Role of 16S Ribosomal RNA in Translation Initiation

Dissertation

Presented in Partial Fulfillment of the Requirements for the Degree Doctor of Philosophy in the Graduate School of The Ohio State University

By

Daoming Qin, B. S.

Ohio State Biochemistry Program

The Ohio State University

2011

Dissertation Committee:

Professor Kurt Fredrick, Advisor
Professor Michael Ibba
Professor Venkat Gopalan
Professor Thomas J. Magliery
Abstract

In the cell, translation is the process by which polypeptides are synthesized on the ribosome based on the genetic information encoded in the messenger RNA (mRNA). The first step of translation, initiation, limits the overall rate of protein synthesis and establishes the translational reading frame. Initiation involves recognition of the start codon by the initiator tRNA (fMet-tRNAfMet in bacteria) in the ribosomal peptidyl (P) site. The correct start codon is selected primarily based on the complementarity of codon-anticodon pairing. Genetic and biochemical studies have suggested that initiation factors IF1 and IF3 negatively regulate translation initiation to promote its accuracy. 16S ribosomal RNA (rRNA) undoubtedly plays a role in translation initiation, but precisely how it participates in the process remains unclear. In this work we identified and characterized the 16S rRNA nucleotides that play an important role in start codon selection in *Escherichia coli*. We also uncovered the structural changes of 16S rRNA upon start codon recognition. Finally, we provided evidence that IF3 kinetically controls initiation to promote the fidelity of this process without enhancing the intrinsic specificity of the programmed P site.

During initiation, fMet-tRNA enters the P site, where 16S rRNA makes extensive contacts with the tRNA. We first screened a number of P-site mutations for
their effects on start codon selection. Our results showed that most of these P-site
mutations, while generally reducing the efficiency of translation, increase the
stringency of start codon selection. We also demonstrated these mutations (e.g.
G1338U, A790U, A790C) confer a defect in fMet-tRNA binding, which likely makes
initiation depend more heavily on the cognate codon-anticodon pairing. Two
exceptions, G1338A and A790G, decrease the fidelity of start codon selection.
G1338A confers a higher affinity for tRNA^{Met}, which could partially compensate for
the mismatches in the codon-anticodon helix and thereby allow initiation from non-
canonical start codon. These data provide evidence that the affinity of the initiator
tRNA for the 30S P site is tuned to balance efficiency and accuracy of initiation.
Unlike G1338A, 30S subunits harboring A790G show defect in IF3 binding, consistent
with the proposal that 790 loop is part of IF3 binding site. We suspect that reduced IF3
binding explains the loss of fidelity in that case.

Next, we employed random mutagenesis in a genetic screen and identified
additional 16S rRNA mutations that increase initiation from non-canonical start codons.
One cluster of mutations maps to the 790 loop, further supporting the previous finding
of A790G. Another cluster localizes to the neck region of the 30S subunit, and these
mutations are suspected to act by altering the dynamics of the head domain. Most
interestingly, a cluster maps to helix 44 (h44) just ‘down’ from the A site, a region
known to be distorted by IF1. Several h44 mutations (e.g. A1413C) are predicted to
change the non-canonical pair to a Watson-Crick pair, suggesting that they may
decrease fidelity of translation initiation by altering the conformation of h44. Indeed,
our subsequent work showed that these h44 mutations stimulate the second step of initiation (50S docking) most dramatically in the presence of a non-canonical start codon, which may explain for their defects in start codon recognition in vivo.

To reveal possible structural changes of 16S rRNA upon start codon recognition, we used chemical probing methods to monitor 16S rRNA in canonical and non-canonical 30S initiation complex. Of ~110 nucleotides targeted, 6 showed altered reactivity in response to the nature of start codon. One of them is A1408, which lies in h44. In particular, the high reactivity of A1408 attributed to the presence of IF1 is specifically reduced in 30S initiation complexes containing AUG, even though the level of IF1 binding remains unchanged. Additionally, either mutation A1413C or streptomycin similarly reduces the reactivity of A1408, even in the presence of a non-canonical start codon. In light of the fact that both A1413C and streptomycin can stimulate premature 50S docking, we propose that the docking is controlled in part by a conformational switch in the 1408 region. IF1 stabilizes the A1408 region in a docking unfavorable conformation, which is reversed upon start codon recognition. Another group of the nucleotides (A532, A794, G926 and G1338) lie in or near the P site. These nucleotides, which are protected by fMet-tRNA in the canonical 30S initiation complexes, are either unprotected or less protected in the non-canonical 30S initiation complexes. The loss of protections suggests that P site is largely unoccupied in the presence of non-canonical start codon. We propose that in this situation, fMet-tRNA associates to the 30S subunits in a liable manner. The last position U701, a binding site
of IF3, is deprotected in canonical 30S initiation complexes, which indicates that IF3 is destabilized.

Finally, to investigate start codon selectivity on the 30S subunit, we compared the effects of start codon on the stability of ternary complex (30S·tRNA·mRNA) by toeprinting. Our results showed that the selectivity factor (AUG versus AUC) is about 100 in the absence of initiation factors, similar to the presence of all initiation factors (IF1, 2 and 3). However, in the presence of IF1 and IF2, both AUG complex and AUC complex are stabilized, with AUC having a more pronounced effects. Under this condition, the selectivity factor drops to 4. These studies provide evidence that 30S has a robust intrinsic selectivity for a canonical start codon. IF2 promotes rapid initiation by stimulating fMet-tRNA binding, thereby shifting the equilibrium for 30S initiation complex formation towards the right. A negative consequence is that the stabilization masks the intrinsic fidelity of the 30S subunit. The role of IF3 is to shift the equilibrium back to the left by increasing $k_{off}$ of fMet-tRNA. Hence, the intrinsic selectivity of the programmed P site can be utilized while the rate of initiation is enhanced.
Dedicated to my family, especially to my wife, Chenguang Wang, father Shende Qin, mother Suhua Qin and brother Gaoming Qin for their support without reservation throughout these years.
Acknowledgements

First and foremost, I owe my deepest gratitude to my adviser Dr. Kurt Fredrick for encouragement, guidance and support over the years. Dr. Fredrick helped me greatly in developing professional skills and insights in research. He also kindly helped me improve my scientific communication and writing skills. It has been a great opportunity to work with Dr. Fredrick.

I would like to thank my committee Dr. Michael Ibba, Dr. Venkat Gopalan, and Dr. Thomas Magliery for serving on my thesis committee and supervising the progress of thesis project.

I am very grateful to Dr. Sarah Walker for her friendship and many helpful discussions in my projects.

I would like to express my gratitude to Nimo A. Abdi for preparing P-site mutations (Chapter 2 and 3) and Qi Liu for providing technique supports in the NMIA probing experiments (Chapter 4). I would also like to thank all other members of the Fredrick lab (past and present), Dr. Xiaofen Zhong, Kevin McGarry, Aishwarya Devaraj, Sean McCloy, Joshua Lesring, Jodie Lee, Mai Dang and Rohan Balakrishnan for being nice with me and helping a lot with my English.
I thank Dr. Ross E. Dalbey and Dr. Jill Rafael-Fortney for their kind support as the Chair of Ohio State Biochemistry Program.

In the end, I want to express my deepest gratitude to my wife, Chenguang Wang, my parents Shende Qin and Suhua Qin, and my brother Gaoming Qin. My wife sacrificed herself to join me in U.S. and is always supportive in my life. My parents and brother are very considerate and offered me a lot of happiness and encouragements over the years.
VITA

DECEMBER 16, 1980 ..........Born in Guilin, China

2004…………………………….B. S. Biology, Peking University, China

2004-present…………………Graduate research associate, The Ohio State

University

Publications


Fields of Study

Major Field: Biochemistry
# Table of Contents

| Page |
|------------|---|
| Abstract | ii |
| Dedication | vi |
| Acknowledgments | vii |
| Vita | ix |
| List of Tables | xv |
| List of Figures | xvi |
| List of Symbols/Abbreviations | xviii |
| Chapters: | |
| 1. Introduction | 1 |
| 1.1 Translation initiation pathway in bacteria | 2 |
| 1.2 Components of the translation initiation pathway | 4 |
| 1.2.1 The ribosome | 5 |
| 1.2.2. The initiator tRNA | 7 |
| 1.2.3. The translation initiation region in mRNA | 8 |
| 1.2.4. Initiation factor | 12 |
| 1.3. Fidelity of translation initiation | 23 |
| 1.3.1. Start codon selection | 23 |
| 1.3.2. Initiation with elongator tRNA | 26 |
| 2. Characterization of 16S rRNA mutations that decrease the fidelity of translation initiation | 29 |
| 2.1 Introduction | 29 |
2.2. Materials and Methods ........................................................................................................32
2.2.1. Specialized ribosome strains ............................................................................................32
2.2.2. Activity of mutant ribosomes in vivo ..............................................................................33
2.2.3. Δ7 prrn strains ....................................................................................................................34
2.2.4. Purification of 30S subunits ...............................................................................................35
2.2.5. P-site tRNA binding experiments ....................................................................................36
2.2.6. IF3 binding experiments ..................................................................................................37
2.3. Results and Discussion .......................................................................................................38
2.3.1. G1338A and A790G increase translation from non-canonical start codons ...........38
2.3.2. G1338A increases the affinity of tRNA^{fMet} for the 30S subunit P site ..................42
2.3.3. A790G decreases the affinity of IF3 for the 30S subunit ..............................................45
2.3.4. Mutations ΔC1400 and G529U do not confer fidelity phenotypes in the specialized ribosome system ..................................................................................................................48
2.4. Conclusions .........................................................................................................................52

