration.
8. Remove and discard supernatant. Invert tube on clean paper towel and allow to drain for 10 min. Add ~100 μL Ribosome Storage Buffer to the ribosome pellet. Place centrifuge tube on ice for ~2-3 hr to allow pellet to dissolve.
9. Measure 50S subunit concentration with UV-Vis (1 A_{260} Unit ≈ 38 nM).
10. Make aliquots, freeze in liquid nitrogen, and store at -80°C.

5.2 Biochemical activity assays

5.2.1 Toeprinting

The toeprinting assay tests the ability of IF2 to promote selection of fMet-$tRNA^{fMet}$ over elongator $tRNA^{Phe}$ during 30S IC assembly [10, 11]. 30S ICs are formed on an mRNA (mRNA #4 in Appendix A) which has been pre-annealed to a ^{32}P-labeled DNA primer. Subsequent reverse transcription of the primer-annealed mRNA generates radiolabeled cDNA products of defined length, which can be separated with single-nucleotide resolution on a 9% denaturing PAGE gel. Reverse transcription is strongly blocked when the reverse transcriptase encounters a 30S IC bound to the mRNA, thereby generating a strong cDNA band, or "toeprint." When the 30S IC contains $tRNA^{fMet}$ bound to the AUG start codon at the P site, reverse transcription is stalled at a position 15 nucleotides downstream of the first nucleotide of the start codon, yielding a +15 toeprint. If, instead, the 30S IC contains $tRNA^{Phe}$ bound to the second, UUC codon at the P site, a +18 toeprint is generated. In the absence of initiation factors, little selectivity is shown towards initiation with fMet-$tRNA^{fMet}$ versus $tRNA^{Phe}$. The addition of IF2, however, promotes selection of fMet-$tRNA^{fMet}$ and stabilization of the resulting 30S IC, thus providing a readout for IF2 activity at the level of 30S IC assembly. The toeprinting assay was performed as previously described [1], with minor modifications.

Buffers and Reagents:

5x Polymix (-Mg^{2+}, -BME): 250 mM Tris-OAc (pH_{RT} = 7.0), 500 mM KCl, 25 mM NH_4OAc, 2.5 mM $Ca(OAc)_2$, 0.5 mM EDTA, 25 mM putrescine-HCl, 5 mM spermidine free-base, and 5% β-D-glucose.

5x Initiation Polymix: 5x Polymix (-Mg^{2+}, -BME), 15 mM $Mg(OAc)_2$, and 18.5 mM BME.

5x Sequencing Polymix: 5x Polymix (-Mg^{2+}, -BME), 50 mM $Mg(OAc)_2$, and 30 mM BME.

10x dNTP Mix: 5 mM dGTP, dTTP, dCTP, and dATP in nanopure water.

Procedure:

1. Primer-annealed mRNA was obtained from Dr. Jingyi Fei and prepared according to the protocol described in her thesis [8].
2. Initiation reactions contained 3.5 pmol 30S subunits, 1.25 pmol primer-annealed mRNA, 5 pmol fMet-$tRNA^{fMet}$, 5 pmol $tRNA^{Phe}$, and 35 pmol IF2 in 1x Initiation Polymix Buffer with 1 mM GTP. The final reaction volume was 5 μL. Reaction components were added in a series of three steps. First, 30S subunits, IF2, buffer, and GTP were mixed and incubated for 10 min at 37°C. Second, primer-annealed mRNA was added, followed by another 10 min incubation at 37°C. Third, fMet-$tRNA^{fMet}$ and $tRNA^{Phe}$ were added and the reaction was incubated for 10 more min at 37°C. Control initiation reactions were performed identically except with the omission of one or more of the components as indicated in Figure 2.4.
3. To the 5 μL initiation reaction, add 4 μL 5x Sequencing Polymix, 2 μL 10x dNTP Mix, 0.25 μL 100 mM ATP, and nanopure water to a final volume of 20 μL. Then add 0.63 μL AMV reverse transcriptase (Promega, final concentration of 0.25 U/μL), mix, and incubate for 15 min at 37°C.
4. Quench reactions with 1x volume phenol. Perform two phenol extractions and one chloroform extraction. Add 0.1x reaction volume 3 M NaOAc (pH = 5.2), then 1x volume 100% ethanol. Incubate at room temperature for 10 min to precipitate cDNA. Centrifuge at 18,000 x g for 10 min. Decant supernatant, wash pellet with 70%

ethanol, and centrifuge again at 18,000 x g for 10 min. Decant supernatant, allow pellet to air dry, then resuspend in 4 μL of denaturing PAGE loading buffer.

5. Separate cDNA fragments on a 9% denaturing PAGE gel run at 55 W constant power. Dry gel on the gel dryer for 2 hr, expose phosphorimaging screen overnight, and scan using the Storm 860 Phosphorimager (Molecular Devices).

5.2.2 GTP hydrolysis assay

The ribosome-dependent GTPase activity of IF2 is tested under multiple-turnover conditions by incubating IF2 and ribosomes with [α-^{32}P]GTP. GTP hydrolysis by IF2 results in formation of [α-^{32}P]GDP, which is separated from [α-^{32}P]GTP by thin layer chromatography (TLC). Ribosomes missing L11 exhibit a four-fold defect in promoting multiple-turnover GTP hydrolysis by IF2, and full activity can be restored by reconstitution with purified L11 [12]. Therefore, this assay allows testing of L11 reconstitution efficiency in addition to IF2's GTP hydrolysis activity. The procedure is based on that described by Brandi et al. [12], with several modifications.

Buffers and Reagents:

Hot/Cold GTP Mix: Mix 977 μL nanopure water, 20 μL of 50 mM GTP, and 3 μL of ~3.33 mM [α-^{32}P]GTP (Perkin Elmer, 3000 Ci/mmol, 10 mCi/mL). This gives final concentrations of 1 mM cold and 10 nM hot GTP. The Hot/Cold GTP Mix is aliquoted and stored at -20°C prior to use.

5x Low-Salt Polymix (-Mg^{2+}, -BME): 50 mM Tris-OAc (pH_{RT} = 7.5), 100 mM KCl, 5 mM NH_4OAc, 0.5 mM $Ca(OAc)_2$, 0.5 mM EDTA, 25 mM putrescine-HCl, 5 mM spermidine free-base, and 5% β-D-glucose.

Procedure:

1. Dilute Hot/Cold GTP Mix three-fold in nanopure water (333 μM final concentration).
2. Mix 6 pmol 70S ribosomes (or equal amounts of 30S and 50S subunits) and 12 pmol IF2 in 1x Low-Salt Polymix Buffer (15 mM Mg^{2+}, 6 mM BME) to a final volume of 13.88 μL. Control reactions are performed by omitting ribosomes, IF2, or both.
3. Add 1.12 μL of 333 μM Hot/Cold GTP Mix, and mix by pipetting up and down.

4. Incubate at room temperature. Remove 2 μL aliquots at 1, 2, 5, 10, and 20 min time points, and quench with 2 μL 100 mM EDTA (pH = 9.5).
5. Heat samples on heat block at 95°C for 1 min.
6. Centrifuge at 18,000 x g for 5 min.
7. Spot 1 μL of the supernatant onto a PEI-F cellulose TLC plate (EMD Chemicals). Wait for the spots to dry.
8. Place TLC plate in tank with 0.9 guanidine HCl mobile phase. Remove plate when solvent front is ~1" from the top of the plate.
9. Allow TLC plate to dry. Expose phosphorimager screen overnight. Scan with Storm 860 Phosphorimager. Quantify [α-^{32}P]GTP and [α-^{32}P]GDP spots using ImageQuant software (Molecular Dynamics).

5.2.3 Dipeptide formation assay

IF2-mediated formation of an elongation-competent 70S IC is tested by reaction with puromycin or Phe-tRNA^{Phe} ternary complex. Mix A (30S ICs containing ^{35}S-fMet-tRNA^{fMet}) is combined with Mix B (50S subunits and either puromycin or ternary complex). Assembly of the 70S IC and subsequent peptide bond formation generates radiolabeled ^{35}S-fMet-puromycin or ^{35}S-fMet-Phe dipeptide. Reaction products can then be separated using electrophoretic TLC (eTLC) [13] and quantified by phosphorimaging.

Buffers and Reagents:

5x Low-Salt Polymix (-Mg^{2+}, -BME): 50 mM Tris-OAc (pH_{RT} = 7.5), 100 mM KCl, 5 mM NH_4OAc, 0.5 mM $Ca(OAc)_2$, 0.5 mM EDTA, 25 mM putrescine-HCl, 5 mM spermidine free-base, and 5% β-D-glucose

Buffer 6: 50 mM Tris-OAc (pH_{RT} = 7.5), 100 mM KCl, 50 mM NH_4OAc, 0.5 mM $Ca(OAc)_2$, 0.1 mM EDTA, 5 mM $Mg(OAc)_2$, and 6 mM BME.

10x GTP Charging Mix: 10 mM GTP, 30 mM phosphoenolpyruvate, and 12.5 U/mL pyruvate kinase (Sigma), prepared in 1x Buffer 6

Procedure:

1. Preparation of 30S ICs

The procedure used to form 30S ICs for the dipeptide formation assay was the same as that used to form 30S ICs for microscope experiments (see Section 5.3.1 below), with the exception that ^{35}S-fMet-tRNA^{fMet}, rather than (Cy3/Cy5)-IF2, was limiting. 30S ICs were prepared by mixing IF1 (0.9 μM), IF2 (0.9 μM), ^{35}S-fMet-tRNA^{fMet} (0.6 μM), non-biotin mRNA (1.8 μM, mRNA #3 in Appendix A), GTP (1 mM), and 30S subunits (0.6 μM) in 1x Low-Salt Polymix Buffer (15 mM Mg^{2+}, 6 mM BME) to a final volume of 20 μL. Add all reaction components except for 30S subunits and mix by pipetting up and down. Then add 30S subunits and mix again. This procedure was chosen so as to not bias the order in which the initiation components associate with the 30S subunit during the assembly process. Incubate the reaction for 10 min at 37°C. Place tube on ice, quickly make 2.5 μL aliquots, snap freeze, and store at -80°C until use.

2. Initiation Dipeptide Reactions

1) Make 5 μL of Mix A: prepare a mixture of pre-formed 30S ICs from above (300 nM final concentration) in 1x Low-Salt Polymix (15 mM Mg^{2+}, 6 mM BME).
2) Prepare 5 μM Phe-tRNA^{Phe} ternary complex: First, mix EF-Tu, EF-Ts, and GTP Charging Mix (1x final concentration) in 1x Buffer 6 to a final volume of 10 μL (1:1 EF-Tu:EF-Ts ratio; specific concentrations should be chosen so as to yield a ~2:1 EF-Tu:Phe-tRNA^{Phe} ratio in the next step). Incubate for 1 min at 37°C, then 1 min on ice. Next, mix this EF-Tu/EF-Ts/GTP mixture with Phe-tRNA^{Phe} (final concentrations of ~11 μM EF-Tu and 5 μM Phe-tRNA^{Phe}; ~2.1-fold excess of EF-Tu). Incubate for 1 min at 37°C, then 1 min on ice. Store ternary complex on ice prior to use.
3) Make 5 μL of Mix B: mix 50S subunits (450 nM final concentration) and either puromycin (2 mM) or pre-formed Phe-tRNA^{Phe} ternary complex (1.2 μM) in 1x Low-Salt Polymix (15 mM Mg^{2+}, 6 mM BME) to a final volume of 5 μL.
4) Incubate Mix A and Mix B separately at room temperature for 5 min.

5) Combine Mix A and Mix B, and mix by pipetting up and down. Incubate the reaction at room temperature.
6) Remove 2 μL aliquots at 15 sec, 30 sec, 1 min, and 5 min time points and quench with 2 μL of 1 M KOH. Place quenched reactions on ice.

3. eTLC analysis

1) Spot 0.5 μL of the quenched reactions in a line along the center of a CCM cellulose TLC plate (EMD Chemicals). Allow spots to dry.
2) Pipet 0.5% pyridine/20% glacial acetic acid buffer onto the edges of the TLC plate above and below the line of spotted samples at the plate's center. Slowly roll the solvent from the edges towards the center of the TLC plate with a 10 mL pipet. Allow the two solvent fronts to migrate towards each other and merge at the center of the plate. Avoid rolling the pipet over the line of spotted samples in the center of the plate, as this will result in poor spot definition in the eTLC analysis.
3) Place TLC plate in the eTLC tank. Run at 1200 V for 30 min.
4) Remove plate and allow to air dry. Expose phosporimaging screen overnight, scan, and quantify ^{35}S-fMet and ^{35}S-fMet-puromycin or ^{35}S-fMet-Phe spots using ImageQuant software.

5.3 Preparation of ribosomal complexes for TIRF imaging

5.3.1 Preparation of 30S ICs

Buffer:

5x Low-Salt Polymix (-Mg^{2+}, -BME): 50 mM Tris-OAc (pH_{RT} = 7.5), 100 mM KCl, 5 mM NH_4OAc, 0.5 mM $Ca(OAc)_2$, 0.5 mM EDTA, 25 mM putrescine-HCl, 5 mM spermidine free-base, and 5% β-D-glucose.

Procedure:

30S ICs were prepared by mixing IF1 (0.9 μM), (Cy3)-IF2 (0.6 μM), IF3 (when included, 0.9 μM), fMet-$tRNA^{fMet}$ (0.6 μM), biotin mRNA (1.8 μM), GTP (1 mM), and 30S subunits (0.6 μM) in 1x Low-Salt Polymix buffer (15 mM Mg^{2+}, 6 mM BME) to a final volume of 20 μL. Add all reaction components except for 30S subunits and mix by

pipetting up and down. Then add 30S subunits and mix again. As mentioned above, this procedure was chosen to avoid biasing the order in which the initiation components associate with the 30S subunit during the assembly process. The specific order of ligand binding events occurring during 30S IC assembly, and, indeed, whether assembly occurs through one major pathway or multiple parallel pathways, is not well understood [1]. By adding 30S subunits last to the reaction mixture, the mRNA, initiation factors, and fMet-tRNA^{fMet} are free to bind in whatever their preferred order may be. Incubate the reaction for 10 min at 37°C. Place tube on ice, quickly make 1 µL aliquots, freeze with liquid nitrogen, and store at -80°C until use.

5.3.2 Preparation of 70SIC$_{GDPNP}$

Buffer:

5x Low-Salt Polymix (-Mg^{2+}, -BME): 50 mM Tris-OAc (pH_{RT} = 7.5), 100 mM KCl, 5 mM NH_4OAc, 0.5 mM $Ca(OAc)_2$, 0.5 mM EDTA, 25 mM putrescine-HCl, 5 mM spermidine free-base, and 5% β-D-glucose.

Procedure:

70S ICs with IF2 stalled in its GDPNP-bound form (70SIC$_{GDPNP}$) were prepared in two steps comprising 30S IC assembly and 50S subunit joining. The final reaction was 20 µL in 1x Low-Salt Polymix buffer (15 mM Mg^{2+}, 6 mM BME). First, prepare 30S ICs by mixing IF1 (0.9 µM), (Cy3)-IF2 (0.6 µM), IF3 (when included, 0.9 µM), fMet-tRNA^{fMet} (0.6 µM), biotin mRNA (1.8 µM), GDPNP (1 mM), and 30S subunits (0.6 µM). Add all reaction components except for 30S subunits and mix by pipetting up and down. Then add 30S subunits and mix again. Incubate for 10 min at 37°C. Next, add (Cy5)-L11 reconstituted 50S subunits (0.6 µM), mix, and incubate for another 10 min at 37°C. Make 1 µL aliquots, freeze with liquid nitrogen, and store at -80°C. 70SIC$_{GDPNP}$ complexes were diluted, surface-immobilized, and imaged without further purification.

5.3.3 Preparation of sucrose-gradient purified 70S ICs

The 70S ICs used for studying IF2's interaction with L11 during multiple-turnover GTP hydrolysis (Section 2.7) were enzymatically prepared and purified by sucrose density gradient ultracentrifugation following a procedure adapted from that described in the thesis of Dr. Jingyi Fei [8]. These 70S ICs contain biotin-mRNA and fMet-tRNA^{fMet} in the P site; they are expected to be free of initiation factors following the sucrose density gradient purification.

Buffer:

8.5x Polymix (-glucose, -Mg^{2+}, -BME): 425 mM Tris-OAc (pH_{RT} = 7.0), 850 mM KCl, 42.5 mM NH_4OAc, 4.25 mM $Ca(OAc)_2$, 0.85 mM EDTA, 42.5 mM putrescine-HCl, and 8.5 mM spermidine free-base.

Procedure:

70S ICs were formed with two separate incubations. Reported concentrations represent the final concentration of each component in the 20 μL reaction. In the first step, 30S subunits and (Cy3)-L11 reconstituted 50S subunits (1.3 μM each) were incubated in 1x Polymix buffer (5 mM Mg^{2+}, 6 mM BME) for 10 min at 37°C in the presence of IF1, IF2, and IF3 (1.7 μM each) and GTP (1 mM). In the second step, fMet-tRNA^{fMet} and biotin-mRNA were added (2 μM each), and the reaction was incubated for 20 min at 37°C. The tube was placed on ice, and the mixture was diluted five-fold by adding 80 μL of ice-cold 1x Polymix (25 mM Mg^{2+}, 6 mM BME). This mixture was loaded onto the top of a chilled SW40 sucrose gradient (10-40% w/v sucrose in 1x Polymix buffer with 20 mM Mg^{2+} and 6 mM BME). The sample was centrifuged for 15 hr at 23,000 rpm and 4°C and then analyzed with the following gradient analyzer settings:

wavelength= 260 nm
pathlength = 5 mm
sensitivity = 0.5
flow rate = 0.75 mL/min
chart speed = 15 cm/h
reference cell = air

slit 1/4 open

The 70S IC peak was collected and its concentration measured by UV-Vis (1 A_{260} Unit ≈ 20 nM). Complexes were aliquoted, flash-frozen, and stored at -80°C.

5.4 Microscope data collection procedures

All single-molecule fluorescence data were collected on a wide-field, prism-based TIRF microscope. The laser beam used for fluorophore excitation is directed through the prism onto the surface of the microscope slide. The angle of incidence is greater than the critical angle, such that the laser beam is totally internally reflected from the interface of the quartz microscope slide and aqueous buffer within the flowcell. This generates a weak, evanescent field that penetrates ~100 nm into solution and is used to selectively excite fluorescence from fluorescently labeled ribosomal complexes tethered to the surface of the flowcell. The shallow penetration depth of the evanescent field greatly reduces background fluorescence from the bulk solution and allows for high signal-to-noise single molecule fluorescence detection.