3. Control of translation initiation involves a factor-induced rearrangement of helix 44 of 16S ribosomal RNA ................................................................................................................53
3.1. Introduction ........................................................................................................................53
3.2. Materials and Methods ......................................................................................................55
3.2.1. Bacterial strains ..............................................................................................................55
3.2.2. Genetic screen .................................................................................................................56
3.2.3. β-galactosidase assay ......................................................................................................56
3.2.4. Purification of 35S-Labeled IF1 ....................................................................................56
3.2.5. Binding assays ...............................................................................................................57
3.3. Results 55
3.3.1. Identification of 16S rRNA mutations that increase translation from AUC ..................58
3.3.2. Mutant ribosomes can support the growth of E. coli ...................................................64
3.3.3. A790G and U789C confer a defect in IF3 binding .......................................................65
3.3.4. A1410G and G886A confer a defect in IF1 binding ........................................66
3.3.5. A1408G causes spurious initiation in vivo .........................................................70
3.3.6. Watson-Crick base pairs at 1413-1487 and 1414-1486 decrease initiation fidelity ...............................................................................................................................70
3.3.7. Most 30S P-site mutations increase the stringency of start codon selection .................................................................................................................................73
3.3.8. Effect of 30S mutations on tRNAf^{Met} binding ......................................................75
3.4. Discussion .................................................................................................................76

4. Role of the 30S subunit in start codon selection .........................................................82
   4.1. Introduction .............................................................................................................82
   4.2. Materials and Methods ..........................................................................................84
      4.2.1. Bacterial strains ............................................................................................84
      4.2.2. Preparation of fMet-tRNAf^{Met} and mRNA ..................................................84
      4.2.3. Initiation factors ............................................................................................85
      4.2.4. Ribosomal subunits .......................................................................................85
      4.2.5. Light scattering to monitor 50S docking .........................................................86
      4.2.6. Chemical modification of 30SIC .....................................................................86
      4.2.7. Primer extension ............................................................................................87
      4.2.8. Toeprinting .....................................................................................................88
      4.2.9. Micro-con assays ............................................................................................89
   4.3. Results ....................................................................................................................89
      4.3.1. Mutations in h44 stimulate 50S docking .........................................................89
      4.3.2. Structural differences in 30S initiation complexes containing the start codon AUG versus AUC .............................................................................................................94
      4.3.3. A1408 exhibits reduced reactivity in the presence of AUG ............................94
      4.3.4. A1408 exhibits reduced reactivity in the presence of mutation A1413C ....98
      4.3.5. A1408 exhibits reduced reactivity in the presence of streptomycin ............98
      4.3.6. Protections attributed to P-site tRNA depend on AUG.................................101
4.3.7. Protection of G369 from hydroxyl radicals is increased in the presence of AUG ................................................................. 104
4.3.8. U701 is less protected from NIMA in the presence of AUG...... 105
4.3.9. Protection of the anticodon of fMet-tRNA is only evident in 30S initiation complexes containing AUG ................................................................. 106
4.3.10. Comparing the extent of fMet-tRNA binding in both complexes........ 107
4.3.11. Stability of the P codon in 30S initiation complexes containing AUG versus AUC ................................................................. 109
4.4. Discussion ................................................................................. 111
  4.4.1. A conformational switch in helix 44 helps control 50S docking........ 111
  4.4.2. The P site is largely unoccupied in the 30SIC(AUC) ....................... 113
  4.4.3. IF3 acts in start codon selection kinetically to reveal the inherent differences of matching and nonmatching codon-anticodon pairing in the P site ........................................................................................................... 115

5. Conclusions and Future Directions.................................................. 118

List of reference ................................................................................. 124
List of Tables

Table
2.1 Effects of 16S rRNA mutations G1338A and A790G on translation of mRNA containing non-canonical start codons ................................................................. 41
2.2. Summary of effects conferred by 16S rRNA mutations G1338A and A790G ...... 44
2.3. Spurious initiation observed upon expression of 16S rRNA containing ΔC1400 and G529U depends on recognition of the reporter mRNA by endogeneous (wild type) ribosomes ........................................................................................................ 50
3.1. Mutations in 16S rRNA recovered in a screen for increased translation from AUC ..................................................................................................................... 61
3.2. Effects of 16S rRNA mutations isolated in the genetic screen on translation from ACG, AUC and AUG ................................................................. 62
3.3. Effects of 16S rRNA mutations on ligand binding .............................................. 69
3.4. Effects of engineered mutations in h44 on translation of lacZ from ACG, AUC and AUG .................................................................................................................. 72
3.5. Effects of P-site mutations on translation from AUC and AUG ...................... 74
4.1. Summary of the apparent rates of 50S docking ................................................. 93
4.2. Qualification of chemical reactivity of 16S rRNA nucleotides in various 30S complexes ............................................................................................................. 97
4.3. Summary of stability of 30S complexes containing AUG versus AUC .......... 110
# List of Figures

## Figure

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>Simplified translation initiation pathway in bacteria</td>
<td>3</td>
</tr>
<tr>
<td>1.2</td>
<td>Overall shape of the 30S subunit</td>
<td>6</td>
</tr>
<tr>
<td>1.3</td>
<td>Interactions between the 16S rRNA and fMet-tRNA</td>
<td>7</td>
</tr>
<tr>
<td>1.4</td>
<td>The secondary structure of fMet-tRNA in <em>E. coli</em></td>
<td>7</td>
</tr>
<tr>
<td>1.5</td>
<td>Structures of the initiation factors</td>
<td>21</td>
</tr>
<tr>
<td>2.1</td>
<td>Mutation G1338A increases the affinity of tRNA&lt;sup&gt;Met&lt;/sup&gt; for the 30S subunit P site</td>
<td>44</td>
</tr>
<tr>
<td>2.2</td>
<td>Mutation A790G decreases the affinity of IF3 for the 30S subunit</td>
<td>47</td>
</tr>
<tr>
<td>2.3</td>
<td>Detection of mutant 16S rRNA in ribosomal fractions separated by sucrose gradient sedimentation</td>
<td>51</td>
</tr>
<tr>
<td>3.1</td>
<td>Nucleotides identified in this study as important for initiation fidelity cluster to h44, the 790 loop and the neck region</td>
<td>63</td>
</tr>
<tr>
<td>3.2</td>
<td>Effects of 16S rRNA mutations on the relative level of translation from AUC and ACG</td>
<td>64</td>
</tr>
<tr>
<td>3.3</td>
<td>Effect of mutations in the 790 loop on the binding of IF3-AF to the 30S subunit</td>
<td>66</td>
</tr>
<tr>
<td>3.4</td>
<td>Effects of 16S rRNA mutations on the binding of IF1 or tRNA&lt;sup&gt;Met&lt;/sup&gt; to the 30S subunit</td>
<td>68</td>
</tr>
<tr>
<td>3.5</td>
<td>Effects of alternative base pairs at 1410-1490, 1413-1487 and 1414-1486 on the relative level of translation from AUC and ACG</td>
<td>73</td>
</tr>
<tr>
<td>4.1</td>
<td>Nucleotides showing different chemical reactivity in 30SIC (AUG) and 30SIC(AUC) locate in h44, P site and IF3 binding site</td>
<td>92</td>
</tr>
<tr>
<td>4.2</td>
<td>16S rRNA mutation A1413C stimulates association of the 50S subunit to the 30SIC(AUC)</td>
<td>93</td>
</tr>
<tr>
<td>4.3</td>
<td>Effects of canonical base paring in the P site on the chemical reactivity of</td>
<td></td>
</tr>
</tbody>
</table>
4.4. Effects of canonical base paring in the P site on the chemical reactivity of G530 and A532 ................................................................. 100
4.5. P-site nucleotides (G1338, G926 and A794) and anticodon of the initiator tRNA are efficiently protected only in 30SIC(AUG) ......................................................... 103
4.6. The co-presence of IF3 and mRNA protect G926 ........................................ 104
4.7. Canonical base paring in the P site de-protects U701 ........................................ 106
4.8. The initiator tRNA associates with the 30SIC(AUC) to a similar extent, but in a less stable manner compared to that in 30SIC(AUG) ........................................ 108
## List of Symbols/Abbreviations