Fluorescently labeled ribosomal complexes are assembled on a 5'-biotinyated mRNA and tethered to the surface of a polyethylene glycol (PEG)-passivated flowcell through the biotin-streptavidin interaction (Figure 1.12). Protocols for the preparation of quartz microscope slides derivatized with PEG/biotin-PEG and construction of flowcells can be found in Section 4.14 of Dr. Jingyi Fei's thesis [8]. Fluorescence emission from Cy3 and Cy5 is collected with the microscope's 1.2 numerical aperture/60x magnification water-immersion objective, spectrally separated with a Dual-View imager (Photometrics), and detected with an EMCCD camera (Cascade II: 512, Photometrics). The camera is cooled to -80°C and operated at a rate of 10 frames sec^{-1} with 2x2 binning. This set-up allows the visualization of donor and acceptor fluorescence from ~300-400 individual ribosomal complexes within a 60 x 120 μm^2 observation area. In the three-color experiments, fluorescence emission from Cy3, Cy5, and Atto488 was collected and spectrally separated onto three independent quadrants of the EMCCD camera using a Quad-View imager (Photometrics).

Detailed protocols and specific considerations for the various single-molecule fluorescence experiments described in this thesis are presented below.

5.4.1 Buffers and reagents

TP 50: 10 mM Tris-OAc (pH_{RT} = 7.5) and 50 mM KCl

100 μM DNA duplex: 50-nucleotide DNA duplex (sequence: 5'-CGT TTA CAC GTG GGG TCC CAA GAC CGC GGC TAC TAG ATC ACG GCT CAG CT-3'), prepared as described in the thesis of Dr. Jingyi Fei (Section 4.14.1) [8].

Block Solution: 10 μM bovine serum albumin (BSA, Molecular Probes) and 10 μM DNA duplex in TP50.

Streptavidin/Block Solution: 1 μM streptavidin (Molecular Probes), 10 μM BSA, and 10 μM DNA duplex in TP50.

5x Low-Salt Polymix (-Mg^{2+}, -BME): 50 mM Tris-OAc (pH_{RT} = 7.5), 100 mM KCl, 5 mM NH_4OAc, 0.5 mM $Ca(OAc)_2$, 0.5 mM EDTA, 25 mM putrescine-HCl, 5 mM spermidine free-base, and 5% β-D-glucose. Store at -20°C.

1x Low-Salt Polymix (+Mg^{2+}, +BME): Dilute 5x Low-Salt Polymix (-Mg^{2+}, -BME) stock with nanopure water. Add X mM $Mg(OAc)_2$ (Mg^{2+} concentrations ranged from 3.5 to 15 mM depending on the experiment; typically 15 mM Mg^{2+} was used) and 6 mM BME. Do not freeze and thaw; prepare fresh every time.

1,000x COT/NBA: Mix 115 μL of 8.7 M COT (1,3,5,7-cyclooctatetraene, Aldrich), 119 μL of 8.4 M NBA (3-nitrobenzyl alcohol, Fluka), and 766 μL ethanol. The NBA needs to be dissolved by incubating at 37°C prior to preparation of the COT/NBA solution. Store at -20°C.

Trolox: 100 mM Trolox (6-Hydroxy-2,5,7,8-tetramethylchroman-2-carboxylic acid, 97% from Sigma) dissolved in DMSO.

80x GOD/CAT:

1. Make GOD/CAT Storage Buffer: 50 mM Tris-OAc (pH_{RT} = 7.0), 50 mM KCl, 500 mM BME, and 50% glycerol.

2. Prepare GOD and CAT stocks: Weigh Glucose Oxidase (GOD, Type VII from *Aspergillus niger*, Sigma) into a tared microcentrifuge tube. Dissolve in GOD/CAT storage buffer to a concentration of 5 U/μL and final volume of 500 μL. Weigh Catalase (CAT, from bovine liver, Sigma) into a separate tared tube. Dissolve in GOD/CAT storage buffer to a concentration of 200 U/μL and final volume of 500 μL. Do not vortex samples; mix by pipetting up and down, trying not to introduce air bubbles. Spin for 30 sec at 18,000 x g in tabletop centrifuge and incubate overnight at 4°C. Next, spin for 5 min at 18,000 x g to pellet any insoluble material. Transfer top 400 μL of solution to a new tube and store in an enzyme box at -20°C.
3. The 80x GOD/CAT solution is prepared by mixing 42.7 μL of GOD stock and 7.3 μL of CAT stock. Mix thoroughly by pipetting up and down. Spin for 30 sec at 18,000 x g, and store at -20°C.

End-labeled DNA oligo: A 15 nucleotide, single-stranded DNA oligo with a 5'-amino modification and 3'-biotin (5'-/5AmMC6/GTA AGT TTT AGG TTG/3Bio-3', purchased from IDT DNA) was labeled with Cy3 or Atto488 NHS ester. The DNA oligo was diluted to 25 μL and 0.1 mM concentration in 0.2 M Na_2CO_3 (pH = 8.4). A 0.1 mg aliquot of dye was resuspended in 20 μL anhydrous DMSO. 25 μL of the DNA solution was mixed with 15 μL of the dye solution, and the reaction was incubated for two hours at room temperature with additional mixing every 30 min. The labeling reaction was passed through two consecutive Micro Bio-Spin P6 gel filtration columns (BioRad) equilibrated with 1x TE Buffer (pH = 7.4) to remove unincorporated dye. The labeling efficiency was determined to be >95% by UV-Vis.

5.4.2 Steady-state experiments

Steady-state conditions were employed to collect smFRET data from $70SIC_{GDPNP}$ complexes (Section 2.5), and to probe IF2's interaction with the GAC during multiple-turnover GTP hydrolysis (Section 2.7).

1. Prepare a microscope slide with five microfluidic flowcells separated by double-sided tape according to the protocol described in Section 4.14 of Dr. Jingyi Fei's thesis [8].

The volume of each flowcell is ~7 μL. Samples are introduced slowly by pipet. Place a folded Kimwipe on the outlet hole to soak up buffer that flows through.

2. Wash flowcell with 200 μL TP50 Buffer
3. Deliver 20 μL Block Solution. Incubate for 5 min.
4. Deliver 20 μL Streptavidin/Block Solution. Incubate for 5 min.
5. Wash with 200 μL TP50 Buffer
6. Equilibrate flowcell with 100 μL 1x Low-Salt Polymix (+Mg^{2+}, +BME)
7. Serially dilute an aliquot of ribosomal complex to ~100-200 pM final concentration in 1x Low-Salt Polymix (+Mg^{2+}, +BME) in two steps. For $70SIC_{GDPNP}$, the buffer was supplemented with 1 mM GDPNP. The final volume of the diluted sample was 20 μL. Add 2 μL of Block Solution to the diluted sample, mix, and pipet into the flowcell. Inclusion of Block Solution helps to reduce non-specific adsorption of the complex to the surface. Incubate for 5 min to allow binding of complexes to the surface via the biotin-streptavidin interaction.
8. Wash flowcell with 100 μL 1x Low-Salt Polymix (+Mg^{2+}, +BME) supplemented with 1x GOD/CAT, 1 mM Trolox, and 1x COT/NBA. For experiments with $70SIC_{GDPNP}$, the buffer was additionally supplemented with 1 mM GDPNP. The enzymatic oxygen scavenging system (50 U/mL glucose oxidase, 365 U/mL catalase, and 1% β-D-glucose) used here helps extend the lifetime of Cy3 and Cy5 fluorophores before photobleaching [14]. Trolox, COT, and NBA are triplet-state quenchers that help suppress fluorophore blinking and photobleaching [15, 16]. Different combinations of triplet-state quenchers in the imaging buffer were tested, and the presence of all three was found to be most effective, markedly improving the lifetime before fluorophore photobleaching for experiments with the (Cy3)-IF2/(Cy5)-L11 FRET pair.
9. For experiments probing the interaction between IF2 and the GAC during multiple-turnover GTP hydrolysis, 40 μL of filling solution was introduced into the flowcell prior to imaging. The filling solution contained (Cy5)-IF2 (concentrations ranged from 2.5-40 nM depending on the experiment), 1 mM GTP, 1 mM Trolox, 1x COT/NBA, and 1x GOD/CAT prepared in 1x Low-Salt Polymix (+Mg^{2+}, +BME).

10. Image sample: collect data with 532 nm laser excitation (6 mW laser power, measured at the prism). Image acquisition is controlled by MetaMorph online software, version 6.3r7 (Molecular Devices). For $70SIC_{GDPNP}$ complexes, which contained (Cy5)-L11, one frame was collected with red, 640 nm excitation followed by stream acquisition with 532 nm excitation. This facilitates rapid identification of ribosomes containing an active Cy5 fluorophore during image analysis.

5.4.3 Real-time subunit joining experiments

1. Follow steps 1-6 from Section 5.4.2 above to prepare flowcell.
2. Serially dilute an aliquot of pre-formed 30S IC containing (Cy3)-IF2 in 1x Low-Salt Polymix (+Mg^{2+}, +BME) in two steps. $30SIC_{-IF3}$ complexes were diluted to a final concentration of ~100-200 pM; the dilution buffer was supplemented with IF1 (0.9 μM) and GTP (1 mM). $30SIC_{+IF3}$ complexes were diluted to a final concentration of ~1-2 nM; the dilution buffer was supplemented with IF1 (0.9 μM), IF3 (0.9 μM), GTP (1 mM), and, when included, fMet-$tRNA^{fMet}$ (0.9 μM). Add 2 μL of Block Solution to the diluted sample, mix, and pipet into the flowcell. Incubate for 5 min.
3. Wash flowcell with 60 μL Wash Buffer: 1x Low-Salt Polymix (+Mg^{2+}, +BME) containing IF1 (0.9 μM), IF3 (0.9 μM), fMet-$tRNA^{fMet}$ (0.9 μM), and GTP (1 mM) when required, and supplemented with 1 mM Trolox, 1x COT/NBA, and 1x GOD/CAT.
4. Fill stopped-flow tubing with Syringe Buffer: Wash Buffer containing (Cy5)-L11 reconstituted 50S subunits (concentrations ranged from 10 to 60 nM depending on the experiment; typically, 20 nM was used). When included, pre-formed Phe-$tRNA^{Phe}$ ternary complex was additionally added to a final concentration of 0.25 to 1 μM depending on the experiment. Ternary complex was prepared as described above in Section 5.2.3, with the exception that the final concentrations of EF-Tu, EF-Ts, and Phe-$tRNA^{Phe}$ in the final reaction mixture were two-fold higher (~21 μM EF-Tu and EF-Ts, 10 μM Phe-$tRNA^{Phe}$). This was done so as to minimize the amount of glycerol from the EF-Tu/EF-Ts stocks that was added into the Syringe Buffer. 60 μL of

Syringe Buffer was prepared and pipetted into the cap of an opened microcentrifuge tube. Hold stopped-flow tubing in the sample, and withdraw 50 μL into the tubing.

5. Place a drop (~1 μL) of Wash Buffer on the microscope slide flowcell's inlet hole and another drop at the end of the stopped-flow tubing to ensure a drop-to-drop connection and to minimize the chance of introducing air bubbles. Connect the stopped-flow tubing to the flowcell's inlet hole. Place a folded filter paper lightly on the outlet hole to soak up buffer that flows through.
6. Acquire data using a MetaMorph journal: Deliver 40 μL of Syringe Buffer at a flow rate of 0.39 mL/min. Delay for 1 sec, then open the green laser shutter, and stream acquire data under 532 nm laser illumination (6 mW power, measured at the prism).

5.4.4 Two-color co-localization experiments

1. Follow steps 1-6 from Section 5.4.2 above to prepare flowcell.
2. Serially dilute an aliquot of pre-formed 30S IC containing (Cy3)-IF2 to ~100 pM final concentration in 1x Low-Salt Polymix (+Mg^{2+}, +BME) supplemented with IF1 (0.9 μM) and GTP (1 mM). To 20 μL of the diluted sample, add 2 μL of Block Solution, mix, and pipet into the flowcell. Incubate for 5 min.
3. Prepare Wash Buffer: 1x Low-Salt Polymix (+Mg^{2+}, +BME) containing IF1 (0.9 μM), GTP (1 mM), Trolox (1 mM), 1x COT/NBA, and 1x GOD/CAT. Wash flowcell with 60 μL Wash Buffer.
4. Prepare 60 μL Syringe Buffer: Wash Buffer containing dark 50S subunits (20 nM), EF-Tu:GTP:Phe-(Cy5)$tRNA^{Phe}$ ternary complex (500 pM), and Block Solution (10% v/v). Ternary complex was prepared fresh before each experiment according to the procedure described in Section 5.2.3 above. The 5 μM ternary complex stock was diluted 25-fold in 1x Low-Salt Polymix (+Mg^{2+}, +BME) supplemented with 1 mM GTP before adding it to the Syringe Buffer.
5. Fill stopped-flow tubing with 50 μL Syringe Buffer.
6. Connect stopped-flow tubing to inlet hole of flowcell on the microscope stage.

7. Acquire data using a MetaMorph journal: Deliver 40 μL of Syringe Buffer at a flow rate of 0.39 mL/min. Delay for 1 sec, then open the green and red laser shutters. Stream acquire data under dual-excitation with 532 and 640 nm lasers. The laser power was adjusted by placing neutral density filters in the optical path, and was 4 mW for both the green and red laser, measured at the prism.
8. On the same day, image several fields-of-view of surface-tethered, (Cy3)-DNA oligo at high 532 nm laser power (~40 mW). These movies can be used for automated calculation of the Dual-View alignment parameters using SFTracer software based on bleedthrough of Cy3 fluorescence into the Cy5 channel (see Section 5.5.2.1 below). The (Cy3)-DNA oligo is diluted to ~10 pM prior to surface-immobilization to yield ~300-400 molecules within the field-of-view.

5.4.5 Three-color experiments

The procedure for three color-experiments was the same as that described above for the two-color co-localization experiments (5.4.4), with the following exceptions: 1) The Syringe Buffer contained EF-Tu:GTP:Phe-(Atto488)tRNA^{Phe} ternary complex instead of (Cy5)-ternary complex, and (Cy5)-50S subunits instead of dark 50S subunits. 2) Data were acquired under dual-illumination with 488 and 532 nm lasers. The 488 nm laser was operated at 30% power, and the 532 nm laser was operated at 6 mW power, measured at the prism. A Quad-View imager was used in place of the Dual-View to separate fluorescence emission from Cy3, Cy5, and Atto488. 3) On the same day, surface-tethered (Cy3)-DNA oligo was imaged at high 532 nm laser power (40 mW) to facilitate alignment of the Cy3 and Cy5 channels during image analysis. In addition, surface-tethered (Atto488)-DNA oligo was imaged at high 488 nm laser power (90% power) to facilitate alignment of the Atto488 and Cy3 channels.

5.5 Data analysis procedures

5.5.1 Analysis of smFRET data

The smFRET data analysis procedures described below make use of a variety of software programs, namely MetaMorph (Molecular Devices), Microsoft Excel, Matlab (MathWorks), vbFRET, OriginPro (OriginLabs), and "R". The Matlab scripts used for data analysis are compiled in Appendix E and the "R" scripts are compiled in Appendix F. Many of the scripts and data analysis procedures employed here were initially developed by Dr. Jingyi Fei, Mr. Pallav Kosuri, Dr. Jonathan Bronson, and Dr. Jiangning Wang. Likewise, many of the procedures outlined in this section have previously been described in the theses of Dr. Fei (reference [8], Section 4.16) and Dr. Wang (reference [3], Section 5.10). They are presented again below in order to give a complete description of the methods used to analyze my smFRET subunit joining data, and to highlight the specific additions and changes I have made to the scripts and procedures.

5.5.1.1 Generation and selection of Cy3 and Cy5 fluorescence versus time trajectories

1. Channel alignment: Open .stk image file in MetaMorph. Split the stack of images into Cy3 and Cy5 fields-of-view using "Display→ Split View" and the appropriate alignment parameters. The parameters were determined during alignment of the Dual-View optics, which was performed following the manufacturer's protocol. The Cy3 and Cy5 fields are called "W1" and "W2" by default.
2. Region Selection: Fluorescent spots are usually selected from the first 300 frames of W2. To do so, select frames 1-300 with "Stack→ Keep Planes". Create a maximum image from these frames using "Process → Stack Arithmetic → Maximum". For each pixel in the image, this operation finds the pixel which has the highest intensity value out of all of the frames, and outputs that value to a new image. Select regions from the maximum image using "Threshold Image," which highlights pixels whose intensity is above a set threshold. The threshold was chosen manually in order to selectively highlight pixels corresponding to fluorescent spots (typically two to four

contiguous bright pixels). Create regions using "Regions → Create Regions Around Objects" and transfer regions to the original W1 and W2 fields using "Regions Transfer Regions." The channel alignment is not perfect, and it will usually be necessary to manually move some of the regions within W1 one or two pixels so that they overlap with the Cy3 spots.

3. Plot the Cy3 and Cy5 intensities for each region as a function of time using "Apps→ Graph Intensities." Place cursor over the graph, right-click, and select "Show Graph Data." Copy and paste Cy3 time and intensity data from W1 into Sheet 1 of an Excel workbook. Copy and paste Cy5 time and intensity data from W2 into Sheet 2 of the same Excel workbook. In older versions of Excel (e.g. Excel 2000), each workbook can hold data from 116 traces; if the dataset contains more than 116 traces, use multiple workbooks. Graph overlayed Cy3 and Cy5 intensity versus time trajectories in Excel by running the macros "graph" and then "FRET" ("Tools → Macro → Macros").
4. Select traces according to the following criteria: (i) The fluorescence intensities of Cy3 and Cy5 should fall within the range expected for single fluorophores; (ii) there should be evidence of anti-correlated changes in Cy3 and Cy5 fluorescence intensity; (iii) loss of the Cy3 and Cy5 signals (via photobleaching or factor dissociation) should occur in a single time step. The Cy3/Cy5 intensity data for the traces that meet these selection criteria are highlighted and then combined in a new Excel workbook. The data is formatted to two decimal places by highlighting, right-clicking, and choosing "Format Cells→ Number → Decimal Places=2 ". Save as a Text (Tab delimited) file and append ".dat" to the file name (e.g. kepttraces.dat). Different files from the same dataset are indicated by adding a file number (n=1,2,3…) to the filename (e.g. kepttraces-1.dat).
5. Import the Cy3/Cy5 intensity data into Matlab. To do so, copy and paste the "loadTraces" m-file into the folder with the .dat files and set the Current Directory to this folder. Load traces into the Matlab workspace:

   >> X = loadTraces('kepttraces-', n);

where “n” is the number of files comprising the dataset to be analyzed.