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>Adenosine or aminoacyl</td>
</tr>
<tr>
<td>aa</td>
<td>amino acid</td>
</tr>
<tr>
<td>ATP</td>
<td>Adenosine 5'-triphosphate</td>
</tr>
<tr>
<td>BSA</td>
<td>bovine serum albumin</td>
</tr>
<tr>
<td>bp</td>
<td>base pair</td>
</tr>
<tr>
<td>C</td>
<td>cytidine</td>
</tr>
<tr>
<td>°C</td>
<td>degree Celsius</td>
</tr>
<tr>
<td>CCase</td>
<td>terminal tRNA nucleotidyl transferase</td>
</tr>
<tr>
<td>CTP</td>
<td>Cytidine 5'-triphosphate</td>
</tr>
<tr>
<td>d</td>
<td>deoxy</td>
</tr>
<tr>
<td>Da</td>
<td>dalton (unit)</td>
</tr>
<tr>
<td>DMS</td>
<td>Dimethyl sulfide</td>
</tr>
<tr>
<td>DMSO</td>
<td>dimethyl sulfoxide</td>
</tr>
<tr>
<td>DNA</td>
<td>deoxyribonucleic acid</td>
</tr>
<tr>
<td>dNTP</td>
<td>deoxynucleotide 5'-triphosphate</td>
</tr>
<tr>
<td>DTT</td>
<td>dithiothreitol</td>
</tr>
<tr>
<td>E</td>
<td>exit</td>
</tr>
</tbody>
</table>

xviii
<table>
<thead>
<tr>
<th>Term</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>E. coli</td>
<td><em>Escherichia coli</em></td>
</tr>
<tr>
<td>EDTA</td>
<td>ethylenediaminetetraacetic acid</td>
</tr>
<tr>
<td>EF-G</td>
<td>elongation factor-G</td>
</tr>
<tr>
<td>EF-Tu</td>
<td>elongation factor-Tu</td>
</tr>
<tr>
<td>EM</td>
<td>electron microscopy</td>
</tr>
<tr>
<td>fMet</td>
<td>formyl methionine</td>
</tr>
<tr>
<td>g</td>
<td>gram</td>
</tr>
<tr>
<td>G</td>
<td>guanosine</td>
</tr>
<tr>
<td>GDP</td>
<td>Guanosine 5’-diphosphate</td>
</tr>
<tr>
<td>GTP</td>
<td>guanosine 5’-triphosphate</td>
</tr>
<tr>
<td>h</td>
<td>hour (unit)</td>
</tr>
<tr>
<td>HEPES</td>
<td>N-[2-hydroxyethyl] piperazine-N’-[2-ethanesulfonic acid]</td>
</tr>
<tr>
<td>IF</td>
<td>initiation factor</td>
</tr>
<tr>
<td>IPTG</td>
<td>isopropyl-β-D-thiogalactoside</td>
</tr>
<tr>
<td>$k_{app}$</td>
<td>apparent rate</td>
</tr>
<tr>
<td>$K_D$</td>
<td>binding constant</td>
</tr>
<tr>
<td>$k_{off}$</td>
<td>dissociation rate</td>
</tr>
<tr>
<td>$k_{on}$</td>
<td>association rate</td>
</tr>
<tr>
<td>LB</td>
<td>Luria-Bertani</td>
</tr>
<tr>
<td>m</td>
<td>mili</td>
</tr>
<tr>
<td>M</td>
<td>molar</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Definition</td>
</tr>
<tr>
<td>--------------</td>
<td>------------</td>
</tr>
<tr>
<td>Met</td>
<td>L-methionine</td>
</tr>
<tr>
<td>Met-RS</td>
<td>methionyl-tRNA synthetase</td>
</tr>
<tr>
<td>min</td>
<td>minute</td>
</tr>
<tr>
<td>mRNA</td>
<td>messenger RNA</td>
</tr>
<tr>
<td>MTF</td>
<td>methionyl tRNA transformylase</td>
</tr>
<tr>
<td>n</td>
<td>nano</td>
</tr>
<tr>
<td>ND</td>
<td>not determined</td>
</tr>
<tr>
<td>NMIA</td>
<td>N-methylisatoic anhydride</td>
</tr>
<tr>
<td>nt</td>
<td>nucleotide</td>
</tr>
<tr>
<td>P</td>
<td>peptidyl</td>
</tr>
<tr>
<td>PAGE</td>
<td>polyacrylamide gel electrophoresis</td>
</tr>
<tr>
<td>PCR</td>
<td>polymerase chain reaction</td>
</tr>
<tr>
<td>RNA</td>
<td>ribonucleic acid</td>
</tr>
<tr>
<td>RT</td>
<td>reverse transcriptase</td>
</tr>
<tr>
<td>sec</td>
<td>second</td>
</tr>
<tr>
<td>TIR</td>
<td>Translation initiation region</td>
</tr>
<tr>
<td>Tris-HCl</td>
<td>tris-(hydroxymethyl) aminomethane hydrochloride</td>
</tr>
<tr>
<td>tRNA</td>
<td>transfer RNA</td>
</tr>
<tr>
<td>U</td>
<td>unit or uridine</td>
</tr>
<tr>
<td>UTP</td>
<td>uridine 5’-triphosphate</td>
</tr>
<tr>
<td>WT</td>
<td>wild-type</td>
</tr>
<tr>
<td>µ</td>
<td>micro</td>
</tr>
</tbody>
</table>

xx
Chapter 1: Introduction

In bacteria, faithful recognition of the canonical start in the mRNA is critical for establishing the translational reading frame of protein synthesis. The process is generally regulated to discriminate against non-canonical start codons (codons other than AUG, GUG and UUG), whose utilization will likely result in alternative and unwanted, potentially harmful products in the cell. In vivo, single substitutions of the canonical start codon AUG (e.g. AUU, AUC and ACG) typically reduce the level of gene expression by 100-fold (129,130,133,164). In bacteria, there are other determinants in the mRNA that help the translation machinery find the correct initiation codon. For example, a purine rich sequence (termed Shine-Dalgano or SD) (148), which is located upstream of start codon in the mRNA, pairs with the 3’ end of 16S rRNA (termed anti-Shine-Dalgano, or ASD) to facilitate the positioning of the start codon in the decoding P site (21,75,79,160). In addition, initiation factors IF1 and IF3 have been implicated in facilitating selection of the start codon from genetic and biochemical studies (84,102,133,164). However, how these factors enhance the stringency of start codon selection remains unclear. Structural and biochemical studies have demonstrated that ribosome (especially the rRNA part) makes extensive contacts
with both the initiator tRNA and mRNA in the P site (12,25,80,141). Some of these interactions are suggested to regulate the fidelity of initiation (86). In this study, we combined genetic and biochemical approaches to elucidate the role of 16S rRNA in start codon selection.

1.1. Translation initiation pathway in bacteria

Translation initiation is a highly regulated process that occurs with an average rate of 0.5 s\(^{-1}\) and limits the overall rate of protein synthesis (168). It involves recognition of the start codon by the initiator tRNA in the ribosomal P site. Three initiation factors IF1, IF2 and IF3 facilitate the process in a manner that depends on a cognate start codon (Figure 1.1) (review in (57,58,87,151)). This process can be divided into two major steps. In the first step, the initiator tRNA (fMet-tRNA) and mRNA bind to the 30S subunit. The initiation factors act together to increase the \(k_{on}\) for fMet-tRNA, while IF1 and IF3 also increase \(k_{off}\) (8). With the establishment of SD-ASD interaction and codon-anticodon pairing, the 30S initiation complex (30SIC) is formed. In the second step, the 50S subunit docks to the 30SIC, which triggers hydrolysis of GTP by IF2 and leads to subsequent release of the factors. At this stage, the 70S initiation complex (70SIC) is formed and ready for the elongation cycle.

Although the two-step model is widely accepted, the mechanistic details in each step remain obscure and are sometimes controversial. Regarding binding of tRNA and mRNA, early studies suggested an ordered pathway with either mRNA (11,171) or fMet-tRNA (74,112) binding first. Other studies proposed a random-ordered model.
(56), which seems to be most accepted today. Another dispute centered on whether IF2 functions as a tRNA carrier in solution, like eukaryotic initiation factor eIF2 (181) or recruits the initiator tRNA on the ribosome (23,59,101). Additionally, the order of release of the initiation factors after fulfilling their jobs remains unclear. For example, does IF3 dissociate before 50S docking (7) or during the formation of 70SIC (3,42,54)? It is often hard to resolve these controversies, since apparently contradictory results can come from different experimental conditions (e.g. temperature, buffer) or use of specific components (e.g. different mRNA, heterogeneous factors). However, the possibility that multiple pathways exist in the cell cannot be ruled out, and may be more common for complex assembly processes (165).

**Figure 1.1. Simplified translation initiation pathway in bacteria.**
Translation initiation occurs at two major steps with the help of three initiation factors (IF1, IF2 and IF3). In the first step, fMet-tRNA and mRNA bind to the 30S subunit in a random order, followed by codon-anticodon pairing, which leads to formation of the 30S initiation complex (30SIC). In the second step, the 50S subunit docks to the 30SIC, accompanied by the release of initiation factor, which forms the 70S initiation complex (70SIC).
Besides gaining components during assembly, the translation initiation complex also undergoes conformational changes during the whole initiation process. It has been proposed that during 30SIC formation, a rate-limiting rearrangement occurs to facilitate the adjustment of mRNA and tRNA leading to codon-anticodon interaction (56). This may involve movement of mRNA to facilitate the start codon positioning in the P site (24,75,82). It has also been proposed that fMet-tRNA may first bind in a codon-independent way and then be adjusted to facilitate the pairing of codon and anticodon (167,180). The fMet-tRNA repositioning may also occur after 50S docking to pave the way for dipeptide formation (85). However, the nature of these events remains unclear.

Clues about the structural basis of rearrangements of initiation complexes come from cryo-electron microscopy (cryo-EM) reconstitution studies. Movements inferred from these studies include subunit-subunit rotation and movement of IF2 on the ribosome upon GTP hydrolysis (3,110). Interestingly, two studies suggest that prior to stable 50S docking, fMet-tRNA occupies a site distinct from the classic P/P binding site (P/P, which stands for binding in 30S P site and 50S P site at the same time) (3,152). The relevance of these rearrangements with respects to translation initiation needs to be established.

1.2. Components of the translation initiation pathway

As mentioned above, the ribosome, fMet-tRNA, mRNA and initiation factors (IF1, IF2 and IF3) are the major components involved in the translation initiation
pathway. Below, I will briefly describe the key characteristics of each component relevant to the fidelity of translation initiation.