6. Set Current Directory to the folder containing the smFRET data analysis scripts.
7. Separate the Cy3 and Cy5 data:

   >> [cy3x, cy5x] = SeparateCy(X);

   where “x” is a suffix of choice.
8. Plot the Cy3 and Cy5 intensity versus time trajectories:

   >> plotTraces(cy3x, cy5x);
9. Visually inspect the traces. If traces are identified that should be discarded from the analysis, do so using the J-filter:

   >> [cy3x, cy5x] = J_Filter(cy3x, cy5x, J);

   where “J” is a vector containing the trace labels of all traces to discard, e.g. J=[1034,1081,2009…].
10. Traces corresponding to a particular sub-population of molecules (e.g. fluctuating traces) to be analyzed separately can be selected and grouped together using the J2-filter:

    >> [cy3x, cy5x] = J2_Filter(cy3x, cy5x, J2);

    where “J2” is a vector containing the trace labels of the traces to keep.
11. Bleedthrough and baseline correct the traces:

    >> [cy3x, cy5x] = correctBaseline_end(cy3x, cy5x, nobleach);

    where “nobleach” is a vector that contains trace labels for those traces that do not exhibit photobleaching of Cy3 and/or Cy5. This script first corrects bleedthrough of Cy3 emission into the Cy5 channel by subtracting 7% of Cy3 intensity from the Cy5 trace at each data point. The bleedthrough coefficient can be changed in the m-file; by default it is set as “BLEED_COEF = 0.07”. The Cy3 and Cy5 traces are then baseline-corrected. For traces that exhibit photobleaching, the average intensity of the last 20 data points is calculated and set as the baseline. This parameter can be changed in the m-file to average more or fewer data points; by default it is set as “AVER = 20.” Trace labels for traces that do not exhibit photobleaching, or that contain aberrant data points within the last 20 frames that would skew the average,

should be stored in the variable “nobleach”. These traces are baseline-corrected using the average of the baseline values calculated for the other traces.

12. Save the traces:

    >> saveTraces(cy3x, cy5x, ‘filename.dat’);

13. Also save the Matlab workspace, so that all variables created during the analysis (not just the final set of traces) can be recalled later.

5.5.1.2 Basic plotting functions for smFRET data

1. Plot the smFRET versus time trajectories:

   >> plotFRETtraces(cy3x, cy5x);

   FRET efficiency is calculated as $E_{FRET} = I_{Cy5}/(I_{Cy3} + I_{Cy5})$ for each data point within a trace. When the sum of Cy3 and Cy5 intensities drop below a threshold, the trace will be truncated. The threshold can be changed in the m-file “getFRET”; by default it is set to “MININT=250”.

2. Plot a one-dimensional FRET histogram:

   >> FH = plotFRET(cy3x, cy5x, bins);

   where “bins” is the number of FRET bins (typically 35), which are equally spaced from E_{FRET} = -0.2 to 1.2. Column 1 (FRET) and Column 2 (normalized population) from “FH” can be copied to Origin 8 and fit to the sum of multiple Gaussians.

3. Plot a two-dimensional time evolution of population FRET histogram:

   >> plotTimeFRET(cy3x, cy5x, FRETbins, Tbinsize, cutoffT);

   where “FRETbins” is the number of bins in the FRET dimension (typically 24); “Tbinsize” indicates over how many time points E_{FRET} is averaged to generate the plot (typically 2); and “cutoffT” is the cutoff time in seconds (typically 20).

4. Plot a post-synchronized, two-dimensional time evolution of population FRET histogram:

   >> plotTimeFRET_ps(cy3x, cy5x, FRETbins, Tbinsize, cutoffT);

This script post-synchronizes all traces to the first data point above a FRET threshold. For subunit joining data, the first data point above E_{FRET} = 0.2 (set in m-file: "LIMIT = 0.2") was post-synchronized to t = 1 sec in the plot.

5.5.1.3 Calculation of subunit joining times

Procedure:

1. Calculate the FRET arrival times:

   >> G = getFRETon(cy3x, cy5x);

   This script returns a row vector "G" containing the frame number at which the onset of FRET occurs for each trace. If a FRET event is not identified for a given trace, that trace's entry will be bookmarked with "NaN". Two consecutive data points must exceed the threshold FRET value (typically E_{FRET} = 0.2), which is set in the m-file (i.e., "LIMIT=0.2"). Before running the script, set the threshold for summed Cy3 and Cy5 intensities below which the trace should be truncated in "getFRET3.m" (typically, "MININT=250").
2. Change subunit joining times to units of seconds by multiplying by the exposure time (sec $frame^{-1}$):

   >> fretON = G * 0.1;
3. Correct for the dead-time of the stopped-flow instrument:

   >> fretON = fretON - 1;

   The dead-time was estimated to be ~2.0 sec based on the time delay before the spike in the fluorescence signal when free Cy5 maleimide was delivered into the flowcell. In a typical subunit joining experiment, (Cy5)-50S subunits were delivered into the flowcell, followed by a 1 sec time delay and then data acquisition (Section 5.4.3). Therefore, 1 sec of the dead-time is already accounted for by the time delay, and an additional 1 sec is subtracted from the raw FRET arrival times to get the dead-time–corrected values.
4. Calculate the average FRET arrival time:

   >> mean(fretON)

Error analysis:

Errors were estimated by splitting the data into three equally sized sets, calculating the average FRET arrival time for each set, and taking the mean and standard deviation of the resultant values. To randomly split the data into three equal sets:

>> [A, B, C] = SplitData(fretON);

where A, B, and C each contain one-third of the FRET arrival times.

5.5.1.4 Calculation of (Cy3)-IF2 signal lifetime subsequent to subunit joining

For stream-acquisition data:

1. Run "getFLUORlifetime" script with two input arguments:

   >> H = getFLUORlifetime(cy3x, cy5x);

   This script returns a row vector "H" where each entry is the total time in frames between the onset of FRET and (Cy3)-IF2 signal loss (resulting from (Cy3)-IF2 dissociation or Cy3 photobleaching) or termination of data acquisition—whichever comes first—for each individual trace. The script identifies the frame at which subunit joining occurs using "getFRETon.m" (see Section 5.5.1.3), and it identifies the frame at which the (Cy3)-IF2 signal is lost when the summed intensity of Cy3 and Cy5 drops below a specified threshold (set in "getFRET3.m", typically "MININT = 250"); two consecutive data points must drop below the threshold to be counted as (Cy3)-IF2 signal loss.

2. Construct a population decay histogram from these dwell times:

   >> [t, N] = PopDecay(H);

   "t" contains the dwell time in units of frames. "N" contains the number of counts, i.e. N(i) is the number of traces for which the (Cy3)-IF2 signal lasts at least i frames after subunit joining.

3. Convert time data into units of seconds:

   >> t = t * 0.1;

4. Copy and paste data from "t" and "N" into Origin 8. Curves were fit with a single exponential decay of the form $y=A_1*\exp(-x/t_1) + y_0$, and the average lifetime of the decay, t_1, was reported.

For shuttered data:

As described in Section 2.6.3, data frames were collected continuously under constant laser excitation for five seconds (i.e. 50 frames) at the beginning of the experiment before starting a shuttering routine in which single 100 msec data frames were collected at regular intervals with the laser light blocked in between. Therefore, data were collected at two different time resolutions within the same experiment. To analyze this data, it is necessary to obtain the timestamps corresponding to each frame of the movie from MetaMorph. The timestamps are imported into Matlab and stored within a row vector. Then, lifetime of the (Cy3)-IF2 signal subsequent to subunit joining was calculated using a procedure similar to that described above:

1. Run "getFLUORlifetime" script with three input arguments:

   >> H = getFLUORlifetime(cy3x, cy5x, timestamps);

   where "timestamps" is a row vector in which timestamps(i) is the timestamp for the i^{th} frame of the movie. As before, set the appropriate thresholds in "getFRETon.m" (i.e., "LIMIT=0.2") and "getFRET3.m" (i.e., "MININT = 250") before running the script. The script returns a row vector "H" where each entry is the total time in seconds between the onset of FRET and (Cy3)-IF2 signal loss.
2. Construct a population decay histogram from the dwell times:

   >> [t, N] = PopDecay(H, timestamps);

   Here, "t" contains the dwell time in units of seconds, and "N" contains the number of counts. For analysis of the shuttering series data (Figure 2.10), a bin size of 15 sec was used to construct the histograms. This was so that dwell times from each dataset would have the same binning and so that all datasets (including that acquired at the slowest shuttering rate of 6 sec $frame^{-1}$) would have at least two data points per bin.

3. Copy and paste data from “t” and “N” into Origin 8. Curves were fit with a single exponential decay of the form $y=A_1*\exp(-x/t_1) + y_0$, and the average lifetime of the decay, t_1, was reported.

5.5.1.5 Idealization of smFRET versus time trajectories with vbFRET

Raw smFRET versus time trajectories were idealized to a hidden Markov model using the vbFRET software package [17], following the procedure outlined in the instruction manual. vbFRET uses a maximum-evidence–based method to identify the number of discrete conformational states within each smFRET trajectory and the most probable path through those states. The resulting idealized trajectories are then processed using a series of Matlab scripts to extract dwell times and calculate transition rates between particular conformational states.

1. Change the Matlab Current Directory to the folder with vbFRET. Open the vbFRET GUI by typing:
   >> vbFRET;
2. Load the .dat file of the traces to be analyzed: “File → Load Data → Add Files.” Check “Relabel Traces” so that traces will be numbered as 1, 2, 3…n, where n is the total number of traces. Click “Load Data.”
3. Under “Analysis Settings” set “Number of FRET states possible” to “Min: 1” and “Max: 5”. For subunit joining data, it is expected that there will be at least three FRET states in the smFRET trajectories, i.e. one zero-FRET state and two non-zero FRET states. These settings therefore allow vbFRET the freedom to potentially model an additional two states. Set “Fitting attempts per trace” to 25.
4. Truncate traces upon fluorophore photobleaching using “Traces → Remove Photobleaching.” Standard settings were as follows:
   Photobleach identification method: summed channel
   Truncate data when Channel 1 + Channel 2 is less than: 100
   Smooth traces before looking for photobleaching: checked
   Smooth over 2 time steps

Truncate an extra 2 time steps

Minimum trace length: 20

Click “Remove Photobleaching!”

5. Click “Analyze Data!”
6. Save data when analysis is finished: “File Save → Save Data”. Select “Save Session (.mat file)” so that analysis can be resumed at a later time. Also select “Save Idealized Traces- Save as concatenated text file (.dat)” to save the path data for subsequent post-processing steps.

5.5.1.6 Dwell time analysis

This section describes the dwell time analysis procedures that were used to calculate average lifetimes of the zero- and non-zero FRET states observed for reversible binding of (Cy5)-IF2 to 70S ICs during multiple-turnover GTP hydrolysis (Section 2.7) and for reversible docking of (Cy5)-50S subunits to $30SIC_{+IF3}$ (Sections 3.3 and 3.4). In both cases, the zero-FRET state represents the unbound, “OFF” state of the complex, while the non-zero FRET state represents the bound, “ON” state of the complex.

Procedure:

1. Import path data from vbFRET into Matlab: “File → Import Data.” The path data is stored in a two-column matrix where the first column contains the trace labels and the second column contains the idealized FRET efficiency at each time point.
2. Change the Current Directory to the folder with the lifetime analysis scripts.
3. Extract dwell times in each FRET state before transitioning to a new FRET state:

   >> dwellData = getRawDwell_all(pathData);

   The output variable “dwellData” is a four-column matrix of the form: [Trace label, FRET(i), FRET(i+1), n] where “FRET(i)” is E_{FRET} for the current dwell, “FRET(i+1)” is E_{FRET} for the next dwell, and “n” is the number of time steps spent in state i before transitioning to state i+1. This script does not discard the first or last dwells of the traces. For the last dwell of a trace, the entry for “FRET(i+1)” will be “NaN.”

4. For analysis of binding and dissociation kinetics, combine contiguous dwells in zero- and non-zero FRET states. To do so, first choose a threshold that optimizes separation between the zero- and non-zero FRET states based on a one-dimensional FRET histogram:

   >> FH = plot_idealizedFRET_hist(dwellData, bins);

   where "bins" is the number of FRET bins, equally spaced from -0.2 to 1.2.

   For all data presented in this thesis, the threshold separating the zero- and non-zero FRET states was determined to be $E_{FRET} = 0.2$. Thus, contiguous dwells with $E_{FRET} \leq 0.2$ and contiguous dwells with $E_{FRET} > 0.2$ were combined by typing:

   >> dwellData = purifyOnOffDwell(dwellData, threshold);

   where "threshold" = 0.2.
5. Plot population decay histograms of dwell times in the zero- and non-zero FRET states:

   >> [ts, N] = getDecay_DM(dwellData, bounds);

   where "bounds" is the range of FRET values defining each state, i.e. [-0.2, 0.2] for the zero-FRET state and [0.2, 1.2] for the non-zero FRET state. "ts" is the dwell time in seconds and "N" is the total number of dwells that last at least that long.
6. Copy and paste data from "ts" and "N" into Origin 8. Typically, curves were fit with a single exponential decay of the form $y=A_1*\exp(-x/t_1) + y_0$, and the average lifetime of the decay, t_1, was reported. However, two different types of 70S IC were formed upon delivery of (Cy5)-50S subunits to $30SIC_{+IF3}$ in the absence of free IF3, with low and intermediate stability, respectively (Section 3.4). Therefore, for this dataset, the population decay histograms of dwell times spent in the non-zero FRET state were fit with a double exponential decay of the form $y=A_1*\exp(-x/t_1) + A_2*\exp(-x/t_2) + y_0$, and the two lifetimes t_1 and t_2 were reported.
7. The procedures used to calculate $k_{a,\,app}$ and $k_{d,\,app}$ from lifetimes of the zero- and non-zero FRET states are described in the main text (Section 3.3).

Error Analysis:

Two different methods were used to estimate errors in the calculated lifetimes. In the first method, the idealized smFRET traces were randomly split into three equally sized sets. Each set was subjected independently to the lifetime analysis procedure described above, and the average and standard deviation of the resulting lifetimes were reported. To split the vbFRET path data into three sets of randomly selected traces, type:

>> [A, B, C] = SplitPathData(pathData);

where A, B, and C each contain one-third of the traces from pathData.

In the second method, resampling of the experimental dwell times was performed with replacement to generate 1000 bootstrap datasets. Population decay histograms were constructed from each dataset, lifetimes were determined by exponential fitting, and the mean and standard deviation of the resultant values were reported. Bootstrapping was performed in "R" version 2.12.2 [18] using the scripts provided in Appendix F and according to the following procedure:

1. Follow steps 1-4 of the dwell time analysis procedure above.
2. Copy and paste purified dwellData from Matlab into WordPad. In the first row, add labels for each of the four columns of data. Save as a .txt file. Copy to a folder for "R" input data files (e.g., "…/R-codes/Data").
3. Open "R" and choose the script to run: "File→ Open Script ." There are three options to choose from: 1) Use "ka_bootstrap_sample_dwells.R" to calculate the lifetime of the unbound state and $k_{a,\ app}$ with single exponential fitting to population decay histograms of dwell times in the zero-FRET state; 2) Use "ON_dwell_bootstrap_sampleDwells_singleExp.R" to calculate the lifetime of the bound state with single exponential fitting to population decay histograms of dwell times in the non-zero FRET state; and 3) Use "ON_dwell_bootstrap_sampleDwells_doubleExp.R" to calculate bound-state lifetimes with double exponential fitting to population decay histograms of dwell times in the non-zero FRET state for experiments in which two bound states with different lifetimes were observed.

4. Update the following lines of the "R" code before running it:
   i) Set the working directory:
      >setwd(dir = "C:/Documents and Settings/Administrator/Desktop/R-codes")
   ii) Update the name and location of the file to be analyzed:
      >dwellData = read.table(file = "data/test.txt", header = TRUE)
   iii) Enter initial guesses for the fitting parameters of the exponential decay:
      >iguessyo = 0
      >iguessAs1 = 150
      >iguesslt1 = 1
      >iguessAs2 = 350
      >iguesslt2 = 10
   iv) Enter the number of bootstrap datasets to construct:
      >nstep = 1000
   v) For "ka_bootstrap_sample_dwells.R", which calculates $k_{a,\,app}$, enter the (Cy5)-50S subunit concentration in units of nM:
      >subunitconc = 20
5. Highlight the whole script and run it by pressing "Ctrl+R"
6. The mean and standard deviations of the lifetimes will be printed. For "ON_dwell_bootstrap_sampleDwells_doubleExp.R", the mean and standard deviation for the percent contribution of fast and slow components to the decay will also be printed.

5.5.1.7 Construction of transition density plots

Transition density plots (TDPs) [19] are generated by plotting the "Starting FRET" versus "Ending FRET" for all transitions within the idealized traces as a contour plot representation of a two-dimensional histogram:

1. Import vbFRET path data into Matlab.
2. Extract dwell times in each FRET state before transitioning to a new FRET state, and remove the last dwell from each trace:

>> dwellData = getRawDwell_DeleteLastDwell(pathData);

3. Plot the TDP:

   >> plotTDP(dwellData, bins);

   where “bins” is the number of FRET bins, equally spaced from -0.2 to 1.2. Typically, bins = 24.

5.5.1.8 Calculation of average transition rate between non-zero FRET states within the 70S IC

This analysis was performed in order to quantify the conformational dynamics observed within $70SIC_{GTP}$ and 70S IC_{GDPNP}, formed upon docking of (Cy5)-50S subunits to $30SIC_{-IF3}$ complexes containing mRNA, fMet-$tRNA^{fMet}$, IF1 and (Cy3)-IF2 (see Section 2.6.4).

1. Manually select smFRET trajectories in which vbFRET models transitions between two or more non-zero FRET states within the 70S IC. The criteria used for trace selection were as follows: (i) At least one transition must occur between contiguous non-zero FRET states. On occasion, vbFRET models dwells in two different non-zero FRET states that are separated by a dwell in a zero-FRET state (~1% of the time); this behavior may represent undocking of one 50S subunit and docking of a new 50S subunit to the same 30S IC, and thus does not qualify as good evidence for the presence of two non-zero FRET states within the same 70S IC. (ii) The non-zero FRET states must be separated by $\Delta E_{FRET} \geq 0.05$. (iii) Dwells in each non-zero FRET state must be longer than one data point, since one-frame events may arise from a camera blurring artifact [17]. (iv) Occasionally, vbFRET does not accurately identify the photobleaching event where the trace should be truncated. When this is the case, aberrant data points downstream of the photobleaching event can be incorrectly modeled with transition(s) to new non-zero FRET states; this behavior was identified and was not considered to be a real transition.

   Scroll through the traces overlayed with their corresponding viterbi paths in the vbFRET GUI and write down the Trace Number for all traces which exhibit transitions between two or more non-zero FRET states based on the above criteria.

Note: The Trace Number is displayed in the lower-right hand corner of the GUI (e.g., "trace 200 of 264").

2. Import vbFRET path data into Matlab.
3. Create a new path data file containing data from the traces that were identified above as exhibiting one or more significant non-zero FRET transitions within the 70S IC:
   >> pathData = J2_Path(pathData, J2);
   where "J2" is a row vector containing the Traces Numbers for the traces that should be kept, e.g., J2=[2, 4, 5, 10,…].
4. For this subset of traces, extract dwell times in each FRET state before transitioning to a new FRET state:
   >> dwellData = getRawDwell_all(pathData);
5. Calculate the average transition rate between non-zero FRET states within the 70S IC:
   >> R = AverageTransitionRate(pathData, signif_trans);
   where "signif_trans" is the minimum change in E_{FRET} considered to represent a significant transition (typically 0.05). The threshold defining zero- and non-zero FRET states is set in the m-file (i.e., "KEY = 0.2"). This script identifies the total number of non-zero FRET transitions that occur and divides by the total time spent in non-zero FRET states, and outputs the average transition rate with units of transitions per second.

5.5.1.9 Analysis of short- and intermediate-lifetime 70S ICs

When (Cy5)-50S subunits were delivered to $30SIC_{+IF3}$ in the absence of free IF3, 50S subunit docking was observed to result in the formation of two different classes of 70S IC, whose respective lifetimes differed by more than an order of magnitude; these were termed short- and intermediate-lifetime 70S ICs, respectively (see Chapter 3.4). Following trace idealization with vbFRET (Section 5.5.1.5), non-zero FRET dwells corresponding to short- and intermediate-lifetime 70S ICs were parsed so that their lifetimes and FRET distributions could be analyzed separately. To do so, short-lifetime

70S ICs were defined as non-zero FRET dwells lasting less than 4 sec, and intermediate-lifetime 70S ICs were defined as non-zero FRET dwells lasting longer then 4 sec. This threshold was chosen because, for experiments in which (Cy5)-50S subunits were delivered to $30SIC_{+IF3}$ in the presence of free IF3—where the dwell time histograms are well-described by a single-exponential decay with a fast time constant—it results in classification of ≥95% of non-zero FRET dwells as short-lived.

1. Import path data from vbFRET into Matlab.
2. Extract dwell times in each FRET state before transitioning to a new FRET state:
   >> dwellData = getRawDwell_all(pathData);
3. Separate dwellData into two matrices containing the data for dwells in zero-FRET states (OFF dwells) and non-zero FRET states (ON dwells), respectively:
   >> [ONdwells, OFFdwells] = separateOnOffDwells(dwellData, threshold);
   where "threshold" defines the boundary between OFF and ON dwells (typically, threshold = 0.2). The output variables "ONdwells" and "OFFdwells" are four-column matrices of the form: [t, FRET(i), FRET(i+1), n] where each row contains the data for an individual ON or OFF dwell. "FRET(i)" is E_{FRET} for the current dwell, "FRET(i+1)" is E_{FRET} for the next dwell, and "n" is the number of time steps spent in the state i before transitioning to state i+1. "t" is a unique identifier for the cumulative ON or OFF dwell to which the individual dwells belong. For example, consecutive dwells in a 0.6 and a 0.8 FRET state within the same trajectory would be assigned the same value of "t".
4. Separate dwellData for the ON dwells further according to whether the individual dwells belong to a relatively long or relatively short cumulative ON dwell:
   >> [longONdwells, shortONdwells] = parse_cumeONdwells_length(ONdwells, cutoff);
   where "cutoff" is the dwell time in seconds above which a cumulative ON dwell is considered to be long; here, cutoff = 4. The output matrices have the same form as those in Step 3.

5. Plot the FRET distributions of the short- and intermediate-lifetime 70S ICs separately:

   >>FH_short = plot_idealizedFRET_hist(shortONdwells, bins);

   >>FH_intermediate= plot_idealizedFRET_hist(longONdwells, bins);

   where “bins” is the number of FRET bins, equally spaced from -0.2 to 1.2; typically, bins = 35. Column 1 (FRET) and Column 2 (normalized population) from “FH” can be copied to Origin 8 and fit to a Gaussian or sum of Gaussians.
6. Calculate the fraction of short- and intermediate-lifetime 70S ICs based on the 4 sec threshold (Table C.3 in the Appendix):

   Start with the original dwellData from Step 2 above. Combine contiguous ONdwells and OFFdwells:

   >> dwellData = purifyOnOffDwell(dwellData, threshold);

   where threshold = 0.2.

   Construct a column vector containing the dwell times of all cumulative ON dwells:

   >> ONdwells = getDwellHist(dwellData, bounds);

   where “bounds” gives the range of FRET efficiencies defining an ON dwell; typically bounds = [0.2, 1.2].

   Calculate the percentage of relatively short-lived and relatively long-lived ON dwells:

   >> getLongDwells(ONdwells, cutoff);

   where “cutoff” is the dwell time in seconds above which a cumulative ON dwell is considered to be long; here, cutoff = 4. This script prints the number of long and short ON dwells, and the percentage of each.

5.5.2 Analysis of two-color fluorescence co-localization data

Image analysis of two-color co-localization data was performed using SFTracer software, currently being developed by Mr. Victor Naumov in our laboratory. Post-processing of Cy3 and Cy5 fluorescence intensity versus time trajectories was performed in Matlab.

5.5.2.1 Analysis of Cy3 and Cy5 fluorescence intensity versus time trajectories

1. Calculate the Dual-View alignment parameters based on bleedthrough from the Cy3 channel into the Cy5 channel:

   Load .stk image file of (Cy3)-DNA oligo imaged at high laser power (see Section 5.4.4) into SFTracer. Click “Load Stack” button in the SFTracer GUI, select the file, and load 20 frames of the movie. To automatically calculate the alignment parameters, click “Alignment”, select “two channels” option, and click “Auto.” The alignment parameters for the Cy3 channel (channel 0) and the Cy5 channel (channel 1) will appear in the dialog box. Repeat for three separate movies to make sure that the same alignment parameters are obtained for each. The SFTracer channel-alignment function takes the highest intensity value for each pixel across the entire stack, histograms these values for the Cy3 and Cy5 channels, and sets a threshold equal to the mean plus three standard deviations. The alignment parameters (i.e., translation, rotation, and skew) are then varied to maximize the overlap of above-threshold pixels between the two channels.
2. Load the .stk file for the two-color experiment into SFTracer.
3. Align the Cy3 and Cy5 channels by entering the parameters determined above into the “Alignment” dialog box and clicking “OK.”
4. Define regions of interest (ROIs) by selecting fluorescent spots from the Cy3 channel. In the upper-right-hand corner of the GUI, select “Channel: green” and “Averaging: first 20 frames,” and click “Find.” SFTracer calculates the average intensity for each pixel within the Cy3 channel over the first 20 frames of the movie. It then uses a scoring function to assign a score to each pixel. The score is increased for pixels within a 2x2 region of foreground-intensity pixels surrounded by a 4x4 square of background-intensity pixels. The background and foreground levels are set in the GUI under “spotBg” and “spotFg”, respectively, and the default values were used for this analysis. The algorithm then searches for high-scoring 2x2 pixel regions and highlights them as the ROIs.

5. Plot Cy3 and Cy5 fluorescence intensity versus time trajectories for each ROI: Click "Calculate traces." Click "Save traces" and save the trajectories as a .tsv file (e.g., "SFTraces.tsv").
6. Copy the "loadSFTraces" m-file into the folder with the saved trajectories. Set the Matlab Current Directory to this folder. Load the trajectories into the Matlab workspace:

   >> [cy3x, cy5x] = loadSFTraces('SFTraces.tsv');
7. Bleedthrough and baseline correct the traces:

   >> [cy3x, cy5x] = correctBaseline (cy3x, cy5x);

   This script first corrects bleedthrough of Cy3 emission into the Cy5 channel by subtracting 7% of Cy3 intensity from the Cy5 trajectory at each data point. The bleedthrough coefficient can be changed in the m-file (i.e., "BLEED_COEF = 0.07"). The script then builds a histogram of Cy3 and Cy5 intensities and uses the most populated bin to baseline correct the traces.
8. Plot the traces:

   >> plotTraces(cy3x, cy5x);
9. Select traces according to the following criteria: (i) The Cy3 signal should exhibit single step photobleaching. If Cy3 does not photobleach, the signal should be stable and fall within the range expected for single Cy3 molecules (typically 1000-3000 units for the experimental conditions employed here). (ii) The Cy3 signal should last for more than 20 frames. (iii) For traces in which bursts of Cy5 fluorescence above the baseline are observed, the Cy5 intensity should increase and decrease in a stepwise manner. (iv) The Cy3 and Cy5 baselines should be roughly centered at zero intensity, deviating no more than approximately ±500 units.

   Store the trace labels for the traces that meet these criteria in a vector, e.g., J2=[10114, 10204, 10378…] and select this subset of traces for further analysis:

   >> [cy3x, cy5x] = J2_filter(cy3x, cy5x, J2);
10. Identify traces which exhibit one or more (Cy5)-T3 binding event. For this purpose, a (Cy5)-T3 binding event was defined as five or more consecutive Cy5 data points

above an intensity threshold. The threshold was chosen by first building a histogram of Cy5 intensity from all traces in the dataset:

```
>> [N, X] = hist(cy5x(:), 500);
>> N=N'; X=X';
```

The histogram typically contains two major peaks, corresponding to the Cy5 fluorescence background and signal from surface-bound (Cy5)-T3, respectively. Copy and paste the data into Origin 8 and fit the histogram to the sum of two Gaussians. The intensity threshold used to define (Cy5)-T3 binding was set as the background peak's center plus 1.25 times the peak width (i.e., xc1 + 1.25*w1). For control experiments in which very little (Cy5)-T3 binding was observed, the histogram contained only one major peak, corresponding to the Cy5 fluorescence background. In this case, the histogram was fit with a single Gaussian function for thresholding purposes.

Once the threshold has been determined, traces are selected which contain bursts of Cy5 fluorescence with five or more consecutive data points above the threshold, indicative of (Cy5)-T3 binding events:

```
>> [cy5y, frameno_cy5] = parseTraces_T3binding(cy5x);
```

The output variable "cy5y" contains Cy5 fluorescence intensity versus time data for those traces which exhibit at least one (Cy5)-T3 binding event. The variable "frameno_cy5" contains the frame number at which the first (Cy5)-T3 binding event occurs for each of these traces. The intensity threshold and number of consecutive data points used to define a (Cy5)-T3 binding event are set within the m-file (e.g., "IntensityThreshold = 1200" and "TimeThreshold = 5").

For datasets with especially low signal-to-noise Cy5 versus time trajectories, it may be necessary to smooth the Cy5 versus time trajectories before conducting this analysis in order to achieve accurate identification of the (Cy5)-T3 binding events:

```
>> cy5x = RollingAvg(cy5x, 3);
```

where the second input argument denotes the span for the rolling average.

The smoothed data can then be used to generate a Cy5 intensity histogram, calculate the intensity threshold, and select traces exhibiting (Cy5)-T3 binding events as described above.

11. Sub-populate Cy3 and Cy5 versus time trajectories based on whether or not (Cy5)-T3 binding events were identified:

    >> [cy3y, cy5y] = J2_filter(cy3x, cy5x, J2);

    where "J2" is a vector containing the trace labels for those traces identified in Step 10 to contain at least one (Cy5)-T3 binding event. The output variables "cy3y" and "cy5y" contain the Cy3 and Cy5 intensity versus time trajectories, respectively, for these traces.

    >> [cy3n, cy5n] = J_filter(cy3x, cy5x, J);

    where "J" is a vector containing the trace labels for the traces that do not exhibit a (Cy5)-T3 binding event, and the output variables "cy3n" and "cy5n" contain the intensity versus time trajectories for this subpopulation of traces.

12. For each trace, identify the frame at which the (Cy3)-IF2 signal is lost:

    >> frameno_cy3 = getCY3lifetime(cy3y);

    This script returns a vector containing the frame number at which the (Cy3)-IF2 signal is lost for each trace. If the (Cy3)-IF2 signal persists for the entire experimental observation time, the total number of data frames collected (typically 1200) will be returned for that trace. Three consecutive data points must drop below a Cy3 intensity threshold to be identified as (Cy3)-IF2 signal loss. The intensity threshold is determined by inspection of a Cy3 intensity histogram constructed from all traces in the dataset, and it is entered in the m-file "getCy3" (e.g., "MININT" = 800).

13. Visually inspect traces to confirm that the first (Cy5)-T3 binding event and the loss of (Cy3)-IF2 signal have been correctly identified:

    >> plotTraces(cy3y, cy5y, frameno_cy3, frameno_cy5);

    Blue and black vertical lines will be overlayed on the Cy3 and Cy5 fluorescence intensity versus time trajectories at the frames where (Cy3)-IF2 signal loss and the first (Cy5)-T3 binding event were identified, respectively.

14. Calculate the time difference between (Cy3)-IF2 signal loss and the first (Cy5)-T3 binding event on a trace-by-trace basis:

    >> diff = frameno_cy5 - frameno_cy3;

    The vector “diff” contains the time difference in frames for each trace. Negative values indicate that (Cy5)-T3 binding precedes (Cy3)-IF2 signal loss, while positive values indicate that (Cy5)-T3 binding occurs after (Cy3)-IF2 signal loss. These values can then be converted into units of time and plotted as a histogram.

5.5.2.2 Procedure for counting Cy3 and Cy5 fluorescent spots as a function of time

SFTracer was used to count the number of fluorescent spots within the Cy5 FOV for each data frame over the course of whole movies in order to assess the specificity of (Cy5)-T3 binding to the surface (Figure 4.2). It is similarly possible to count the number of Cy3 spots per FOV for each frame of a movie. The procedure used for automated spot-counting with SFTracer is described below:

1. Load .stk files from two-color co-localization experiments into the SFTracer GUI.
2. Calculate the spot background and spot foreground values (“spotBg” and “spotFg”) to be used for ROI identification:

    If Cy5 spots are to be counted, calculate the spotBg and spotFg values by selecting “channel: red” and “averaging: last frame” and then clicking “Find.” The last frame of the movie was used for the calculation since it typically contained a relatively large number of (Cy5)-T3 spots within the FOV. Record the calculated values. Repeat for all movies recorded on the same day and under the same imaging conditions for which (Cy5)-T3 spots were observed to accumulate within the FOV over the course of the movie. Take the average of the resulting spotBg and spotFg values, and use these average values for spot-counting analysis of all movies to be compared.

    If Cy3 spots are to be counted, calculate the spotBg and spotFg values by selecting “channel: green” and “averaging: first frame” and then clicking “Find.” Here, the first frame should be used since the FOV contains a large number of (Cy3)-IF2 spots at the beginning of the movie prior to (Cy3)-IF2 signal loss via photobleaching or factor

dissociation. Record the calculated values. As before, average spotBg and spotFg values should be calculated and used uniformly for spot-counting analysis of all movies to be compared.

3. Access the SFTracer directory on the lab server via the command line with PuTTY. Then, count the number of Cy3 or Cy5 spots for each frame of a movie by typing a command of the following form:

   ~/sftracer$ java SFTracer -numspots -channel right -sbg 1835 -sfg 2561 ../Data/movie1.stk > numspots_movie1.txt

   Here, “right” refers to the right channel of the Dual-View image and indicates that Cy5 spots should be counted. Enter “left” to count Cy3 spots instead. The numbers following “-sbg” and “-sfg” are the spotBg and spotFg values from Step 2 above. The next entry (i.e., “../Data/movie1.stk”) is the pathname of the .stk image file to be analyzed. The final entry (i.e., “numspots_movie1.txt”) is the name of the output data file, which will automatically be saved in the current directory. The second column of the output .txt file contains the number of fluorescent spots identified for each frame of the movie.
4. Download all output files from the server using WinSCP, and plot the number of Cy3/Cy5 spots versus time.

5.5.3 Analysis of three-color fluorescence data

Preliminary image analysis, comprising channel alignment, region selection, and generation of three-color fluorescence intensity versus time trajectories was performed using SFTracer software. Bleedthrough and baseline corrections and plotting of the trajectories was carried out in Matlab.

1. Channel alignment: Quad-View alignment parameters were calculated based on bleedthrough of Cy3 fluorescence into the Cy5 channel and bleedthrough of Atto488 fluorescence into the Cy3 channel. First, load the .stk image file of (Cy3)-DNA oligo imaged at high green laser power (see Section 5.4.5) into the SFTracer GUI. Click “Alignment”, select “three channels” option, and click “Auto.” Record the alignment

parameters for the Cy3 channel (channel 0) and the Cy5 channel (channel 1). Then, load the .stk image file of (Atto488)-DNA oligo imaged at high blue laser power (Section 5.4.5), and auto-calculate the alignment parameters as before. Record the parameters for the Cy3 channel (channel 0) and the Atto488 channel (channel 2).

Open the three-color movie to be analyzed with SFTracer. Perform channel alignment by entering the alignment parameters calculated above into the "Alignment" dialog box and clicking "OK."

2. Region selection: Define ROIs by picking fluorescent spots from the Cy5 channel. In the upper-right-hand corner of the GUI, select "Channel: red" and "Averaging: max all frames," and click "Find."
3. Plot the Cy3, Cy5, and Atto488 intensity versus time trajectories for each ROI by clicking "Calculate traces." Click "Save traces" and save the trajectories in a .tsv file (e.g., "SFTraces.tsv").
4. Load the trajectories into the Matlab workspace:

   >> [cy3x, cy5x, atto488x] = loadSFTraces('SFTraces.tsv');
5. Plot the trajectories:

   >> plotTraces_3color(cy3x, cy5x, atto488x);
6. Select traces that exhibit FRET as indicated by anticorrelation of the Cy3 and Cy5 signals, i.e., a drop in Cy3 intensity and concomitant rise in Cy5 intensity upon 50S subunit docking to the 30S IC:

   >> [cy3x, cy5x, atto488x] = J2_Filter_3color(cy3x, cy5x, atto488x, J2);

   where "J2" is a vector containing the trace labels of the traces to keep.

   Alternatively, traces can be discarded from the analysis using the J-filter:

   >> [cy3x, cy5x, atto488x] = J_Filter_3color(cy3x, cy5x, atto488x, J);

   where "J" is a vector containing the trace labels for the traces to be discarded.
7. Bleedthrough- and baseline-correct the traces. First, bleedthrough- and baseline-correct the Cy3 and Cy5 intensity versus time trajectories:

   >> [cy3x, cy5x] = correctBaseline_end(cy3x, cy5x, nobleach_cy3cy5);

where “nobleach_cy3cy5” is a vector that contains trace labels for those traces that do not exhibit photobleaching of Cy3 and/or Cy5, or that contain aberrant data points within the last 20 frames that would affect calculation of the baseline. This script first subtracts 7% of Cy3 intensity from the Cy5 trajectory at each data point to correct for bleedthrough of Cy3 emission into the Cy5 channel. It then performs the baseline correction by taking the average intensity over the last 20 data points for each Cy3 trajectory and each Cy5 trajectory and subtracting the resultant values from each data point of the trajectories. For traces that do not exhibit photobleaching, the average Cy3 and Cy5 baselines from the rest of the traces in the dataset is used for the baseline correction.