1.2.1. The ribosome

The ribosome is a ribonucleoprotein machine, composed of a small subunit (30S in bacteria) and a large subunit (50S in bacteria). Each subunit has three tRNA binding sites; namely, the aminoacyl (A), peptidyl (P) and exit (E) sites. The 30S subunit can be divided into four regions: head, neck, body, and platform (Figure 1.2). As mentioned above, fMet-tRNA binds the ribosomal P site, where extensive contacts between the ribosome and tRNA occur. Recent high-resolution structures of ribosome reveal the details of these interactions (Figure 1.3) (12,25,80,141). (1) The P site codon is held in place with C1402, C1403 and U1498 of 16S rRNA. (2) The phosphate of nucleotide +1 (AUG, with A being +1 position; this numbering scheme is used throughout) is H bonded to N1 and N2 of G926. (3) The base pair at position +3 is stabilized by stacking interactions from C1400 and G966. (4) Two G:C pairs (positions 30-40 and 29-41) from the anticodon stem are involved in Type I and II A-minor interactions with G1338 and A1339, respectively. (5) On the other side, the ribose of position 38 contacts 790 of 16S rRNA. (7) Apart from RNA-RNA contacts, P-site tRNA also interacts with the C termini of ribosomal proteins S9 and S13.
Figure 1.2. Overall shape of the 30S subunit.
The 30S subunit is viewed from the subunit interface (image based on (141), PDB 2J02). Only the 16S rRNA is shown and ribosomal proteins are omitted. The domains (head, neck, platform, and body) are labeled.

Figure 1.3. Interactions between the 16S rRNA and fMet-tRNA.
The image is based on (141) (PDB: 2J02). 16S rRNA P-site nucleotides are shown in light purple. fMet-tRNA is shown in red, while mRNA is shown in green. Codon-anticodon base pairs are highlighted with hydrogen bond (yellow).
1.2.2. The initiator tRNA

In bacteria, the initiator tRNA is tRNA\textsuperscript{Met}. It possesses several unique structural features not found in the elongator tRNAs (Figure 1.4) (169). These features include: (1) a mismatch at position 1 and 72; (2) three conserved consecutive G:C base pair at the anticodon stem; and (3) purine-11-pyrimidine-24 in the dihydrouridine stem, instead of pyrimidine-11-purine-24, which is typical in elongator tRNA. The tRNA\textsuperscript{Met} is aminoacylated by methionyl tRNA synthetase (MetRS), which recognizes the anticodon as the major identity element (138). The resulting Met-tRNA\textsuperscript{Met} is then formylated by methionyl tRNA transformylase (MTF) (39). The mismatch or weak base pair between 1-72 is critical for MTF recognition (61,137,169).

![Figure 1.4. The secondary structure of fMet-tRNA in E. coli.](image)

The unique features of fMet-tRNA are highlighted in green; the anticodon is in red.
The unique sequence and/or structural features of fMet-tRNA guide its exclusive usage in initiation. The fMet-tRNA specifically enters the 30S P site by two anchors. At the anticodon stem loop (ASL), the 30S P-site makes extensive interactions with fMet-tRNA as described above. At the acceptor end, IF2 recognizes the formyl group and holds the otherwise free CCA end (152). The three consecutive G:C pairs at the anticodon stem seem to be essential for initiation (64,143). G:C base pairs, into which A1338 and G1339 dock (12,25,80,141), are optimal for stabilizing these types of tertiary contacts, providing an explanation for the importance of the conserved G:C pairs in initiation (40). These G:C pairs could be also associated with a unique conformation of the ASL, which may act as another specificity determinant (64,143).

The unique features of the initiator tRNA also prevents its usage in elongation. For example, formylation and mismatch at 1-72 also block its binding to elongation factor Tu (EF-Tu)(144).

1.2.3. The translation initiation region in mRNA

Initiation rate of a specific mRNA is determined by its translation initiation region (TIR), the region surrounding the start codon. Several elements in the TIR contribute to its ability to promote initiation, as described below:

The initiation codon

In bacteria, AUG is the predominant start codon, while UUG and GUG are quite commonly used (41). The frequency with which the three canonical start codons
are used varies between different organisms. For example, in *E. coli*, 83% of the genes begin with AUG, 14% with GUG, and 3% with UUG. In *Borrelia burgdorferi*, 69% of the genes use AUG, 9% use GUG, and 22% use UUG. In highly expressed genes, the AUG is not preferentially utilized, suggesting that high-level translation does not require AUG. For example, *tufA*, which encodes EF-Tu, one of the most abundant proteins, starts with GUG (182). AUG, GUG and UUG, which are not discriminated against by the translation initiation machinery, are considered canonical start codons. (133,164). Despite the fact that changing a canonical start codon to non-canonical start codon generally reduces the gene expression, non-canonical start codons can sometimes be used. A well-known example is *infC*, the gene encoding IF3, which starts with AUU (135). The non-canonical start codon is key to auto-regulation of *infC* as described below (20).

**Shine-Dalgarno sequence**

The Shine-Dalgarno (SD) sequence, first identified by Shine and Dalgarno, is a purine rich sequence (e.g. AGGA), located 6-12 nucleotides upstream of the start codon (148,150). The SD pairs with the 3′ end of the 16S rRNA, termed anti-Shine-Dalgarno (ASD), and facilitates positioning of the start codon in the P site (21,160). The SD-ASD interaction was demonstrated *in vivo* in 1987 by Hui and de Boer, who showed that complementary substitutions in the ASD can rescue mutations in the SD (72). The length of SD-ASD helix ranges from 3-9 nucleotides (nt), averaging 6 nt (139). The length or strength of the predicted SD-ASD does not correlate with
translation efficiency in a simple way (89). For example, SD sequence longer than average can inhibit translation initiation (78,173). Recently, the SD-ASD interaction has been visualized in various high-resolution ribosome structures (75,79,185,186). These studies localized the SD-ASD helix between the head and platform domains of the 30S subunit, in a cavity formed by multiple rRNA helices (h28, h37, h23a and h26). Interestingly, the SD-ASD appears to preferentially bind two positions, SD_in with SD-ASD close to ribosomal protein S2, and SD_out, with SD-ASD close to S18 (75). The occupancy of these sites may represent intermediate states during the initiation process (see below).

It is noteworthy that besides the sequence of SD, the spacing between the SD and start codon also affects the efficiency of translation initiation (29,170). In *E. coli* genes, the spacing ranges from 6-12 nt (150). Recently, our lab has provided biochemical evidence that short spacing destabilizes the codon-anticodon interaction in the P site (38).

Although the role of SD-ASD has been well documented, analysis of 162 completed prokaryotic genomes showed that the percentage of genes containing SD sequences (%SD) varies from 11.6% to 90.8%, which supports the idea that the SD sequence is loosely conserved in prokaryotic genes (28). In *E. coli*, the %SD is estimated to be around 60% (28,92,158). In theory, the SD-less mRNAs include leaderless mRNAs, which lack 5’ untranslated regions (UTR), or non-SD-led mRNAs, which have 5’ UTR without SD sequence. In *E. coli*, one example of non-SD-led gene is *rpsA*, which encodes ribosomal protein S1. Despite the absence of a SD element, S1
is a highly expressed protein (16). The lack of a SD in many genes suggests an alternative mechanism to recruit ribosomes to the mRNA. In the view of evolution, the SD-independent pathway of leaderless translation seems to be the ancestral, based on the fact that leaderless genes can be efficiently translated in all three domains (91).

Other auxiliary elements

Besides the SD, other sequence elements have been proposed to contribute to the binding of the ribosome to mRNA. One is a pyrimidine-rich sequence 5’ of the SD, which appears to be recognized by ribosomal protein S1 (17,142). Another is a region downstream from the start codon, named the downstream box (DB) (156). DB was original proposed to pair to 16S rRNA bases 1469-1483 and thereby promote translation initiation, in a manner mimicking the SD-ASD interaction. However, recent studies have disproved this hypothesis (107,117). Other mechanisms may also exist for recruitment of ribosomes to mRNA (154). All these ribosome recruitment mechanisms may be complementary to one another. For example, in Bacillus, which lacks S1, the SD-ASD pathway seems to be more dominant (89.4%) (41,92).

Secondary structure in the TIR

In order for efficient translation initiation to occur, the start codon and SD should be accessible for pairing with the 30S subunit and initiator tRNA. Stable secondary structure formed within the TIR can block translation initiation by occluding the SD and or start codon. The ability of mRNA secondary structure to inhibit
translation has in many cases been exploited by the cell to regulate gene expression. For example, regulatory proteins, or metabolites or temperature can influence secondary structure of the TIR and thereby affect translation (46,51,113). How does secondary structure in the TIR affect the kinetics of initiation? Fluorescence-based studies suggest that 30SIC assembly occurs in two steps (162). First, the 30S subunit binds to a single-stranded region of mRNA in a kinetically labile manner. Next, the TIR is unfolded in a way that requires the SD, a canonical start codon, IF2·GTP and the initiator tRNA, which results in the 30SIC with stably bound mRNA.

The adjustment of TIR on the ribosome

As described above, the ribosome initially interacts with the mRNA in a nonspecific manner. An anchoring SD-ASD interaction can form if the SD is accessible and the 30S subunit to “find” the SD is thought to be factor-independent (24). It has been proposed based on site-directed crosslinking studies that the start codon is not yet positioned at the decoding P site (82). This is supported by chemical probing studies, which showed that the region after position +5 is not protected, leading to hypothesis that the region downstream of SD is free, out of the mRNA channel (73). The initiation factors do not affect the SD-ASD interaction nor the affinity of the mRNA, however, they (mostly IF3) help the mRNA proceed to the channel and relocate the start codon closer to the P site (82).