Baseline-correction of the Atto488 intensity versus time trajectories was performed separately:

>> atto488x = correctBaseline_end_1color(atto488x, nobleach_atto488);

where “nobleach_atto488” is a vector that contains trace labels for those traces that do not exhibit photobleaching of Atto488, or that contain aberrant data points within the last 20 frames that would affect calculation of the baseline. This script calculates the baseline for each Atto488 trajectory based on the average of the last 20 data points as before.

5.6 References

1. 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.

2. Elvekrog, M.M., *The Role of Initiation Factor Dynamics in Translation Initiation.* 2011, Columbia University.

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

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

5. Simonian, M.H. and J.A. Smith, *Spectrophotometric and colorimetric determination of protein concentration.* Curr Protoc Mol Biol, 2006. **Chapter 10**: p. Unit 10 1A.

6. Bowen, W.S., N. Van Dyke, E.J. Murgola, J.S. Lodmell, and W.E. Hill, *Interaction of thiostrepton and elongation factor-G with the ribosomal protein L11-binding domain.* J Biol Chem, 2005. **280**(4): p. 2934-43.

7. Van Dyke, N., W. Xu, and E.J. Murgola, *Limitation of ribosomal protein L11 availability in vivo affects translation termination.* J Mol Biol, 2002. **319**(2): p. 329-39.

8. Fei, J., *Coupling of Ribosome and tRNA Dynamics during Protein Synthesis*. 2010, Columbia University: New York.

9. Seo, H.S., S. Abedin, D. Kamp, D.N. Wilson, K.H. Nierhaus, and B.S. Cooperman, *EF-G-dependent GTPase on the ribosome. conformational change and fusidic acid inhibition.* Biochemistry, 2006. **45**(8): p. 2504-14.

10. 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.

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. Brandi, L., S. Marzi, A. Fabbretti, C. Fleischer, W.E. Hill, C.O. Gualerzi, and J. Stephen Lodmell, *The translation initiation functions of IF2: targets for thiostrepton inhibition.* J Mol Biol, 2004. **335**(4): p. 881-94.

13. Youngman, E.M., J.L. Brunelle, A.B. Kochaniak, and R. Green, *The active site of the ribosome is composed of two layers of conserved nucleotides with distinct roles in peptide bond formation and peptide release.* Cell, 2004. **117**(5): p. 589-99.

14. 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.

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. Rasnik, I., S.A. McKinney, and T. Ha, *Nonblinking and long-lasting single-molecule fluorescence imaging.* Nat Methods, 2006. **3**(11): p. 891-3.

17. 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.

18. *R: A language and environment for statistical computing*. 2011, R Development Core Team.

19. 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.

Appendix A – mRNA sequences

Table A.1: mRNA sequences.
mRNAs used for all biochemical and microscope work were derived from the mRNA encoding gene product 32 from T4 bacteriophage. mRNAs #1 and 2 contain a 5'-biotin modification (Bi) and an AUG or AUU start codon, respectively. They were used to assemble ribosomal complexes for smFRET experiments. mRNA #3 was used in the GTP hydrolysis and dipeptide formation assays, while mRNA #4 was used in the toeprinting assay. mRNAs #1-3 were chemically synthesized and purchased from Dharmacon, Inc., and mRNA #4 was generated by *in vitro* transcription. The Shine-Dalgarno (SD) sequence is underlined, the spacer region is italicized, and the start codon is bold and underlined.

#	**mRNA nucleotide sequence**
1	5'-Bi.CAACCUAAAACUUACACAAAUUAAAAAGGA*AAUAGAC***AUG**UUCAAAGUCGAAAAAU CUACUGCU-3'
2	5'-Bi.CAACCUAAAACUUACACAAAUUAAAAAGGA*AAUAGAC***AUU**UUCAAAGCGAAAAAUC UACUGCU-3'
3	5'-GCAACCUAAAACUUACACAGGGCCCUAAGGA*AAUAAAA***AUG**UUUAAA-3'
4	5'-GGCAACCUAAAACUUACACAGGGCCCUAAGGA*AAUAAAA***AUG**UUUAAAGAAGUAUACA CUGCUGAACUCGCUGCACAAAUGGCUAAACUGAAUGGCAAUAAAGGUUUUUCUUCUGAA GAUAAAGGCGAGUGGAAACUGAAACUCGAUAAUGCGGGUAACGGUCAAGCAGUAAUUC GUUUUCUUCCGUCUAAAAAUGAUGAACAAGCACCAUUCGCAAUUCUUGUAAAUCACGGU UUCAAGAAAAAUGGUAAAUGGUAUAUUGAAACAUGUUCAUCUACCCAUGGUGAUUACG AUUCUUGCCCAGUAUGUCAAUACAUCAGUAAAAAUGAUCUAUACAACACUGACAAUAAA GAGUACAGUCUUGUUAAACGUAAAACUUCUUACUGGGCUAACAUUCUUGUAGUAAAAG ACCCAGCUGCUCCAGAAAACGAAGGUAAAGUAUUUAAAUACCGUUUCGGUAAGAAAAUC UGGGAUAAAAUCAAUGCAAUGAUUGCGGUUGAUGUUGAAAUGGGUGAAACUCCAGUUG AUGUAACUUGUCCGUGGGAAGGUGCUAACUUUGUACUGAAAGUUAAACAAGUUUCUGG AUUUAGUAACUACGAUGAAUCUAAAUUCCUGAAUCAAUCUGCGAUUCCAAACAUUGACG AUGAAUCUUUCCAGAAAGAACUGUUCGAACAAAUGGUCGACCUUUCUGAAAUGACUUCU AAAGAUAAAUAAGG-3'

Appendix B – FRET arrival times for $30SIC_{+IF3}$

Table B.1: FRET arrival time for $30SIC_{+IF3}$ as a function of (Cy5)-50S concentration.
Different concentrations of (Cy5)-L11 reconstituted 50S subunits were stopped-flow delivered to surface-immobilized $30SIC_{+IF3}$ complexes containing IF1, (Cy3)-IF2-GTP, IF3, biotin-mRNA, and fMet-$tRNA^{fMet}$. The stopped-flow buffer did not contain IF3 or fMet-$tRNA^{fMet}$. FRET arrival times were calculated for each smFRET versus time trajectory as the time of the first data point > 0.2 FRET, minus the estimated dead time of our stopped-flow instrument (~2.0 sec). Data were compiled from three independent experiments. Errors were estimated by splitting the data into three equal parts and calculating the average and standard deviation. The decrease in the mean FRET arrival time with increasing (Cy5)-50S concentrations is consistent with a bimolecular association reaction.

[Cy5-50S], nM	Number of Molecules	Mean FRET Arrival Time (sec)
10	209	5.8 ± 0.3
20	262	4 ± 1
40	231	3.5 ± 0.4
60	217	3.0 ± 0.4

Appendix C – Lifetime analysis of zero- and non-zero FRET dwells for 30SIC$_{+IF3}$

Table C.1 Dwell times spent in the zero-FRET state for 30SIC$_{+IF3}$ complexes under varying conditions. 30SIC$_{+IF3}$ complexes contained IF1, (Cy3)-IF2-GTP, IF3, fMet-tRNA^{fMet}, and biotinylated mRNA with either an AUG or AUU start codon as indicated.

Start codon	IF3 in solution[a]	fMet-tRNA^{fMet} in solution[b]	[Cy5-50S], nM	Number of dwells	t_1, sec[c]
AUG	-	-	10	515	8.85
AUG	-	-	20	744	6.98
AUG	-	-	40	592	5.59
AUG	-	-	60	570	4.55
AUG	-	+	20	713	7.11
AUU	-	-	20	392	6.87
AUU	-	+	20	462	9.45
AUG	+	-	10	501	8.52
AUG	+	-	20	1024	5.76
AUG	+	-	40	1641	3.8
AUG	+	-	60	1858	3.44
AUG	+	+	20	1091	7.02
AUU	+	+	20	826	7.55

(a) In all experiments, IF3 (0.9 μM) was included in the buffers used to dilute 30SIC$_{+IF3}$s and to rinse the flowcell following surface immobilization. The stopped-flow buffer delivered into the flowcell contained (Cy5)-50S subunits, IF1, GTP, either in the presence or absence of IF3 as indicated.

(b) fMet-tRNA^{fMet} (0.9 μM) was either included or omitted from all dilution, rinsing, and stopped-flow buffers as indicated.

(c) Dwell times spent at FRET ≤ 0.2 were extracted from idealized smFRET trajectories and plotted as a population decay histogram. The number of dwells comprising each histogram is indicated. The histograms were fit with a single exponential decay of the form $y=A_1*\exp(-x/t_1) + y_0$, and the average lifetime, t_1, is reported here.

Table C.2 Dwell times spent in the non-zero FRET state for $30SIC_{+IF3}$ complexes under varying conditions.

Start codon	IF3 in solution[a]	fMet-tRNA^{fMet} in solution[b]	[Cy5-50S], nM	Number of dwells	t_1, sec (A1, %)[c]	t_2, sec (A2, %)[c]
AUG	-	-	10	322	0.7 (64%)	10.9 (34%)
AUG	-	-	20	497	1.1 (68%)	13.3 (32%)
AUG	-	-	40	436	0.7 (61%)	9.6 (39%)
AUG	-	-	60	484	0.7 (59%)	9.2 (41%)
AUG	-	+	20	458	0.8 (67%)	15.6 (33%)
AUU	-	-	20	249	1.3 (67%)	14.4 (33%)
AUU	-	+	20	298	1.6 (69%)	19.7 (31%)
AUG	+	-	10	281	0.9	N.A.
AUG	+	-	20	643	1.0	N.A.
AUG	+	-	40	1099	0.7	N.A.
AUG	+	-	60	1331	0.8	N.A.
AUG	+	+	20	601	0.6	N.A.
AUU	+	+	20	462	0.6	N.A.

(a) In all experiments, IF3 (0.9 μM) was included in the buffers used to dilute $30SIC_{+IF3}$s and to rinse the flowcell following surface immobilization. The stopped-flow buffer delivered into the flowcell contained (Cy5)-50S subunits, IF1, GTP, either in the presence or absence of IF3 as indicated.

(b) fMet-tRNA^{fMet} (0.9 μM) was either included or omitted from all dilution, rinsing, and stopped-flow buffers as indicated.

(c) Dwell times spent at FRET > 0.2 were extracted from idealized smFRET trajectories and plotted as a population decay histogram. The number of dwells comprising each histogram is indicated. When IF3 was not included in solution, dwell time histograms were fit with a double exponential decay of the form $y=A1*\exp(-x/t_1) + A2*\exp(-x/t_2) + y_0$. The percent contribution of the fast and slow components to the decay were estimated as A1/(A1+A2)*100 and A2/(A1+A2)*100, respectively. When IF3 was included in solution, the dwell time histograms were fit with a single exponential decay of the form $y=A_1*\exp(-x/t_1) + y_0$.

Table C.3 Quantification of the partitioning between short- and long-lived (Cy5)-50S subunit docking events based on a 4 sec threshold. This threshold was chosen because, for experiments in which IF3 was kept in solution, where the dwell time histograms are well-described by a single-exponential decay with a fast time constant, it results in classification of ≥95% of non-zero FRET dwells as short-lived.

Start codon	IF3 in solution[a]	fMet-tRNA^{fMet} in solution[b]	[50S], nM	Number of dwells	Short dwells (%)[c]	Long dwells (%)[c]
AUG	-	-	10	367	76	24
AUG	-	-	20	574	76	24
AUG	-	-	40	506	76	24
AUG	-	-	60	560	76	24
AUG	-	+	20	483	84	16
AUU	-	-	20	297	76	24
AUU	-	+	20	337	73	27
AUG	+	-	10	337	95	5
AUG	+	-	20	800	97	3
AUG	+	-	40	1426	95	5
AUG	+	-	60	1757	95	5
AUG	+	+	20	831	98	2
AUU	+	+	20	657	98	2

(a) In all experiments, IF3 (0.9 μM) was included in the buffers used to dilute 30SIC$_{+IF3}$s and to rinse the flowcell following surface immobilization. The stopped-flow buffer delivered into the flowcell contained (Cy5)-50S subunits, IF1, GTP, either in the presence or absence of IF3 as indicated.

(b) fMet-tRNA^{fMet} (0.9 μM) was either included or omitted from all dilution, rinsing, and stopped-flow buffers as indicated.

(c) Dwell times spent at FRET > 0.2 were extracted from the idealized smFRET trajectories and classified as short- or long-lived if they lasted less or more than 4 sec, respectively.

Appendix D – Example smFRET traces of rare subunit-joining dynamics observed for 30SIC$_{-IF3}$

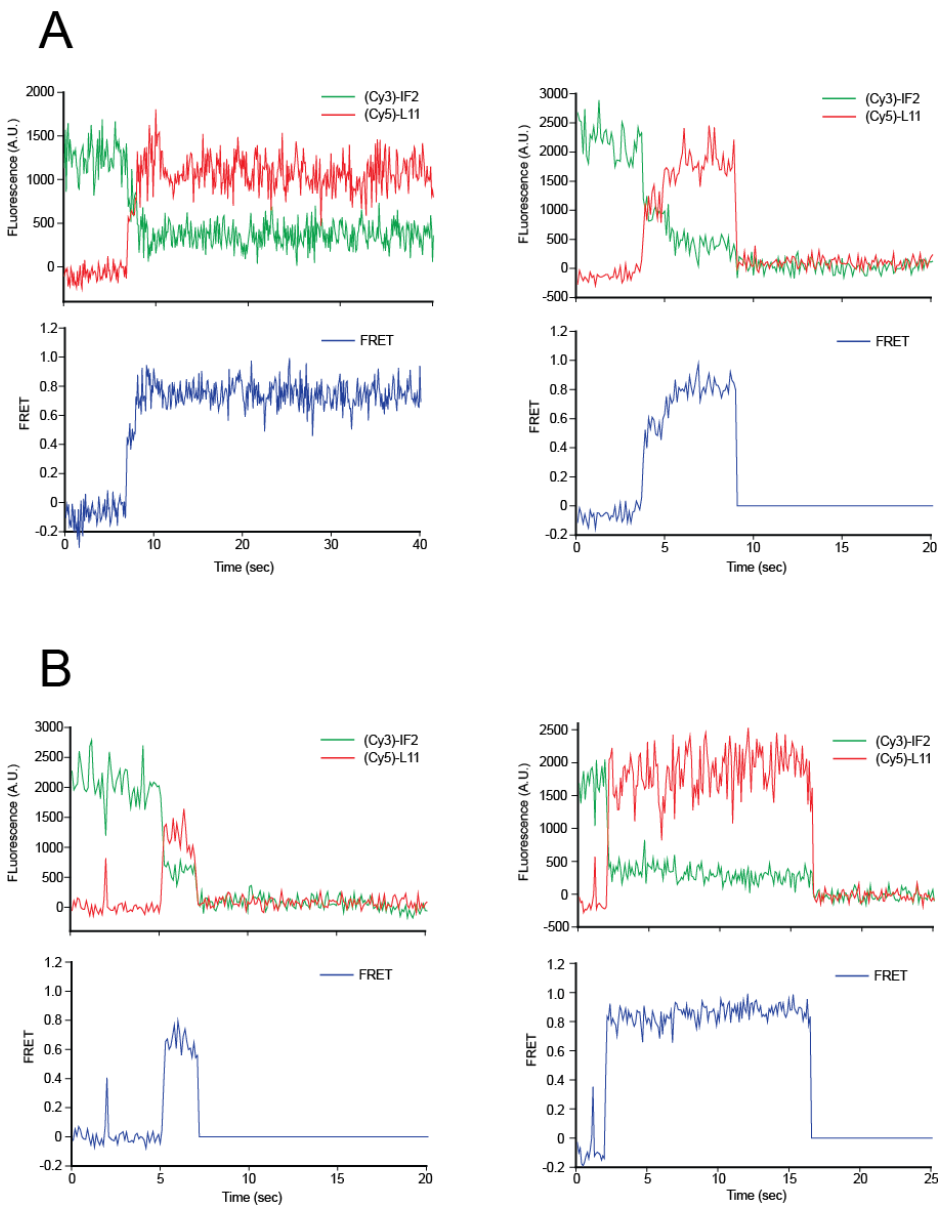

Figure D.1: Example smFRET traces from 30SIC$_{-IF3}$ suggesting the presence of a transient subunit-joining intermediate.
(Cy5)-L11 labeled 50S subunit were stopped-flow delivered to 30SIC$_{-IF3}$ complexes containing IF1, (Cy3)-IF2-GTP, biotin-mRNA, and fMet-tRNA^{fMet}. In a small subpopulation of the traces, FRET behaviors indicative of a transient conformational intermediate along the 70S IC formation pathway were observed. (A) Upon subunit joining, ~1 sec dwells in a mid-FRET state were followed by a transition to a longer lived high-FRET state. (B) Short, ~100 msec FRET events occasionally preceded stable subunit joining, reminiscent of the transient subunit docking events observed with 30SIC$_{+IF3}$.