1.2.4. Initiation factor
Unlike eukaryotic translation initiation, which requires more than dozen protein factors (76), only three major factors are involved in bacterial translation initiation, kinetically controlling the process to optimize efficiency and accuracy (reviewed in (151)). They bind to the 30S subunit in a cooperative way, with high affinity ($K_D$ in nM range). Considering their estimated concentration in the cell is $>1$ mM (70,71), it is believed that 30S subunits unengaged in translation are saturated with IFs.

**Initiation factor 1**

Initiation factor 1 (IF1) is an 8.2 kDa protein, the smallest of the three initiation factors. IF1 is encoded by *infA* gene and homologous proteins are present in three domains of life (eIF1A in Eukarya and aIF1A in Archaea). The structure of IF1 has been determined by NMR spectroscopy (Figure. 1. 5A) (145). It is characterized by a five-stranded b-barrel with a flexible loop between stands 3 and 4. The fold of IF1 resembles an oligomer-binding (OB) motif found in other OB-fold family proteins (e.g., CspA, aspartyl-tRNA synthetase and S1).

The crystal structure of the 30S subunit bound with IF1 has also been solved (26). IF1 binds the 30S A site, in a cleft between the 530 loop, helix 44 and ribosomal protein S12. The binding of IF1 flips out bases A1492 and A1493 in the A site. At the same time, it disrupts base pairs around A1413$\circ G1487$, causes a long-range conformational change along h44 and induces a subtle global movement of the major domains of the subunit. The structural data correlate well with chemical probing data.
IF1 protects G530, A1492 and A1493, and increases the reactivity of A1413 and A1487 (34,106).

Although it is essential for cell survival, the specific function of IF1 in initiation remains obscure. Several activities have been attributed to IF1. (1) IF1 stimulates the association and dissociation rates of the 70S ribosome without affecting the equilibrium constant (53). (2) IF1 works synergistically with IF3 to split and recycle the 70S ribosome, thus increasing the subunit pool (121). (3) IF1 increases the rate of IF2-dependent binding of fMet-tRNA to the 30S subunit, stimulates 30SIC formation, and helps IF2 discriminate elongator tRNA (7,8,180). (4) With IF2, it catalyses the drop off of short peptidyl tRNA from the ribosome (77). (5) IF1 fine tunes the effects of IF2 and IF3 and promotes the unfolding of structured TIRs (162). (6) IF1 possesses RNA chaperone activity (31,50,176). (7) IF1 negatively regulates translation initiation, inhibiting 50S docking in the presence of certain TIRs (102,163).

It is worth noting that most of these activities of IF1 are associated with IF2 or IF3, which probably originates from cooperative binding of the three factors on the 30S subunit. Also, the effects of IF1 seem to vary in response to different TIRs (102,163). Although highly conserved, no IF1 homolog has been detected in mammalian mitochondria. It has been suggested that an extra-extended loop from mitochondrial IF2, which inserts into space normally occupied by IF1, functionally replaces IF1 in this organelle (49). This observation provides some evidence that the essential role of IF1 involves translation initiation and IF2 function.
Initiation factor 2

Initiation factor (IF2), a guanosine nucleotide binding protein, is the largest of the initiation factors. It is encoded by the infB gene. Two major isoforms IF2-α (97.3 kDa) and IF2-β (79.7 kDa) exist at similar level in E. coli (69,125,134). Although either form alone is able to support cell survival, the presence of both forms is required for optimal growth (136). No obvious difference in function has been detected between the two forms (27). More recently, a third isoform IF2-γ was resolved, which differs from IF2-β only by the absence of 7 amino acids at the N terminus (114). These isoforms are translated from three different in-frame start sites at codon 1(AUG), codon 158(GUG) and codon 165(AUG) of infB, respectively.

IF2 can be divided into 5 structural domains: the N-terminal region, a central G domain, domain II, domain III and domain IV (also named C-terminal domains) (Figure 1.5B). The portion from the G domain through the C domain is conserved, while the N terminal region is highly variable, serving as a functionally dispensable anchor to the ribosome (27,155). The overall structure of IF2 has not been successfully determined yet. However, the structures of homologous protein aIF5B in its free and bound (GDP or GDPNP) form from the archaeon M. thermoautotrophicum have been solved (132)(Figure 1.5B). This factor lacks the corresponding N terminal region in E. coli. It has an overall “chalice” shape, with the “cup” (domains G, II and III) connecting through a long a helix to the “base” (domain IV). The relative spatial arrangement of the G domain and domain II are similar to other GTPases (e.g., EF-Tu and EF-G).
In contrast to the other initiation factor, IF2 stably interacts with both ribosomal subunits. While chemical probing studies have shown that IF2 binds to 50S at the α-sarcin loop and H89, similar approaches to define IF2 binding sites in the 30S subunit were unsuccessful (83,106,174). This raised the possibility that IF2 contacts the 30S subunit primarily via protein-protein interactions. Hitherto, no high-resolution structure is available for ribosome bound with IF2. However, cryo-EM structures containing 70S ribosome with IF1, IF2, non-hydrolysable GTP analog GDPNP, fMet-tRNA and mRNA have been described that allow IF2 to be modeled on the ribosome (3).

Accordingly, the N terminal region contacts the shoulder of the 30S subunit, close to G423 of 16S rRNA. The G domain of IF2 lies in the vicinity of the GTPase association center (GAC) of the 50S subunit, a location also occupied by the G domains of other translation factors (EF-Tu and EF-G). Domain II interacts with the 30S subunit, contacting 16S rRNA at positions 55-59, 357, and 366-369. Domain III contacts ribosomal protein L14. Domain IV cover a region of the 50S subunit including rRNA bases 2550-2556, 2601-2602 and ribosomal protein L16. These extensive contacts between IF2 and the ribosome result in burial of 2600 Å² on the 50S subunit and 1300 Å² on the 30S subunit, explaining how IF2 can stimulate subunit association. Other available cryo-EM structures (70S ribosome with IF2, GDPNP, tRNA and mRNA; 70S ribosome with IF2, GDP, fMet-tRNA and mRNA; and the 30S subunit with IF1, IF2, GTP, fMet-tRNA and mRNA) show somewhat different positions of IF2 on the ribosome (110,152). These different positions may reflect movements of IF2 during different stages of translation initiation. For example, it has been proposed that upon
transition from GTP to GDP bound form, IF2 moves relative to the ribosome, resulting in reduced contacts with fMet-tRNA and the GAC of the 50S subunit. These rearrangements lead to a “ready-to-leave” state, which rationalizes the role of GTP hydrolysis in release of IF2.

As a central protein in translation initiation, IF2 is conserved among three kingdoms (aIF5B in Archaea and eIF5B in Eukarya). One function of IF2 is to stimulate the binding of fMet-tRNA to the ribosome (7,8,23,101,180). Biochemical studies have shown that the C domain of IF2 interacts directly with the acceptor end of fMet-tRNA (60,157). Mutations in the C domain can suppress the defect of non-formylated Met-tRNA\(^{fMet}\), probably by increasing the affinity of the C domain for the acceptor end of Met-tRNA\(^{fMet}\) (159). Cryo-EM structures of the 30S subunit with IF1, IF2, GTP, fMet-tRNA and mRNA have helped explain how IF2 promotes fMet-tRNA binding on the ribosome (152). IF2 anchors the otherwise free fMet-tRNA acceptor end, confirming earlier predictions about the basis of cooperative binding (7,8,23,101,180). Besides promoting the binding of fMet-tRNA, IF2 also plays a role in adjustment of fMet-tRNA for during the late stage of 70SIC formation (82). Another essential function of IF2 is to promote the association of subunits in the presence of the initiator tRNA (3,5,53). Recent studies indicate that IF2 exists on the 30S subunit in two conformations, one of which supports 50S docking. The fMet-tRNA interacts with IF2 on the 30S subunit and shifts the equilibrium to the docking favorable state, thereby stimulating 70SIC formation (120,187).
It is generally agreed that fast subunit association requires IF2 in the GTP bound form, and upon 50S docking GTP hydrolysis occurs rapidly, which triggers the subsequent release of IF2 (6,110). Dissociation of IF2 is essential prior to elongation, since IF2 will sterically block EF-Tu-GTP-aminoacyl-tRNA ternary complex due to their overlapping binding sites. However, some studies provide contradicting evidences about whether the transition from initiation to elongation is affected by the nature of the guanosine (GTP vs GDP) (167). Further experiments are needed to resolve this controversy about the role of GTP hydrolysis.

**Initiation factor 3**

Initiation factor 3 (IF3) is the only bacterial initiation factor, which lacks orthologs in other domains of life. IF3 is a 20 kDa protein, encoded by \textit{infC} gene. The structures of IF3 have been solved by X-ray crystallography and NMR (13,47,48). It is composed of two domains (denoting IF3N and IF3C), which are bridged by a ~45 Å lysine-rich flexible linker. IF3N and IF3C are approximately equal in size and both consist of a compact \(\alpha/\beta\) fold (Figure 1.5C).