Appendix E – Matlab scripts

The Matlab scripts used for the data analysis procedures described in Section 5.5 are presented below. These include the scripts that were specifically mentioned as well as those that are executed through internal function calls. Many of the scripts were written by Mr. Pallav Kosuri and Dr. Jingyi Fei as part of the FRET data analysis package (FDAP v1.7). Others were written by Dr. Jiangning Wang and yet others by myself. Specific authorship details are included in the first comment field of all scripts below. The scripts are presented in alphabetical order as follows:

addLabels.m

```
%From FDAP v1.7

function [cy3, cy5]=addLabels(cy3, cy5, labels)

        %add labels
    cy3=[labels; cy3];
    cy5=[labels; cy5];
```

addLabels_1color.m

```
%same as addLabels from FDAP v1.7, modified for 1-color data

function [cy3]=addLabels_1color(cy3, labels)

        %add labels
    cy3=[labels; cy3];
```

addLabels_3color.m

```
%same as addLabels from FDAP v1.7, modified for 3-color data

function [cy3, cy5, cy2]=addLabels_3color(cy3, cy5, cy2, labels)

        %add labels
    cy3=[labels; cy3];
    cy5=[labels; cy5];
    cy2=[labels; cy2];
```

AverageTransitionRate.m

```
%AverageTransitionRate will calculate the average number of transitions
%sec-1 for transitions between two non-zero FRET states within the same
%70SIC. Input argument "pathData" should have the four column format,
i.e.
%[fileno FRET(i) FRET(i+1) n] and should be generated by using
getRawDwell_all
%such that first and last dwells of each trace are kept. The latter is
%important since AverageTransitionRate will identify the last dwell of
a
%given trace by the fact that FRET(i+1) is NaN. This script sums
together
%the length of all ON dwells that contain at least one significant
%transition between non-zero FRET states (deltaFRET for what qualifies
as a significant
%transition is defined with input parameter "signif_trans") and divides
the
%total number of such transitions observed by this value.
%Written by DDM

function R = AverageTransitionRate(pathData, signif_trans)

KEY = 0.2;  %Threshold separating bound and unbound states

pathData = pathData(:,[1,2,3,4]);

fileno = pathData(:,1);

t = 1; %current dwell

numtransitions = 0;  %counter for number of transitions between non-
zero FRET states

spf = 0.1;  %exposure time

%to hold path data for dwells comprising ON dwells that have at least
one
%significant transition between non-zero FRET states
```

```
SavedData = ones(1,4);

while t <= length(fileno)

    flag = 0;  %flags if a significant transition has been detected

    if pathData(t, 2) <= KEY %If current dwell is an OFF dwell
        t = t + 1;  %Go to next dwell without saving entry
    else
        if pathData(t, 3) <= KEY | isnan(pathData(t,3))  %If current
dwell is an isolated ON dwell followed by an OFF dwell or a new trace
            t = t +1;  %Go to next dwell without saving entry

        elseif pathData(t, 3) > KEY   %If current dwell is an ON dwell
followed by another ON dwell
            i = 1;
            while pathData(t,2) > KEY

                tempSavedData(i, :) = pathData(t,:);

                if pathData(t,3) > KEY && abs(diff([pathData(t,2)
pathData(t,3)])) > signif_trans
                    flag = 1;
                    numtransitions = numtransitions + 1;
                end

                i = i + 1;
                t = t + 1;

                if t > length(fileno)
                    break;
                end
            end
        end
    end

    if flag == 1
        SavedData = [SavedData; tempSavedData];
    end
end

SavedData = SavedData(2:end, :);

totalONdwell_frames = sum(SavedData(:,4));

totalONdwell_sec = totalONdwell_frames * spf;

R = numtransitions / totalONdwell_sec;

disp(['R =  ' num2str(R) ' transitions per sec']);
```

correctBaseline.m

```
%From FDAP v1.7

function [cy3, cy5] = correctBaseline(cy3, cy5)

    BLEED_COEF = 0.07

    [cy3, cy5, labels]=removeLabels(cy3, cy5);

    X = linspace(0,18000,360);

    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);
```

correctBaseline_end.m

```
%the argument noBleach is a list of all the traces that do not show
%photobleaching
%From FDAP v1.7

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 = 20

    [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, cy5]=addLabels(cy3, cy5, labels);

%base_list = [labels; cy3_baseline; cy5_baseline];
```

correctBaseline_end_1color.m

```
%the argument noBleach is a list of all the traces that do not show
%photobleaching
%correctBaseline_end from FDAP v1.7, modified for 1-color data

function [cy3] = correctBaseline_end_1color(cy3, noBleach)

        %AVER specifies how many datapoints at the end of the trace
that
        %are averaged to get the baseline.
    AVER = 20

    [cy3, labels]=removeLabels_1color(cy3);

    ms = size(cy3);
```

```
        %individually correct all traces that show photobleaching
    for i = 1:ms(2)
        if find(noBleach==labels(i))
            cy3_baseline(i)=0;
        else
            cy3_baseline(i) = mean(cy3((ms(1)-AVER):ms(1), i));
            cy3(:, i) = cy3(:, i) - cy3_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));
    no = find(cy3_baseline == 0);
    cy3(:, no) = cy3(:, no) - cy3_average;

    [cy3]=addLabels_1color(cy3, labels);
```

getCy3.m

```
%getCy3 will take a matrix of Cy3 intensity traces, remove labels, and
%enter a 0 whenever Cy3 intensity drops below the minimum. This
facilitates
%identifying the specific frame where the minimum threshold is
breached.
%To be used in conjunction with getCY3lifetime.
%Written by DDM

function F = getCy3(cy3)

    cy3=cy3(2:end,:); %Remove labels

    %minimum total intensity
    MININT=800;

    ms=size(cy3);
    cols=ms(2);
    rows=ms(1);

    for j=1:cols
        for i=1:rows
            if cy3(i,j) > MININT
               F(i,j)=cy3(i,j);
            else
                    %too low total intensity, Cy3 has photobleached
               F(i,j)=0;
            end
        end
    end
```

getCY3lifetime.m

```
%getCY3lifetime returns a row vector where each entry gives the
lifetime
%of the Cy3 fluorescence signal before photobleaching
%Written by DDM

function L = getCY3lifetime(cy3,X);
    %Input variable X is a row vector containing the timestamps in sec
    %corresponding to each frame of the movie

    F=getCy3(cy3);

    ms=size(F);
    cols=ms(2);
    rows=ms(1);

    if nargin==1 %Results output in units of frame number
        for j=1:cols
        %If cy3 does not photobleach, the array entry will be the total
        %number of frames in the movie
            L(j)=rows;
        %If cy3 does photobleach, the array entry will be the number of
        %frames before photobleaching.
            for i=1:(rows-1)
                if F(i,j)==0 & F(i+1,j)==0 & F(i+2,j)==0
                L(j)=i;
                break;
                end
            end
        end

    elseif nargin==2 %Results output in units of seconds
        for j=1:cols
        %If cy3 does not photobleach, the array entry will be the final
        %timestamp of the movie
            endtime=X(rows);
            L(j)=endtime;
        %If cy3 does photobleach, the array entry will be the time in
        %seconds before photobleaching.
            for i=1:rows
                if F(i,j)==0 & F(i+1,j)==0 & F(i+2,j)==0
                L(j)=X(i);
                break;
                end
            end
        end
    end
```

getDecay.m

```
%From Jiangning Wang, modified by DDM

function [ts, N] = getDecay_DM(dwellData, bounds)

    start_low = bounds(1);
    start_high = bounds(2);

    selection = find(dwellData(:,2) > start_low & dwellData(:,2) <
start_high);
    maxT = max(dwellData(selection,4));
    t = 1:(maxT-1);
    ts = t*0.1;
    for i = t
        N(i) = length(find(dwellData(selection, 4) >= t(i)));
    end

    figure, plot(ts, N, '.')
    title(['Decay curve for state bounded by: start=['
num2str(start_low) ', ' num2str(start_high) ']'])
    ylabel('Population')
    xlabel('t (s)')
    DC=[ts;N].';
    save('decaycurvedata.dat', '-ascii', 'DC');

    ts=ts';
    N=N';
```

getDwellHist.m

```
%getDwellHist takes purified dwell data in the four column format, i.e.
%[fileno, FRET(i), FRET(i+1), dwelltime_in_frames] and returns a column
vector N
%containing dwelltimes (in seconds) for all dwells with a starting FRET
value that
%falls within the range defined by input argument "bounds"
%Written by DDM

function N = getDwellHist(dwellData, bounds)

    start_low = bounds(1);
    start_high = bounds(2);

    selection = find(dwellData(:,2) > start_low & dwellData(:,2) <
start_high);

    fps= 0.1; %enter frame rate

    N = dwellData(selection, 4);
```

```
    N = N*fps;
```

getFLUORlifetime.m

```
%getFLUORlifetime returns a row vector where each entry gives the total
%duration of the fluorescence signal (including cy3 fluorescence after
%cy5 photobleaching) for a given trace. To be used for stopped-flow
%delivery experiment with cy3-IF2 and cy5-50S subunits
%Written by DDM

function H = getFLUORlifetime(cy3,cy5,X);
    %Input variable X is a row vector containing the timestamps in sec
    %corresponding to each frame of the movie

    F=getFRET3(cy3,cy5);

    ms=size(F);
    cols=ms(2);
    rows=ms(1);


    if nargin==2 %Results output in units of frame number
        G=getFRETon(cy3,cy5);
        for j=1:cols
        %If cy3 does not photobleach, the array entry will be the
number
        %of frames after initial FRET onset until data acquisition was
        %terminated. This will necessarily be an underestimation of the
        %true fluorescence lifetime.
            H(j)=rows-G(j);
        %If cy3 does photobleach, the array entry will be the number of
        %frames from the onset of FRET to the cy3 photobleaching event.
            for i=1:(rows-1)
                try if F(i,j)==0 & F(i+1,j)==0 & i>G(j);
                    H(j)=i-G(j);
                    break;
                    end
                catch
                    H(j)=rows-G(j);
                    disp(lasterr)
                end
            end
        end
    elseif nargin==3 %Results output in units of seconds
        Gframes=getFRETon(cy3,cy5);
        G=getFRETon(cy3,cy5,X);
        for j=1:cols
        %If cy3 does not photobleach, the array entry will be the time
in
        %seconds between initial FRET onset and termination of data
        %acquisition
```

```
            endtime=X(rows);
            H(j)=endtime-G(j);
        %If cy3 does photobleach, the array entry will be the time in
        %seconds between the onset of FRET and the cy3 photobleaching
        %event.
            for i=1:(rows-1)
                try if F(i,j)==0 & F(i+1,j)==0 & i>Gframes(j);
                    H(j)=X(i)-G(j);
                    break;
                    end
                catch
                    H(j)=endtime-G(j);
                    disp(lasterr)
                end
            end
        end
    end
```

getFRET.m

```
%getFRET will identify the first instance where the sum of cy3+cy5
drops below the threshold and mark
%that data frame and all subsequent data frames with NaN.
%In contrast, getFRET2 checks each data frame individually to see if
sum of cy3+cy5
%intensity drops below threshold and marks them as NaN accordingly.
%From FDAP v1.7

function F = getFRET(cy3, cy5)

    [cy3, cy5, labels]=removeLabels(cy3, cy5);

    %minimum total intensity
    MININT=250;

    ms=size(cy3);
    cols=ms(2);
    rows=ms(1);

    summ=cy3+cy5;

    for j=1:cols
        for i=1:rows
            if summ(i,j) > MININT
                F(i,j)=cy5(i,j)/summ(i,j);
            else
                    %too low total intensity to be FRET
                    %delete rest of trace
                F(i:rows,j)=NaN;
                break;
            end
```

```
        end
    end
```

getFRET2.m

```
%getFRET2 checks each data frame individually to see if sum of cy3+cy5
%intensity drops below threshold and marks them as NaN accordingly; in
contrast, getFRET will identify the
%first instance where the sum of cy3+cy5 drops below the threshold and
mark
%that data frame and all subsequent data frames with NaN.
%From FDAP v1.7

function F = getFRET2(cy3, cy5)

    [cy3, cy5, labels]=removeLabels(cy3, cy5);

    %minimum total intensity
    MININT=250;

    ms=size(cy3);
    cols=ms(2);
    rows=ms(1);

    summ=cy3+cy5;

    for j=1:cols
        for i=1:rows
            if summ(i,j) > MININT
                F(i,j)=cy5(i,j)/summ(i,j);
            else
                    %too low total intensity to be FRET
                F(i,j)=NaN;
            end
        end
    end
```

getFRET3.m

```
%getFRET3 is the same as getFRET2 except instead of entering NaN when
the
%sum of cy3 and cy5 intensities drops below the threshold, it enters 0.
%This facilitates identifying the specific frame where the minimum
%threshold is breached. To be used in conjunction with getFRETon and
%getFLUORlifetime
%Written by DDM

function F = getFRET3(cy3, cy5)
```

```
    [cy3, cy5, labels]=removeLabels(cy3, cy5);

    %minimum total intensity
    MININT=250

    ms=size(cy3);
    cols=ms(2);
    rows=ms(1);

    summ=cy3+cy5;

    for j=1:cols
        for i=1:rows
            if summ(i,j) > MININT
                F(i,j)=cy5(i,j)/summ(i,j);
            else
                    %too low total intensity to be FRET
                F(i,j)=0;
            end
        end
    end
```

getFRETon.m

```
%getFRETon returns a row vector where each entry is either the frame
number
%or time in seconds at which the onset of FRET occurs. Frame number is
%returned when two inputs are specified, and seconds are returned if
three
%inputs are specified. If FRET does not occur above the chosen
threshold
%for a given trace, NaN will be inserted as the entry for that trace
%Written by DDM

function G = getFRETon(cy3,cy5,X);
    %getFRET3 is used here. It will mark data points that drop below
min
    %intensity as 0, but will not delete subsequent data.
    %Input variable X is a row vector containing the timestamps in sec
    %corresponding to each frame of the movie
    F=getFRET3(cy3,cy5);

    %Threshold FRET value
    LIMIT=0.20

    ms=size(F);
    cols=ms(2);
    rows=ms(1);
```

```
    for j=1:cols
        %Entry will be NaN if trace does not attain threshold FRET
value
        G(j)=NaN;
        %Otherwise entry will be the frame number at which FRET
        %value above threshold is first observed
        for i=1:rows
            if F(i,j)>LIMIT & F(i+1,j)>LIMIT
                G(j)=i;
                break;
            end
        end
    end
    %If the array of timestamps X is provided as the third function
input,
    %output entries are reassigned to be the time in seconds at which
FRET
    %above the threshold is first observed
    if nargin==3
        for j=1:cols
            if ~isnan(G(j))
                ontime=X(G(j));
                G(j)=ontime;
            end
        end
    end
```

getLongDwells.m

```
%getLongDwells returns a column vector containing all dwell times
%that are longer than a specified threshold.
%Input argument ONdwells is a column vector containg dwell times in
units of
%sec, i.e. the the output from getDwellHist. Input argument "threshold"
%defines long versus short dwells and should be given in sec.
%On-screen display provided of the number of long and short dwells, and
the
%percentage of the total for each
%Written by DDM

function N = getLongDwells(ONdwells, threshold)

LongDwells = find(ONdwells(:) > threshold);

N = ONdwells(LongDwells);

NumLong = length(LongDwells);

disp(['Number of long dwells = ' int2str(NumLong)])

ShortDwells = find(ONdwells(:) <= threshold);
```

```
NumShort = length(ShortDwells);

disp(['Number of short dwells = ' int2str(NumShort)])

TotalDwells = NumLong + NumShort;

PerLong = (NumLong/TotalDwells)*100;
PerShort = (NumShort/TotalDwells)*100;

disp([int2str(PerLong) '% Long']);
disp([int2str(PerShort) '% Short']);
```

getRawDwell_all.m

```
%From Jiangning Wang, modified by DDM

function dwellData = getRawDwell_all(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)) | ((i==(length(FRET)-1)) &
(fileno(i)==fileno(i+1)));
            dwellData(t, :) = [fileno(i) FRET(i) NaN n];
            t = t + 1;
            n = 1;
            continue;
        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
```

getRawDwell_DeleteLastDwell.m

```
%From Jiangning Wang, modified by DDM

function dwellData = getRawDwell_DeleteLastDwell(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))
            n = 1;
            continue; % skip the last dwell of each trace without
saving it,
            %equivalent to removing the last dwell
        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
```

getTimeFRET.m

```
%From FDAP v1.7

function H = getTimeFRET(F, FRETbins, Tbs)

    ms=size(F);

        %no of timesteps after binning
    nt=floor(ms(1)/Tbs);

        %apply binning in time dimension
    Fb = zeros(nt, ms(2));
    for i = (1:nt)
        Fb(i,:) = mean(F((i-1)*Tbs+1:i*Tbs,:));
    end
```

```
        %build histogram
    Y = linspace(-0.2, 1.2, FRETbins);
    H = zeros(FRETbins, nt);
    for i = 1:nt
        H(:, i) = hist(Fb(i, :), Y)';
    end

%Final matrix H: Number of rows is equal to number of FRET bins.
%Number of columns is equal to the number of time bins. Any given
column
%is the population FRET distribution for a particular binned timepoint
```

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.
%From FDAP v1.7

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
```

J_filter_3color.m

```
%Takes cy3, cy5, and cy2 traces as well as a vector with the
%J-selections (bad traces) as arguments and returns the filtered
traces.
%same as J_filter from FDAP v1.7, modified for 3-color data

function [cy3, cy5, cy2] = J_filter_3color(cy3, cy5, cy2, J)

    ms=size(cy3);

        %selective filter
    for i = 1:length(J)
```

```
        [cy3, cy5, cy2]=removeTrace_3color(cy3, cy5, cy2, J(i));
    end

        %display results
    ms2=size(cy3);
    total=ms(2)
    discarded=ms(2)-ms2(2)
    kept=total-discarded
```

J_Path.m

```
%J_Path will discard a subset of the idealized traces from vbFRET.
%"pathData" is the two-column vector of path data output from vbFRET.
%"selection" is a vector containing the trace labels for the idealized
traces that
%you want to delete. It is a good idea to make sure that the trace
labels you
%wrote down correspond to the trace labels that show up in the path
data
%file, since things can get messed up if traces were relabeled prior to
or during
%the vbFRET session.
%Written by DDM

function pathData = J_Path(pathData, selection)

pathData = pathData(:, [1,2]);
fileno = pathData(:, 1);
sizePath = size(pathData);
numrows = sizePath(1);
selind = 1;

for i=1:numrows
    if find(selection == fileno(i));
    else selind = [selind; i];
    end
end

lensel = length(selind);
rowind = selind(2:lensel);

pathData = pathData(rowind, :);
```

J2_filter.m

```
%Keep selected traces and discard the rest.
%From FDAP v1.7
```

```
function [cy3, cy5] = J2_filter(cy3, cy5, select)

    [cy3a, cy5a, labels]=removeLabels(cy3, cy5);
    ms = size(cy3);
    for i = 1:ms(2)
        if find(select==labels(i))
        else
            [cy3, cy5]=removeTrace(cy3, cy5, labels(i));
        end
    end
    ms2=size(cy3);
    total_kept=ms2(2)
```

J2_filter_3color.m

```
%Keep selected 3-color traces and discard the rest.
%same as J2_filter from FDAP v1.7, modified for 3-color data

function [cy3, cy5, cy2] = J2_filter_3color(cy3, cy5, cy2, select)

    [cy3a, cy5a, cy2a, labels]=removeLabels_3color(cy3, cy5, cy2);
    ms = size(cy3);
    for i = 1:ms(2)
        if find(select==labels(i))
        else
            [cy3, cy5, cy2]=removeTrace_3color(cy3, cy5, cy2,
labels(i));
        end
    end
    ms2=size(cy3);
    total_kept=ms2(2)
```

J2_Path.m

```
%J2_Path will select a subset of the idealized traces from vbFRET.
Input
%should be the two-column vector of path data output from vbFRET.
%"selection" is a vector containing the trace labels for the idealized
traces that
%you want to keep. It is a good idea to make sure that the trace labels
you
%wrote down correspond to the trace labels that show up in the path
data
%file, since things can get messed up if traces were relabeled prior to
or during
%the vbFRET session.
%Written by DDM
```

```
function pathData = J2_Path(pathData, selection)

pathData = pathData(:, [1,2]);
fileno = pathData(:, 1);
sizePath = size(pathData);
numrows = sizePath(1);
selind = 1;

for i=1:numrows
    if find(selection == fileno(i));
    selind = [selind; i];
    end
end

lensel = length(selind);
rowind = selind(2:lensel);

pathData = pathData(rowind, :);
```

loadSFTraces.m

```
%From Victor Naumov

function [green, red, blue, magenta] = loadSFTraces(filePrefix,
numFiles)
  traces=[];
  if (nargin<2)
    numFiles=1;
  end
  numChannels=1;
  for fileNum = 1:numFiles
    if (nargin==1)
      fileName=filePrefix
    else
      fileName=[filePrefix int2str(fileNum) '.tsv']
    end
    handle=fopen(fileName);
    if (~(handle>0))
      error('file cannot be opened');
    end
    token=fscanf(handle,'%s',1);
    while ~strcmp(token,'volumes')
      token=fscanf(handle,'%s',1);
    end

    [region, hasdata]=fscanf(handle, '%d', 1);
    while hasdata>0
      channel=fscanf(handle, '%d', 1);
      if (channel+1>numChannels) %channel numbering starts at zero
    numChannels=channel+1;
```

```
      end
      area=fscanf(handle, '%d', 1);
      length=fscanf(handle, '%d', 1);
      background=fscanf(handle, '%d', 1);

      %unlike in loadTraces.m, the labels contain the region ID rather
than count up sequentially
      %also, file IDs increase in increments of 10000, not 1000
      label=fileNum*10000+region;
      volumes=[label; fscanf(handle, '%d', length)-0];
      sizedif=size(volumes,1)-size(traces,1);
      if (sizedif > 0)
        traces=[traces; zeros(sizedif, size(traces,2))];
      end
      if (sizedif < 0)
        volumes=[volumes; zeros(-sizedif,1)];
      end
      traces=[traces volumes];
      [region, hasdata]=fscanf(handle, '%d', 1);
    end
    fclose(handle);
  end

  for i = (1:size(traces,2)/numChannels)
    green(:,i) = traces(:, (i-1)*numChannels+1);
    if (numChannels>=2)
      red(:,i) = traces(:, (i-1)*numChannels+2);
    end
    if (numChannels>=3)
      blue(:,i) = traces(:, (i-1)*numChannels+3);
    end
    if (numChannels>=4)
      magenta(:,i) = traces(:, (i-1)*numChannels+4);
    end
  end

  traces=size(traces,2)/numChannels

end
```

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#>"
%From FDAP v1.7

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
```

parse_cumeONdwells_length.m

```
%parse_cumeONdwells_length will separate dwellData for ON dwells into
%two separate matrices: "longONdwells," which contains the data for
those individual
%dwells that comprise long cumulative ON dwells, and "shortONdwells,"
which contains
%the data the for those individual dwells that comprise short
cumulative ON dwells.
%The input variable "threshold" is used to define short and long
cumulative ON dwells
%and should be entered in seconds. The input "ONdwells" should be the
%output from separateOnOffDwells, which is a four-column matrix of the
form
%[i FRET(t) FRET(t+1) n] where i is the unique identifier for each
%cumulative ON dwell and n is the number of time points spent within
each
%individual dwell.
%Written by DDM

function [longONdwells, shortONdwells] =
parse_cumeONdwells_length(ONdwells, threshold)

    fps = 10; %time resolution

    ONdwells = ONdwells(:, [1,2,3,4]);
    dwellno = ONdwells(:, 1);
    numdwells = length(unique(dwellno));

    longONdwells = ones(1,4);
    shortONdwells = ones(1,4);

    for t=1:numdwells
        rowind = find(dwellno == t);
        tempDwellData = ONdwells(rowind, :);
        n_cumedwell = sum(tempDwellData(:,4));
        if n_cumedwell >= threshold*fps
            longONdwells = [longONdwells; tempDwellData];
        else
            shortONdwells = [shortONdwells; tempDwellData];
        end
    end

    longONdwells = longONdwells(2:end, :);
    shortONdwells = shortONdwells(2:end, :);

end
```

parseTraces_T3binding.m

```
%Written by DDM

function [cy5, frameno] = parseTraces_T3binding(cy5)

ms=size(cy5);
rows=ms(1);
cols=ms(2);

IntensityThreshold = 1200;
TimeThreshold = 5; %Number of consecutive data points above intensity
threshold to be counted as a binding event.

keptTraces = ones(rows,1); %To store those traces that show a Cy5-T3
binding event
frameno = ones(1,1); %To store the corresponding frame number where
Cy5-T3 binding event is first observed.

for j = 1:cols
    for i=2:(rows - TimeThreshold + 1)
        if cy5(i:(i + TimeThreshold - 1),j) > IntensityThreshold
            keptTraces = [keptTraces, cy5(:,j)];
            frameno = [frameno, i - 1];
            break
        end
    end
end

cy5 = keptTraces(:,2:end);
frameno = frameno(2:end);

end
```

plotFRET.m

```
%calculates FRET and plots a histogram with 'n' bins.
%returns the histogram data as a two-column matrix
%with x-values in column 1 and y-values in column 2
%From FDAP v1.7

function FH = plotFRET(cy3, cy5, n)

    F=getFRET(cy3, cy5);

    X=linspace(-.2,1.2,n);

    H=hist(F(:),X);
        %normalize
    H=H/max(H);
```

```
    figure, bar(X, H, 'k'); %third argument indicates the color~ black
in this case
    axis([-0.2 1.2 0 max(H)*1.05]);
    xlabel('FRET');
    ylabel('normalized frequency');
    title('FRET histogram')

    FH=[X', H'];
```

plotFRETtraces.m

```
%From FDAP v1.7

function plotFRETtraces(cy3, cy5)

    F = getFRET(cy3, cy5);
    [cy3, cy5, labels]=removeLabels(cy3, cy5);

    ms=size(F);
    cols=ms(2);
    rows=ms(1);

        j=1;
        while j <= cols
            figure;
            for n = 1:9
                subplot(3,3,n);
                plot(F(:,j), 'b');
                axis([0 rows -0.2 1.2]);
                title(labels(j));
                j=j+1;
                if j > cols, break, end
            end
        end
```

plot_idealizedFRET_hist.m

```
%plot_idealizedFRET_hist will plot a 1D FRET histogram based on data
from
%the idealized traces. Input ONdwells can be dwellData post-processed
to
%any extent, so long as the second column is the current dwell's FRET
value
%and the fourth column is n (the number of time points spent in that
%dwell).
%Written by DDM

function FH = plot_idealizedFRET_hist(ONdwells, bins);
```

```
    ONdwells = ONdwells(:,[1,2,3,4]);
    FRETvalue = ONdwells(:,2);
    numtimepoints = ONdwells(:,4);
    numdwells = length(FRETvalue);
    i = 1;

    %Extract each data point of the viterbi path from dwellData
    for t=1:numdwells
        n = numtimepoints(t);
        F(i:i+n-1) = FRETvalue(t);
        i = i+n;
    end

    %Plot the data as a histogram
    X=linspace(-.2,1.2,bins);

    H=hist(F(:),X);
        %normalize
    H=H/max(H);
    figure, bar(X, H, 'k'); %third argument indicates the color~ black
in this case
    axis([-0.2 1.2 0 max(H)*1.05]);
    xlabel('FRET');
    ylabel('normalized frequency');
    title('FRET histogram')

    FH=[X', H'];
End
```

plotTDP.m

```
%From Jiangning Wang

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(:, 2);
    stop = dwellData(:, 3);

    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

        %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 tight square
    colorbar;
    hold on
    MINCOUNT = max(max(Z))*0.3
    MAXCOUNT = max(max(Z))*1.0;

        %minimum intensity: MINCOUNT
        %minimum intensity: MAXCOUNT
    caxis([MINCOUNT MAXCOUNT]);
```

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 seconds. If no cutoff time is given,
%no cutoff is applied.
%From FDAP v1.7

function [T, Y, H] = plotTimeFRET(cy3, cy5, FRETbins, Tbinsize,
cutoffT)
```

```
    %Exposure time = 100 ms
FPS = 10;

    %Minimum count shown
MINCOUNT = 1;
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

    %creation of interpolated data
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(4,3); jet]);
hold on

MAXCOUNT = max(max(H))*0.85

    %minimum intensity: MINCOUNT
    %maximum intensity: MAXCOUNT
caxis([MINCOUNT MAXCOUNT]);
axis([min(T) max(T) min(Y) max(Y)]);

colorbar; shading interp; axis tight square;

    %add labels
    xlabel('T (seconds)') ; ylabel('FRET') ; title('FRET Time Evolution
Histogram');
```

plotTimeFRET_ps.m

```
%Generates and plots a post-synchronized 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.

%default parameters:
%plotTimeFRET_ps(cy3x, cy5x, 24, 2, 20);
%Use postSync3 function- delete the traces that don't show the
%transition to the limit FRET value

%From FDAP v1.7

function [T, Y, H] = plotTimeFRET_ps(cy3, cy5, FRETbins, Tbinsize,
cutoffT)

        %Post-synchronization limit
    LIMIT = 0.2;

        %Exposure time = 100 msec
    FPS = 10;

        %Minimum count shown
    MINCOUNT = 1;
    RESOLUTION = 800;

    F = postSync_3(cy3, cy5, LIMIT);
    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

        %generate interpolated data
    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(4, 3); jet]);
    hold on

    MAXCOUNT = max(max(H))*0.85;

        %minimum intensity: MINCOUNT
        %maximum intensity: MAXCOUNT
    caxis([MINCOUNT MAXCOUNT]);
```

```
    axis([min(T) max(T) min(Y) max(Y)]);

    colorbar; shading interp; axis tight square;

        %add labels
    xlabel('T (seconds)') ; ylabel('FRET') ; title('FRET Time Evolution
Histogram') ;
```

plotTraces.m

```
%From FDAP v1.7

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 -1000 5000]);
                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);
```

plotTraces_3color.m

```
%plotTraces from FDAP v1.7, modified for 3-color data

function plotTraces_3color(cy3, cy5, cy2)

    [cy3, cy5, cy2, labels]=removeLabels_3color(cy3, cy5, cy2);

    ms=size(cy3);
    cols=ms(2);
    rows=ms(1);

    if nargin==3
        j=1;
        while j <= cols
            figure;
            for n = 1:9
                subplot(3,3,n);
                plot(cy3(:,j), 'g'), hold on, plot(cy5(:,j), 'r'), hold
on,
                plot(cy2(:,j), 'b');
                %axis([0 rows -1000 5000]);
                title(labels(j));
                j=j+1;
                if j > cols, break, end
            end
        end

    end

    [cy3, cy5, cy2]=addLabels_3color(cy3, cy5, cy2, labels);
```

PopDecay.m

```
%PopDecay is used to build a histogram of population decay for dwell
times
%or total fluorescence on times. For fluorescence on times, input
should
%be matrix H from getFLUORlifetime. Gives the total dwell time in
number of
%frames or seconds (t) and the corresponding number of traces (N) with
a
%dwell that is at least that long
%Written by DDM
```

```
function [N, t] = PopDecay(H,X)
    if nargin == 1 %Array H should contain dwell times in frames
        %Find longest dwell time (in frames)
        maxT=max(H(:));
        %Dwell time (in frames)
        t=0:1:maxT;

        l=length(t);
        %Construct histogram
        for i=1:l
            N(i)= length(find(H(:)>=t(i)));
        end

    elseif nargin == 2 %Array H should contain dwell times in seconds
        %Array X should contain frame timestamps
        %histogram parameters
        bins= 81;
        endtime= round(X(end));
        t= [0:15:1224]; %linspace(0,endtime,bins);
        %Find longest dwell (in sec)
        maxT= max(H(:));

        i=1;
        %Construct histogram
        while t(i)<= maxT
            N(i)= length(find(H(:)>=t(i)));
            i=i+1;
            if i==bins+1
                break;
            end
        end

        l=length(N);
        %Dwell times (in sec)
        t=t(1:l);

    end
```

postSync_3.m

```
%From FDAP v1.7

function F = postSync_3(cy3, cy5, LIMIT)

    Fs = getFRET2(cy3, cy5);

    DELETE=1;
    AVER=0; %Mean of AVER+1 consecutive frames must be above LIMIT
    KEEP=10; %First FRET event above LIMIT will occur at frame KEEP+1
in post-synchronized plot
```

```
    ms=size(Fs);
    cols=ms(2);
    rows=ms(1);

        %delete the first x frames (x=DELETE-1), right now it doesn't
delete
        %anything, the whole trace is kept.
    Fs=Fs(DELETE:ms(1), :);

    ms=size(Fs);
    cols=ms(2);
    rows=ms(1);
    k=1;

    for j = 1:cols
        for i = 1:(rows-AVER)
            if mean(Fs(i:i+AVER,j))>LIMIT
                    %make timepoint (i-KEEP) time zero if it is
positive
                newTrace = Fs(max((i-KEEP), 1):rows, j);
                F(:, k) = NaN;
                F(1:length(newTrace), k) = newTrace;
                k=k+1;
                    %skip to next trace
              break;
            end
        end
    end
```

purifyOnOffDwell.m

```
%combine all the successive high and low FRET states(defined by the
threshold)
%From Jiangning Wang, modified by DDM

function dwellData = purifyOnOffDwell(dwellData, Filter)

    if nargin==2
        KEY = Filter;
    else
        KEY = 0.20;
    end

    dwellData = dwellData(:,[1,2,3,4]);
    fileno = dwellData(:,1);

    sizedw = size(dwellData);
    selind = 1;

    for t = 1:sizedw(1)-1
```

```
        if diff(fileno(t:t+1)) %If last dwell of trace
            continue;
        elseif dwellData(t,2)>KEY %ON dwells
            if dwellData(t,3)>KEY && dwellData(t+1,2)>KEY
                selind = [selind; t];
                dwellData(t+1,4) = dwellData(t,4)+dwellData(t+1,4);
            end
        else
            if dwellData(t,3)<=KEY && dwellData(t+1,2)<=KEY %OFF dwells
                selind = [selind; t];
                dwellData(t+1,4) = dwellData(t,4)+dwellData(t+1,4);
            end
        end
    end

    lensel = length(selind);

    selind = selind(2:lensel);

    mergeind = selind;
    lenmerg = length(mergeind);

    rowind = [1:sizedw(1)];
    for k = 1:lenmerg
        rowind = removeEntry(rowind, mergeind(k));
    end
    dwellData = dwellData(rowind, :); %Keep all 4 columns in case trace
number needs to be identified
```

removeEntry.m

```
%From Jiangning Wang

function [labels] = removeEntry(labels, key)

    mc=length(labels);
    no=find(labels==key);

    if no

        labels=[labels(1:no-1) labels(no+1:mc)];

    else

        disp(['Trace ' int2str(key) ' does not exist'])

    end
```

removeLabels.m

```
%From FDAP v1.7

function [cy3, cy5, labels]=removeLabels(cy3, cy5)

        %remove labels
    ms=size(cy3);
    labels=cy3(1,:);
    cy3=cy3(2:ms(1), :);
    cy5=cy5(2:ms(1), :);
```

removeLabels_1color.m

```
%same as removeLabels from FDAP v1.7, modified for 1-color data

function [cy3, labels]=removeLabels_1color(cy3)

        %remove labels
    ms=size(cy3);
    labels=cy3(1,:);
    cy3=cy3(2:ms(1), :);
```

removeLabels_3color.m

```
%same as removeLabels from FDAP v1.7, modified for 3-color data

function [cy3, cy5, cy2, labels]=removeLabels_3color(cy3, cy5, cy2)

        %remove labels
    ms=size(cy3);
    labels=cy3(1,:);
    cy3=cy3(2:ms(1), :);
    cy5=cy5(2:ms(1), :);
    cy2=cy2(2:ms(1), :);
```

removeTrace.m

```
%From FDAP v1.7

function [cy3, cy5] = removeTrace(cy3, cy5, selected)
    ms=size(cy3);
    labels=cy3(1,:);
    no=find(labels==selected);
```

```
    if no
        cy3=[cy3(:,1:no-1) cy3(:,no+1:ms(2))];
        cy5=[cy5(:,1:no-1) cy5(:,no+1:ms(2))];

    else
        disp(['Trace ' int2str(selected) ' does not exist'])

    end
```

removeTrace_3color.m

```
%same as removeTrace from FDAP v1.7, modified for 3-color data

function [cy3, cy5, cy2] = removeTrace_3color(cy3, cy5, cy2, selected)
    ms=size(cy3);
    labels=cy3(1,:);
    no=find(labels==selected);

    if no
        cy3=[cy3(:,1:no-1) cy3(:,no+1:ms(2))];
        cy5=[cy5(:,1:no-1) cy5(:,no+1:ms(2))];
        cy2=[cy2(:,1:no-1) cy2(:,no+1:ms(2))];

    else
        disp(['Trace ' int2str(selected) ' does not exist'])

    end
```

RollingAvg.m

```
%Written by DDM

function cy5 = RollingAvg(cy5, span);

ms = size(cy5);
rows = ms(1);
cols = ms(2);
labels = cy5(1,:);

cy5=cy5(2:rows,:);

for j=1:cols
    cy5(:,j) = smooth(cy5(:,j), span);
end

cy5=[labels;cy5];
end
```

saveTraces.m

```
%From FDAP v1.7

function saveTraces(cy3, cy5, filen)
    merged=mergeCy(cy3, cy5);
    save(filen, 'merged',  '-ASCII');
```

separateCy.m

```
%From FDAP v1.7

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
```

separateOnOffDwells.m

```
%separateOnOffDwells parses dwellData into two separate matrices, one
containing
%the ON dwells and the other containing the OFF dwells. The threshold
FRET value
%separating ON and OFF dwells is the input parameter "KEY". Input
"dwellData"
%should be in the four column format, i.e. [fileno FRET(t) FRET(t+1) n]
and
%should be generated by using getRawDwell_all. Contiguous ON and OFF
dwells within
%the same trace will be identified, and all dwells comprising the same
%cumulative ON dwell will be marked with the same index i. The output
%matrix ONdwells will have a four-column format different from that of
the
%input matrix, i.e. [i FRET(t) FRET(t+1) n].
%Written by DDM

function [ONdwells, OFFdwells] = separateOnOffDwells(dwellData, KEY)

    dwellData = dwellData(:,[1,2,3,4]);
    fileno = dwellData(:,1);
    sizedw = size(dwellData);
```

```
    t = 1; %current dwell
    i = 1; %new index for cumulative ON dwells
    j = 1; %new index for cumulative OFF dwells