Co-crystal structures with IF3 bound to the 30S subunit at its primary site are unavailable. However the general location of IF3 binding has been derived from other approaches. Chemical probing studies place IF3 close to h23 and h24 in the platform of central domain (106). A ~27\(\AA\) resolution cryo-EM reconstitution, generally consistent with the observed chemical protection pattern, localized IF3 at the subunit interface, covering a region of the platform (98). Using an approach combining hydroxyl radical
footprinting, site-directed hydroxyl radical probing and structural modeling, Dallas et al assigned the positions of IF3C and IF3N on the 30S subunit. They placed the IF3C at the interface of the platform (contacting h23, h24 and h45), in agreement with cryo-EM structure, while IF3N was modeled near the E site between S7 and S11, instead of near the neck, as the cryo-EM study suggested (35). Additionally, Dallas et al placed IF3C and IF3N on opposite sides of the initiator tRNA and proposed mutually exclusive interactions between IF3 and H69 of the 50S subunit, explaining the “anti-association” activity of IF3. More recently, different cryo-EM structures lent support to the topology of IF3 proposed by Dallas and Noller, whereas another probing study suggested an alternative orientation, with IF3N contacting the 790 loop and IF3C near the 700 region (3,42).

IF3 is an essential protein, which has been strongly implicated in the fidelity of translation initiation. Several activities have been attributed to IF3: (1) IF3 acts as an anti-association factor, shifting the equilibrium of 30S+50S ⇔ 70S to the left (53); (2) IF3 acts during ribosome recycling, preventing subunit reassociation after subunit splitting (122); (3) IF3 works synergistically with IF1 to split of ribosome containing mRNA (121); (4) IF3 increases both $k_{on}$ and $k_{off}$ for fMet-tRNA binding, facilitating 30SIC formation (8,56,65,84); (5) IF3 adjusts the position of a mRNA from a “stand-by” site to a “decoding” site during translation initiation (24,82); (5) IF3 kinetically controls initiation to ensure the fidelity of initiation site selection, favoring complexes containing fMet-tRNA and canonical start codons (7,64,65,115,133,164).
The two-domain structure of IF3 led to speculation that each function can be assigned to the individual domains (36,98). However, a biochemical study showed that the isolated C domain can perform all activities of IF3, when added in amounts to compensate for its reduced affinity. In contrast, the N domain can neither bind the subunit nor display detectable activities (123). Hence, it was proposed that all the activities reside in IF3C, while IF3N simply provides an additional anchor, which facilitates the recycling of IF3 on and off the ribosome. In line with this proposal, a recent study combining chemical probing and fast kinetics demonstrated that during association, IF3 binds to the 30S subunit first via its C domain and then via N domain; in dissociation, the interaction of N domain is lost first, followed by the C domain (42).
Figure 1. 5. Structures of the initiation factors

All the structures are generalized by PYMOL (Delano Scientific Freeware) using deposited pdb file from Protein Data Bank (PDB). (A) Structure of IF1 (PDB ID: 1AH9); (B) Top left, structure of aIF5B bound with GDPNP, an IF2 homolog (PDB ID: 1G7T); Top right, structure of EF-Tu bound with GDPNP (PDB ID: 1TTT); both G domain is highlighted in magenta, which GDPNP in green; Bottom, the schematic view of IF2 domains; (C) structures of N and C domain of IF3.

Other factors involved in translation initiation
Although an *in vitro* protein-synthesizing system showed that IF1, IF2 and IF3 are essential and sufficient to stimulate translation initiation, other factors have been implicated in this process to various extents as described below (147,179).

EF-P is encoded by *efp* gene in *E. coli* and is conserved in bacteria. It is homologous to eukaryotic translation initiation factor eIF-5A (9,81). *In vitro*, EF-P can enhance the extent of peptidyl transfer in reactions employing small A-site analogs (e.g., puromycin) (52). Recent X-ray crystal structure of EF-P bound to the 70S ribosome revealed that EF-P binds in a location between P site and E site and appears to stabilize the initiator tRNA in the P site for the first round of peptide bond formation (14).

YciH is homologous to eIF1/SUI1, a eukaryotic translation initiation fidelity factor involved in start codon selection (81). Although YciH is found in all known archaeal genomes and some enteric bacteria, it is not essential and has no known function. An *in vitro* study has shown that YciH binds to the similar part of small subunit as IF3 and can perform some functions of IF3 (90).

During cold stress, translation is arrested by YfiA, which is released upon the restoration of normal conditions (2,94). An early model suggested that YfiA affects elongation by blocking aminoacyl-tRNA incorporation (2). However, recent structural and biochemical studies suggest that the arrest happens at the initiation stage (172). YfiA competes with the binding of the initiator tRNA and one or more initiation factors by occupying the ribosomal P and A sites, which rationalizes its inhibitory effects on initiation.
1.3. Fidelity of translation initiation

To ensure faithful translation, the ribosome must initiate only at the correct start codon. Errors in initiation can potentially occur if: (1) the 70SIC forms at a codon other than the canonical start codon; (2) the 70SIC forms with elongator tRNA (115). The former case represents an error in start codon selection, while the latter represents an error in the tRNA selection. Below, I will summarize the current views of how the translation initiation machinery obtains its high fidelity, with the focus on start codon selection.

1.3.1. Start codon selection

The base complementarity is the primary source of the selection of the correct start codon. The anticodon of fMet-tRNA, 5’-CAU-3’ forms perfect match with canonical start codon 5’-AUG-3’. The +1 position of the start codon seems to be tolerant of substitutions, thereby allowing efficient use of GUG and UUG. The predicted energetic cost of a single base pair mismatch in solution is insufficient to account for the level of discrimination observed in start codon selection (45). However, the effects of mismatch on the stability of codon-anticodon pairing in the P site of the 30S subunit have not been well characterized. It has been shown that in the 30S subunit bound with IFs, replacing AUG with AUU greatly reduces the stability of 30S·mRNA·tRNA ternary complex (84). This effect was attributed to IF3. In the absence of IF3, no obvious differences were detected between the two complexes. This study suggested: (1) The 30S in the presence of IF1·IF2 performs poorly in selecting
correct start codon; (2) IF3 plays an important role in start codon selection, as demonstrated in genetic studies (133,164). As described above, IF2 stimulates fMet-tRNA binding and promotes formation of 30SIC (8,180), probably by stabilizing the acceptor end of fMet-tRNA (152). Since 30S·mRNA·tRNA is likely stabilized regardless of the start codon, and an increase in the rate of the forward reaction is predicted to occur at the cost of fidelity. Thus, a fidelity factor, like IF3, is essential to regulate this process.

How IF3 improves the stringency of start codon selection is not clear yet. One prevalent view is that IF3 monitors codon-anticodon pairing and only discriminates against mismatches. It has been suggested that IF3 selects the start codon AUG, GUG and UUG itself (64). However, a later study argued that IF3 selects the canonical codon-anticodon pairing in the P site by showing complementary substitution at the anticodon can suppress the mutation in the start codon (99). It was suggested that IF3 distinguishes correct from incorrect base-pairing geometry, in a similar way that enzymes selects correct substrate. However, this proposal requires that IF3 contact the codon-anticodon helix. Hitherto, evidence that IF3 directly interacts the codon-anticodon helix has been lacking. Alternatively, IF3 could indirectly use P site contacts to monitor the codon-anticodon pairing. UV-induced crosslinking study showed that structural changes (alteration of frequency of intramolecular RNA crosslinks C1402XC1501, C967XC1400 and U793XG1517) occur in the presence of IF3, implying that IF3 can restructure 30S P site (146). However, this idea is difficult to
reconcile with the structural data, which showed 16S rRNA nucleotides do not directly contact the minor groove of codon-anticodon pairing (Figure 1.3) (12, 80, 141).

A kinetic study investigating tRNA selection (see below) showed that IF3 destabilizes 30S·mRNA·fMet-tRNA complex programmed with AUG (7), arguing against the idea that IF3 only destabilizes non-canonical complexes. Then, how does IF3 increase the fidelity of start codon selection? An alternative mechanism was proposed in the same study. The intrinsic stability of 30S·AUU·fMet-tRNA is smaller than 30S·AUG·fMet-tRNA. However, when reaction of forming 70SIC is shifted in the forward direction by IF2, the dissociation rates of both complexes are too slow compared to forward rates (e.g., 50S docking), thereby making the intrinsic stability differences irrelevant for the selectivity of the reaction. IF3 kinetically tunes the reaction with little cost of efficiency. It slows down 50S docking on the one hand, and generally speeds up dissociation of 30SICs on the other, allowing the intrinsic differences of 30S·AUU·fMet-tRNA and 30S·AUG·fMet-tRNA to be utilized for discrimination. While attractive, this model has not been rigorously tested.

How IF3 destabilizes tRNA in the P site is not yet known. IF3 could rearrange the P site, thereby removing some non-specific contacts between the 30S subunits and fMet-tRNA. For example, it has been hypothesized that S9 and S13 lose contacts with fMet-tRNA in the presence of IF3 (68). It has also been suggested that IF3 opens the gate formed by G1338 and A790, facilitating dissociation of fMet-tRNA (42).

Recently, studies of fidelity of initiation have been extended to the later events of translation initiation (50S docking, GTP hydrolysis, Pi release and dissociation of
initiation factors). While the sequence of the start codon (AUG vs AUU) had little apparent effect on 30SIC formation in their hands, Grigoriadou et al. noticed that there is ~ 20-fold difference in the rate of 70S formation (mostly in a slow phase of 50S docking) (55). This difference disappears in the absence of IF3, which leads to the hypothesis that IF3 retards the conversion of a loose 70SIC into a more stable 70SIC in the presence of a non-canonical start codon. Another group demonstrated that omission of IF1 also stimulates 70SIC formation in the presence of AUU by increasing the rate of docking of the 50S subunit and release of IF3 (102). This study provided additional evidence that the start codon is monitored at later stages of translation initiation. Furthermore, these studies suggested that IF1 assists IF3 by inducing a docking-unfavorable state, which is rescued or overcome in the presence of a canonical start codon. These studies suggest that identity of the start codon impacts the initiation process before and after GTP hydrolysis, raising the possibility that a proofreading mechanism may also contribute to initiation fidelity.

1.3.2. Initiation with elongator tRNA

The efficiency of translation initiation is generally high when the start codon is sitting in an optimal distance downstream from the SD (29,170). However, the window of this spacing is broad, thus triplets other than the start codon within this window are also potential targets for initiation with elongator tRNA species. How does the translation machinery restrict access of elongator tRNA into the P site during translation initiation? Selection of fMet-tRNA over elongator tRNAs is due in part to
IF2, which specifically binds charged tRNA with a blocked NH$_2$ group (e.g. formylmethionine) (7,64,65). Another specificity determinants of fMet-tRNA is a row of 3 G:C base pairs in the ASL, which are required for efficient initiation (64,95,143). Notably, it has been proposed that IF3 is not only responsible for selecting the start codon but also fMet-tRNA. It has been suggested based on S1nuclease protection data that fMet-tRNA possesses a special conformation associated with the 3 G:C base pairs (64). IF3 selects fMet-tRNA by directly recognizing this type of conformation. This model gained support from recent X-ray study that showed fMet-tRNA has a non-canonical conformation at the ASL (10). However, other studies using similar S1 protection techniques produced contradictory results (95).

Another view regarding how these G:C pairs are recognized is that tRNAs with G:C pairs have intrinsic high affinities for the ribosome. Indeed, G:C pairs are predicted to form optimal interactions with 16S rRNA P-site nucleotides G1338 and A1339 (12). A biochemical study showed that mutations in G1338 and A1339 cause the 30S P site more readily to release the tRNA with G:C pairs in the presence of IF3 (86). This study proposed that the role of IF3 in selecting fMet-tRNA is to unmask an inherent preference of the 30S subunit for tRNA with G:C pairs, probably by removing non-specific contacts between the 30S subunit and P-site tRNA (e.g., S9 and S13 contacts). In line with this proposal, another study showed that IF3 by itself uniformly increases the dissociation rate three tRNA species containing G:C pairs, including fMet-tRNA (7). Additionally, this study suggested that the formyl group on fMet-tRNA
is the major determinant of tRNA selection; however, an increase in the rate of tRNA
dissociation from the 30S subunit by IF3 is required to realize the selection potential.
Chapter 2

Characterization of 16S rRNA Mutations that Decrease the Fidelity of Translation Initiation

2.1. Introduction

Initiation of translation involves recognition of the start codon by initiator tRNA in the P site of the small ribosomal subunit. This process requires both selection of the initiator tRNA from all other species of tRNA and selection of the start codon from other similar or identical triplet sequences in the mRNA. In bacteria, molecular determinants that specify the initiator tRNA (fMet-tRNA\textsuperscript{fMet}) and the start codon have been identified (reviewed in (57,97). One unique feature of initiator tRNA\textsuperscript{fMet} is a mismatch at the last base pair of the acceptor stem (C1×A72). This mismatch is important for formylation of Met-tRNA\textsuperscript{fMet} (88), without which specific recognition by initiation factor 2 (IF2) is lost (7). The C1×A72 mismatch also prevents Met-tRNA\textsuperscript{fMet} from binding EF-Tu and participating in the elongation phase of translation (144).

Another unique feature of tRNA\textsuperscript{fMet} is a series of three adjacent G-C base pairs in the anticodon stem. Two of these G-C base pairs are particularly important for efficient initiation (95), and these (G29-C41 and G30-C40) interact with 16S rRNA nucleotides G1338 and A1339, respectively (12,80,141). The importance of these G-C pairs may
stem from their ability to stabilize fMet-tRNA^{fMet} in the 30S P site (86,95,143).

Regarding start codon selection, an important determinant is a purine-rich element (e.g., AGGA) termed the Shine-Dalgarno sequence (SD) (148), which is positioned 6-12 nucleotides upstream from the start codon (150). The SD sequence base-pairs with the 3' end of the 16S rRNA (termed the anti-Shine Dalgarno sequence or ASD) to facilitate positioning of the start codon in the 30S P site (160). The sequence of the SD, sequence of the start codon, number of nucleotides between the SD and start codon, and the mRNA secondary structure influence the efficiency of initiation (162,170), indicating the importance of each of these determinants in the initiation process.

Translation initiation in bacteria is kinetically controlled by three initiation factors (IF1, IF2, and IF3), and recent studies have helped clarify the roles of these factors in the initiation process (6,7,8,162). The first major step in the process is assembly of the 30S pre-initiation complex. In this complex, fMet-tRNA^{fMet} is paired to the start codon of mRNA in the 30S subunit P site, IF1 occupies the 30S A site, IF2•GTP binds an adjacent position and interacts specifically with the acceptor end of fMet-tRNA^{fMet}, and IF3 binds the platform of the 30S subunit near the P and E sites (3,26,35,96,98,106). The order of assembly remains unclear, and it has been proposed that mRNA and fMet-tRNA^{fMet} associate in random order (57). During pre-initiation complex formation, the initiation factors act synergistically to increase the association rate of fMet-tRNA^{fMet} (by 400-fold), while IF1 and IF3 also increase the dissociation rate of fMet-tRNA^{fMet} (by 300-fold) (7,8). Other species of tRNA exhibit considerably slower association rates and faster dissociation rates, suggesting the basis for fMet-
tRNA^fMet selection at this stage of initiation (7). The second major step of the initiation process is 50S subunit docking. The GTP-bound form of IF2 accelerates 50S docking, while IF3 inhibits the event (6,8). It has been proposed that 50S docking can occur only after release of IF3 from the complex, and thus IF3 acts to prevent premature 50S docking when fMet-tRNA^fMet is absent (8). After 50S docking, hydrolysis of GTP facilitates release of IF2, generating a 70S initiation complex ready for the first round of elongation (6).

A number of mutations have been isolated in *E. coli* that increase translation from non-canonical start codons (i.e., codons other than AUG, UUG, and GUG) (62,133,164). These mutations mapped to *infC*, the gene encoding IF3, strongly implicating the factor in the fidelity of initiation. Reducing the concentration of wild-type IF3 in the cell also increased spurious initiation (119), suggesting that the isolated mutations may reduce either the activity or level of IF3. Over-expression of IF3 conferred the opposite phenotype: repression of translation from non-canonical start codons relative to control strains (19,133). The ability of IF3 to influence start codon selection is well exemplified by its ability to negatively regulate its own gene, *infC*. Translation of *infC* mRNA depends on initiation from the non-canonical start codon AUU (19,20).

In theory, spurious initiation could involve either an elongator tRNA paired to a cognate codon (i.e., an error in tRNA selection) or fMet-tRNA^fMet paired to a non- or near-cognate codon (i.e., an error in codon selection). *N*-terminal sequencing of polypeptide products of spurious initiation events indicated that, in *infC* and wild-type
strains, translation began with the initiator tRNA (115,116). Moreover, spurious initiation was observed specifically from codons termed Class IIA (i.e., AUA, AUC, AUU, ACG, and CUG) (133,164), which differ from cognate start codons (AUG, GUG, or UUG) at only one position. This sequence dependence would not be expected if elongator tRNAs played a significant role in spurious initiation. These observations suggest that IF3 increases fidelity in vivo primarily by preventing errors in start codon selection.

Mutations in the 16S rRNA genes were not recovered in the aforementioned genetic studies, even though IF3 makes extensive contact with the platform domain of the 30S subunit (35). There are seven copies of the 16S rRNA gene in E. coli, which raises the possibility that these mutations were missed because their effects were masked by the remaining wild-type copies. Here, using a specialized ribosome system, we show that mutations G1338A and A790G of 16S rRNA increase translation from non-canonical start codons and do so in distinct ways. Mutation G1338A stabilizes the tRNA^fMet in the 30S P site, whereas A790G decreases the affinity of IF3 for the 30S subunit. Additionally, we provide evidence that 16S rRNA mutations previously reported to decrease initiation fidelity confer this phenotype indirectly.

2.2. Materials and methods

2.2.1. Specialized ribosome strains

Translation activity of specialized ribosomes was measured in indicator strains similar to those described previously (1). Each strain contained the lacZ gene (with the
alternative SD sequence 5'-ATCCC-3') in single copy on the chromosome. To construct these strains, DNA fragments containing a consensus variant of the P_{ant} promoter (108), the alternative SD sequence 5'-ATCCC-3', and a start codon (ATG, ACG, ATC, or CTG) were generated as described (1). These fragments were digested with EcoRI and BamHI and cloned upstream of lacZ in pRS552, and the resulting fusions were transferred to λRS45 by homologous recombination \textit{in vivo} (153). The recombinant λ phage were then used to lysogenize strain JK382 [F- ara-600 Δ(gpt-lac)5 relA1 spoT1 thi-1 zdi-925::Tn10], JK378 [F- ara-600 Δ(gpt-lac)5 relA1 spoT1 thi-1 zdi-925::Tn10 infC362] and JK530 [F- ara-600 Δ(gpt-lac)5 relA1 spoT1 thi-1 zdi-925::Tn10 infC561] (164), which were obtained from the Coli Genetic Stock Center (Yale University). Each of the resulting lysogens was confirmed by PCR to contain a single prophage (127).