    ONdwells = ones(1,4);
    OFFdwells = ones(1,4);

    while t <= sizedw(1);
        if dwellData(t,2)>KEY %ON dwells
            tempONdwells = ones(1,4);
            while dwellData(t,2)>KEY
                tempONdwells = [tempONdwells; i, dwellData(t,2),
dwellData(t,3), dwellData(t,4)];
                if t==sizedw(1) | diff(fileno(t:t+1))
                    t=t+1;
                    break;
                else
                    t=t+1;
                end
            end
            tempONdwells = tempONdwells(2:end, :);
            ONdwells = [ONdwells; tempONdwells];
            i=i+1;

        elseif dwellData(t,2)<=KEY %OFF dwells
            tempOFFdwells = ones(1,4);
            while dwellData(t,2)<=KEY
                tempOFFdwells = [tempOFFdwells; j, dwellData(t,2),
dwellData(t,3), dwellData(t,4)];
                if t==sizedw(1) | diff(fileno(t:t+1))
                    t=t+1;
                    break;
                else
                    t=t+1;
                end
            end
            tempOFFdwells = tempOFFdwells(2:end, :);
            OFFdwells = [OFFdwells; tempOFFdwells];
            j=j+1;
        end
    end

    ONdwells = ONdwells(2:end, :);
    OFFdwells = OFFdwells(2:end, :);
End
```

SplitData.m

```
%SplitData will split FRET on times or IF2 residency times into three
random groups so
%that mean and standard deviation can be calculated. Input should be G
```

```
%containing FRET on times or H containing IF2 residency times (units of
%seconds). Returns three separate vectors containing one third of the
randomized data.
%Written by DDM

function [A, B, C] = SplitData(G);

numtraces = length(G);

randomizedData = randsample(G, length(G));

numSplitTraces = floor(numtraces/3);

A = randomizedData(1:numSplitTraces);

B = randomizedData((numSplitTraces+1):(numSplitTraces*2));

C = randomizedData(((numSplitTraces*2)+1):(numtraces));
```

SplitPathData.m

```
%SplitPathData takes pathData from vbFRET and splits the idealized
%trajectories into three separate arrays, each containing the same
number
%of idealized trajectories. Trace sorting is done randomly so that
output arrays
%can be treated as three independent datasets. Input data should be
concatenated
%idealized traces where col1=tracelabel and col2=FRETvalue. Traces
should be
%labeled consecutively as 1->n, where n is the total number of traces.
%Output traces will have the same format, i.e. with 2 cols, and the
traces
%will be relabeled as 1->n/3.
%Written by DDM

function [A,B,C] = SplitPathData(pathData);

    traceno = pathData(:, 1);
    FRET = pathData(:, 2);

    numtraces = max(traceno);

    rdmsample = randsample(numtraces,numtraces); %randomize traces

    lensample = length(rdmsample);

    numsplittraces = floor(lensample/3); %number of traces to be put in
A,B,and C
```

```
    extratraces = rem(lensample,3); %leftover traces after dividing
total by 3

    A=ones(1,2);
    B=ones(1,2);
    C=ones(1,2);

    i = 1;

    for n = 1:numsplittraces  %Separate path data into A, B, and C

        rowind = find(traceno == rdmsample(i));
        A = [A; pathData(rowind,:)];

        i=i+1;

        rowind = find(traceno == rdmsample(i));
        B = [B; pathData(rowind,:)];

        i=i+1;

        rowind = find(traceno == rdmsample(i));
        C = [C; pathData(rowind,:)];

        i=i+1;
    end

    %Place extra traces (if total number of traces was not divisible by
3)
    if extratraces == 0
    elseif extratraces == 1
        rowind = find(traceno == rdmsample(i));
        A = [A; pathData(rowind,:)];

    elseif extratraces == 2
        rowind = find(traceno == rdmsample(i));
        A = [A; pathData(rowind,:)];

        i=i+1;

        rowind = find(traceno == rdmsample(i));
        B = [B; pathData(rowind,:)];
    end

    A= A(2:end,:);
    B= B(2:end,:);
    C= C(2:end,:);

    %Reformat trace labels for matrix A
    tracenoA = A(:,1);
    lenA = length(tracenoA);

    rows = [1:lenA];
```

```
    for n=1:lenA-1
        if tracenoA(n)== tracenoA(n+1)
            rows = removeEntry(rows,n);
        end
    end
    oldtraceidA = tracenoA(rows);  %Original trace labels from vbFRET
    numtracesA = length(oldtraceidA);

    for n=1:numtracesA  %Relabel traces from 1->n
        rowind = find(tracenoA == oldtraceidA(n));
        A(rowind,1) = n;
    end

  %Reformat trace labels for matrix B
    tracenoB = B(:,1);
    lenB = length(tracenoB);

    rows = [1:lenB];

    for n=1:lenB-1
        if tracenoB(n)== tracenoB(n+1)
            rows = removeEntry(rows,n);
        end
    end
    oldtraceidB = tracenoB(rows);
    numtracesB = length(oldtraceidB);

    for n=1:numtracesB
        rowind = find(tracenoB == oldtraceidB(n));
        B(rowind,1) = n;
    end

    %Reformat trace labels for matrix C
    tracenoC = C(:,1);
    lenC = length(tracenoC);

    rows = [1:lenC];

    for n=1:lenC-1
        if tracenoC(n)== tracenoC(n+1)
            rows = removeEntry(rows,n);
        end
    end
    oldtraceidC = tracenoC(rows);
    numtracesC = length(oldtraceidC);

    for n=1:numtracesC
        rowind = find(tracenoC == oldtraceidC(n));
        C(rowind,1) = n;
    end
```

Appendix F – “R” scripts

Three “R” scripts were used to estimate errors for the lifetimes calculated using dwell time analysis, as outlined in Section 5.5.1.6. These scripts were written by Dr. Jiangning Wang and modified by myself, and are presented below.

ka_bootstrap_sample_dwells.R

```
#Script for bootstrapping OFF dwell lifetimes obtained from single
exponential fits to dwell time population decay histograms.
#Also returns ka values in units of uM-1sec-1
#Sampling of dwells with replacement

#Script from Jiangning Wang, modified by DDM

#Input data are saved in .txt file in matrix form. There are four
columns, the 1st is 'Trace ID'
#the 2nd is 'starting FRET' the 3rd is 'ending FRET' and the fourth is
'number of frames'

#Set working directory

setwd(dir = "C:/Documents and Settings/Administrator/Desktop/R-codes")

#Import data, data is saved in folder named 'data' as a .txt file
dwellData = read.table(file = "data/test.txt", header = TRUE)

assign("traceID", dwellData[,1])
assign("startF", dwellData[,2])
assign("endF", dwellData[,3])
assign("Ldwell", dwellData[,4])

#Set boundaries for start FRET value
start_low <- -0.25
start_high <- 0.2

selection <- which(dwellData[,2] > start_low & dwellData[,2] <
start_high)
dwelltimes <- dwellData[selection, 4]
maxT <- max(dwelltimes)
t <- c(1:1:(maxT-1))
time <- t*0.1
population <- rep(0, length(t))
for(i in 1:length(t)){
population[i] <- length(which(dwelltimes >= t[i]))
}

Decaycurvedata <- as.data.frame(cbind(time, population))
```

```
#Initial guess on the parameters (y_off(yo), Asym(As) and lifetime(lt))
for the exponential decay
iguessyo = 0
iguessAs = 3000
iguesslt = 4

#fit the raw data with single exponential decay, return the fitting
result
mod1 <- nls( population ~ y_off + Asym*exp( -time/lifetime ),
                  data = Decaycurvedata,
                  start = list( y_off = iguessyo, Asym = iguessAs,
lifetime = iguesslt ),
                  trace = TRUE )

#the fitting result output
summary( mod1)
lifetime = coef(mod1)[3]

#calculate ka in units of uM-1sec-1. Subunit concentration given in
units of nM
subunitconc = 20
ka = 1000/(as.numeric(coef(mod1)[3])*subunitconc)
ka

#plot the decay fitting graph
par(lwd = 3, cex.axis = 1.5, cex.lab = 1.5, mar = c(5,5,2,2))
plot(population ~ time, lwd = 3, pch = 20, bty = "n",xlab = "time",
ylab = "population")
lines(predict(mod1)~time, lwd = 3, pch = 20, bty = "n", xlab = "time",
ylab = "population")


#Bootstrap to obtain SD of lifetime and rate
n = length(dwelltimes)
lifetime_bs = c()
ka_bs = c()
number = (1:1:n)

#number of bootstrap datasets to construct
nstep = 1000

for(i in 1:nstep) {
  numbernew = sample(number, n, replace=TRUE)
  dwelltimesnew = dwelltimes[numbernew]
  maxT <- max(dwelltimesnew)
  t <- c(1:1:(maxT-1))
  time <- t*0.1
  population <- rep(0, length(t))
  for(j in 1:length(t)){
    population[j] <- length(which(dwelltimesnew >= t[j]))
  }

  Decaycurvedata <- as.data.frame(cbind(time, population))
```

```
  mod2 = nls( population ~ y_off + Asym*exp( -time/lifetime ),
                   data = Decaycurvedata,
                   start = list( y_off = iguessyo, Asym = iguessAs,
lifetime = iguesslt ),
                   trace = TRUE )

 lifetime_bs = append(lifetime_bs, coef(mod2)[3])
 ka_bs = append(ka_bs, 1000/(as.numeric(coef(mod2)[3])*subunitconc))
}

#plot the rate frequency histogram
hist(ka_bs)

#compare the mean of the bootstrapped parameters to those estimated
from sample and get sd of bootstrapped parameters
lifetime
mean(lifetime_bs)
sd(lifetime_bs)

ka
mean(ka_bs)
sd(ka_bs)
```

ON_dwell_bootstrap_sampleDwells_singleExp.R

```
#Script for bootstrapping ON dwell lifetimes obtained from single
exponential fits to dwell time population decay histograms.
#Also returns the kd values
#Sampling of dwells with replacement

#Script from Jiangning Wang, modified by DDM

#Input data are saved in .txt file in matrix form. There are four
columns, the 1st is 'Trace ID'
#the 2nd is 'starting FRET' the 3rd is 'ending FRET' and the 4th is
'number of frames'

#Set working directory
setwd(dir = "C:/Documents and Settings/Administrator/Desktop/R-codes")

#Import data, data is saved in folder named 'data' as a .txt file
dwellData = read.table(file = "data/AUG_plusIF3_60nM_data.txt", header
= TRUE)

assign("traceID", dwellData[,1])
assign("startF", dwellData[,2])
assign("endF", dwellData[,3])
assign("Ldwell", dwellData[,4])

#Set boundaries for start FRET value
start_low <- 0.2
```

```
start_high <- 1.2

selection <- which(dwellData[,2] > start_low & dwellData[,2] <
start_high)
dwelltimes <- dwellData[selection, 4]
maxT <- max(dwelltimes)
t <- c(1:1:(maxT-1))
time <- t*0.1
population <- rep(0, length(t))
for(i in 1:length(t)){
population[i] <- length(which(dwelltimes >= t[i]))
}

Decaycurvedata <- as.data.frame(cbind(time, population))

#Initial guess on the parameters for the exponential decay (y_off(yo),
Asym(As), lifetime(lt))
iguessyo = 0
iguessAs = 1800
iguesslt = .8

#fit the raw data with exponential decay, return the fitting result
mod1 <- nls( population ~ y_off + Asym*exp( -time/lifetime ),
                  data = Decaycurvedata,
                  start = list( y_off = iguessyo, Asym = iguessAs,
lifetime = iguesslt ),
                  trace = TRUE )

#the fitting result output
summary(mod1)
lifetime = coef(mod1)[3]
kd = 1/as.numeric(lifetime)


#plot the decay fitting graph
par(lwd = 3, cex.axis = 1.5, cex.lab = 1.5, mar = c(5,5,2,2))
plot(population ~ time, lwd = 3, pch = 20, bty = "n",xlab = "time",
ylab = "population")
lines(predict(mod1)~time, lwd = 3, pch = 20, bty = "n", xlab = "time",
ylab = "population")


#Bootstrap to obtain SD of lifetime and rate
n = length(dwelltimes)
lifetime_bs = c()
kd_bs = c()
number = (1:1:n)

#number of bootstrap datasets to construct
nstep = 1000

for(i in 1:nstep) {
  numbernew = sample(number, n, replace=TRUE)
  dwelltimesnew = dwelltimes[numbernew]
```

```
  maxT <- max(dwelltimesnew)
  t <- c(1:1:(maxT-1))
  time <- t*0.1
  population <- rep(0, length(t))
  for(j in 1:length(t)){
    population[j] <- length(which(dwelltimesnew >= t[j]))
  }

  Decaycurvedata <- as.data.frame(cbind(time, population))

  mod2 = nls( population ~ y_off + Asym*exp( -time/lifetime ),
                  data = Decaycurvedata,
                  start = list( y_off = iguessyo, Asym = iguessAs,
lifetime = iguesslt ),
                  trace = FALSE )

 lifetime_bs = append(lifetime_bs, coef(mod2)[3])
 kd_bs = append(kd_bs, 1/as.numeric(coef(mod2)[3]))
}

#compare the mean of the bootstrapped parameters to those estimated
from sample and get sd of bootstrapped parameters
lifetime
mean(lifetime_bs)
sd(lifetime_bs)

kd
mean(kd_bs)
sd(kd_bs)
```

ON_dwell_bootstrap_sampleDwells_doubleExp.R

```
#Script for bootstrapping ON dwell lifetimes obtained from double
exponential fits to dwell time population decay histograms.
#Also returns the corresponding kd values
#Sampling of dwells with replacement

#Script from Jiangning Wang, modified by DDM

#Data are saved in .txt file in matrix form. There are four columns,
the 1st is 'Trace ID'
#the 2nd is 'starting FRET' the 3rd is 'ending FRET' and the 4th is
'number of frames'

#Set working directory
setwd(dir = "C:/Documents and Settings/Administrator/Desktop/R-codes")

#Import data, data is saved in folder named 'data' as a .txt file
dwellData = read.table(file = "data/AUU_minIF3_plusTRNAi_20nM.txt",
header = TRUE)
```

```
assign("traceID", dwellData[,1])
assign("startF", dwellData[,2])
assign("endF", dwellData[,3])
assign("Ldwell", dwellData[,4])

#Set boundaries for start FRET value
start_low <- 0.2
start_high <- 1.2

selection <- which(dwellData[,2] > start_low & dwellData[,2] <
start_high)
dwelltimes <- dwellData[selection, 4]
maxT <- max(dwelltimes)
t <- c(1:1:(maxT-1))
time <- t*0.1
population <- rep(0, length(t))
for(i in 1:length(t)){
population[i] <- length(which(dwelltimes >= t[i]))
}

Decaycurvedata <- as.data.frame(cbind(time, population))

#Initial guess on the parameters for the double exponential decay
(y_off(yo), Asym1(As1), lifetime1(lt1), Asym2(As2), lifetime2(lt2))
iguessyo = 0
iguessAs1 = 150
iguesslt1 = 1
iguessAs2 = 350
iguesslt2 = 10

#fit the raw data with double exponential decay, return the fitting
result
mod1 <- nls( population ~ y_off + Asym1*exp( -time/lifetime1 ) +
Asym2*exp( -time/lifetime2 ),
                  data = Decaycurvedata,
                  start = list( y_off = iguessyo, Asym1 = iguessAs1,
lifetime1 = iguesslt1, Asym2 = iguessAs2, lifetime2 = iguesslt2 ),
                  trace = TRUE )

#the fitting result output
summary(mod1)
Asym1_percent = (coef(mod1)[2])/((coef(mod1)[2])+(coef(mod1)[4]))
Asym2_percent = (coef(mod1)[4])/((coef(mod1)[2])+(coef(mod1)[4]))
Asym1_percent
Asym2_percent

#ON dwell times
lifetime1 = coef(mod1)[3]
lifetime2 = coef(mod1)[5]

#Dissociation rates
kd1 = 1/as.numeric(lifetime1)
kd2 = 1/as.numeric(lifetime2)
```

```
#plot the decay fitting graph
par(lwd = 3, cex.axis = 1.5, cex.lab = 1.5, mar = c(5,5,2,2))
plot(population ~ time, lwd = 3, pch = 20, bty = "n",xlab = "time",
ylab = "population")
lines(predict(mod1)~time, lwd = 3, pch = 20, bty = "n", xlab = "time",
ylab = "population")


#Bootstrap to obtain SD of lifetimes and amplitudes
n = length(dwelltimes)
Asym1_percent_bs = c()
lifetime1_bs = c()
kd1_bs = c()
Asym2_percent_bs = c()
lifetime2_bs = c()
kd2_bs = c()

number = (1:1:n)

#number of bootstrap datasets to construct
nstep = 1000

for(i in 1:nstep) {
  numbernew = sample(number, n, replace=TRUE)
  dwelltimesnew = dwelltimes[numbernew]
  maxT <- max(dwelltimesnew)
  t <- c(1:1:(maxT-1))
  time <- t*0.1
  population <- rep(0, length(t))
  for(j in 1:length(t)){
    population[j] <- length(which(dwelltimesnew >= t[j]))
  }

  Decaycurvedata <- as.data.frame(cbind(time, population))

  mod2 = nls( population ~ y_off + Asym1*exp( -time/lifetime1 ) +
Asym2*exp( -time/lifetime2 ),
                  data = Decaycurvedata,
                  start = list( y_off = iguessyo, Asym1 = iguessAs1,
lifetime1 = iguesslt1, Asym2 = iguessAs2, lifetime2 = iguesslt2 ),
                  trace = FALSE )

 Asym1_percent_new = (coef(mod2)[2])/((coef(mod2)[2])+(coef(mod2)[4]))
 Asym1_percent_bs = append(Asym1_percent_bs, Asym1_percent_new)
 Asym2_percent_new = (coef(mod2)[4])/((coef(mod2)[2])+(coef(mod2)[4]))
 Asym2_percent_bs = append(Asym2_percent_bs, Asym2_percent_new)
 lifetime1_bs = append(lifetime1_bs, coef(mod2)[3])
 kd1_bs = append(kd1_bs, 1/as.numeric(coef(mod2)[3]))
 lifetime2_bs = append(lifetime2_bs, coef(mod2)[5])
 kd2_bs = append(kd2_bs, 1/as.numeric(coef(mod2)[5]))
}

#compare the mean of the bootstrapped parameters to those estimated
from sample and get sd of bootstrapped parameters
```

```
Asym1_percent
mean(Asym1_percent_bs)
sd(Asym1_percent_bs)

lifetime1
mean(lifetime1_bs)
sd(lifetime1_bs)

kd1
mean(kd1_bs)
sd(kd1_bs)

Asym2_percent
mean(Asym2_percent_bs)
sd(Asym2_percent_bs)

lifetime2
mean(lifetime2_bs)
sd(lifetime2_bs)

kd2
mean(kd2_bs)
sd(kd2_bs)
```