Plasmid pKF207 contains the 16S rRNA gene with the altered ASD 5'-'GGGGT-3' (89) under transcriptional control of the P_{BAD} promoter. The construction of pKF207 and its derivatives was described previously (1). Expression of these plasmid-borne 16S rRNA alleles in the indicator strains described above allowed translation activity of the corresponding mutant ribosomes to be measured.

\textbf{2.2.2. Activity of mutant ribosomes \textit{in vivo}}

Cells from a saturated culture were diluted 500-fold into fresh Luria Broth (LB) containing ampicillin (Amp; 100 μg/mL), kanamycin (Kan; 30 μg/mL), and L-arabinose (5 mM) and grown for 6 h at 37°C. The cells were washed once in Z buffer
(100 mM sodium phosphate [pH 7.0], 10 mM KCl, 10 mM MgSO₄), and β-galactosidase activity was measured as described (1), except that Chlorophenol Red-β-D-galactopyranoside (CPRG; Sigma) was used as the substrate. Reported units were defined by the equation: 1 unit = 1000 × (A₅₇₃) / (OD₆₀₀ × v × t), where A₅₇₃ is absorbance at 573 nm (characteristic of the product), OD₆₀₀ is optical density of the cell suspension used, v is the volume of cell suspension used (in mL), and t is time of incubation (in min) at room temperature. Differences were deemed statistically significant based on Student's t test with the Bonferroni correction.

The level of mutant versus endogenous (wild-type) 30S subunits in various ribosomal fractions was determined as described previously (1). A radiolabeled primer was annealed to 16S rRNA at a position 3' of the mutation site such that primer extension in the presence of a specific dideoxynucleotide triphosphate resulted in distinct products corresponding to mutant and wild-type templates. To detect ΔC1400, primer #1401 (5'-CCCATGGTGTGACGGGCG-3') was extended in the presence of ddGTP, dATP, dCTP, and dTTP. To detect G529U, primer #530 (5'-TTGCACCCTCCGTATTA-3') was extended in the presence of ddATP, dCTP, dGTP, and dTTP. To detect G1338A, primer #1342 (5'-GATCCACGATTACTAGC-3') was extended in the presence of ddCTP, dATP, dGTP, and dTTP. To detect A790G, primer #791 (5'-CGTGGACTACCAGGGTAT-3') was extended in the presence of ddTTP, dATP, dCTP, and dGTP.

2.2.3. Δ7 prrn strains
Mutations G1338A and A790G were introduced into plasmid p278MS2 (183), which contains (a tagged version of) the \textit{rrnB} operon, to generate plasmids pNA56 and pNA58, respectively. Plasmids p278MS2, pNA56 and pNA58 were then transformed into the \textbf{Δ7 prrn} strain SQZ10 (kindly provided by C. Squires), selecting for resistance to Amp. Resulting transformants were grown in liquid media and spread onto LB plates containing Amp and sucrose (5\%) to select against pHKrrnC-sacB (Kan\textsuperscript{R}), the resident plasmid of SQZ10. Isolates resistant to sucrose and Amp but sensitive to Kan were identified in which the p278MS2 derivative had replaced pHKrrnC-sacB. Plasmid replacement was confirmed in strains expressing mutant ribosomes by (1) purifying the plasmid and sequencing the region containing the mutation and (2) purifying 16S rRNA and confirming the presence of the mutation using primer extension (Fig. 3B-C, lanes 11-12)

\textbf{2.2.4. Purification of 30S subunits}

Each \textbf{Δ7 prrn} strain was grown to mid-logarithmic phase (OD\textsubscript{600} ≈ 0.4) at 37°C in 1L LB, the culture was chilled on ice, and all subsequent steps were performed at 4°C. The cells were collected by centrifugation, washed in 30 mL of buffer A [20 mM Tris-HCl (pH 7.5), 10.5 mM MgCl\textsubscript{2}, 100 mM NH\textsubscript{4}Cl, 0.5 mM EDTA, 6 mM β-ME], resuspended in 30 mL of the same buffer, and lysed by passage through a French press. The cell lysate was clarified by centrifugation at 15,000 rpm for 15 min. The ribosomes were pelleted through a 10 mL sucrose cushion (1.1 M) in buffer B [20 mM Tris-HCl (pH 7.5) 10.5 mM MgCl\textsubscript{2}, 500 mM NH\textsubscript{4}Cl, 0.5 mM EDTA, 6 mM β-ME] by
centrifugation at 41,000 rpm for 18 h in a Beckman Ti60 rotor. The ribosome pellet was dissolved in buffer C [50 mM Tris-HCl (pH 7.5), 1 mM MgCl$_2$, 100 mM NH$_4$Cl, 6 mM β-ME] and dialyzed against the same buffer to dissociate the ribosomes into subunits. Subunits were separated by centrifugation through 34 mL sucrose gradients (10-30% sucrose in buffer C) at 20,000 rpm for 13 h in a Beckman SW32 rotor. 30S subunits were collected, the magnesium concentration was raised to 10 mM, and the subunits were pelleted by centrifugation in a Beckman Ti60 rotor at 38,000 rpm for 17 h. Finally, the subunits were dissolved in buffer D [50 mM Tris-HCl (pH 7.5), 10 mM MgCl$_2$, 100 mM NH$_4$Cl, 6 mM β-ME], separated into small aliquots, flash frozen, and stored at -70°C.

2.2.5. P-site tRNA binding experiments

The affinity of tRNA$^{fMet}$ for the 30S subunit P site was measured using a double-filter method described previously (43). 3’-[$^{32}$P]-tRNA$^{fMet}$ (10 nM), mRNA (5’-UAAACGAGGAAACAAAUGGCUCGCACAACA-3’; 2 µM) and heat-activated 30S subunits at various concentrations were incubated at 37°C for 10 min in 20 µL TNM buffer [50 mM Tris-HCl (pH 7.5), 20 mM MgCl$_2$, 100 mM NH$_4$Cl and 6 mM β-ME). Binding reactions were diluted with 100 µL of cold TNM and immediately filtered through bi-layer of nitrocellulose (GE Corporation) and Hybond-N$^+$ (Amersham Biosciences) membranes, using a 96-well dot-blot apparatus (Schleicher and Schuell). The membranes were separated, dried, and exposed to a phosphorimager screen for quantification. The fraction of counts trapped on the nitrocellulose was taken as the
fraction of tRNA\(^{\text{fMet}}\) bound and plotted as a function of 30S subunit concentration using Kaleidagraph. Dissociation constants (\(K_D\)) were determined using the equation:

\[
[30S\cdot\text{tRNA}] = \frac{([\text{tRNA}]_{\text{input}} + [30S]_{\text{input}} + K_D - (([\text{tRNA}]_{\text{input}} + [30S]_{\text{input}} + K_D)^2 - 4[\text{tRNA}]_{\text{input}}[30S]_{\text{input}})^{1/2}}{2}.
\]

### 2.2.6. IF3 binding experiments

IF3 containing a C-terminal hexahistidine tag was purified as described previously (35) and labeled with an amino-reactive fluorophore (Alexa Fluor® 488) as recommended by the supplier (Invitrogen). In a 0.1 mL reaction, IF3 (1 mg) was stirred with Alexa Fluor® 488 carboxylic acid, 2,3,5,6-tetrafluorophenyl ester (50 \(\mu\)g; Invitrogen) in 0.1 M sodium bicarbonate buffer (pH 9.0) for 1 h at room temperature. The reaction was stopped by adding 10 \(\mu\)L of freshly prepared 1.5 M hydroxylamine (pH 8.5). The labeled protein (IF3-AF) was separated from unreacted labeling reagent by extensive dialysis in PBS (137 mM NaCl, 2.7 mM KCl, 10 mM Na\(_2\)HPO\(_4\) and 2 mM KH\(_2\)PO\(_4\)) buffer (pH 7.4), separated into small aliquots, flash frozen, and stored at -70°C. The degree of labeling, determined by measuring the absorbance of the protein at 280 nm and the absorbance of the dye at 495 nm, was estimated to be 0.6 dyes per IF3.

IF-AF (1 \(\mu\)M) was incubated with control or mutant 30S subunits (1 \(\mu\)M) in 10 mM Tris-HCl (pH 7.5), 60 mM NH\(_4\)Cl, 10 mM Mg(OAc)\(_2\), 0.03% Nikkol and 6 mM \(\beta\)-mercaptoethanol at room temperature for 10 min. Samples were then loaded onto 11 mL sucrose gradients (10-30% sucrose in the same buffer) and centrifuged at 26,000 rpm for 16 h in a Beckmann SW41 rotor. Gradients were fractionated while \(A_{260}\) was
monitored using an in-line UV detector (ISCO/Brandel system). Fractions (0.5 mL) were collected and the fluorescence intensity (excitation and emission at 495 nm 520 nm, respectively) of each was measured.

Steady-state fluorescence measurements were made using a Fluorolog®-3 spectrofluorometer (Jobin Yvon, Inc.) and Type 16F Sub-Micro cuvettes (160 µL volume; Starna Cells, Inc.). The excitation and emission wavelengths were 495 nm and 520 nm, respectively, and the slit widths were set at 5 nm. To estimate equilibrium dissociation constants, fluorescence anisotropy was employed as described (4). IF3-AF (10 nM) was incubated with 30S subunits (at various concentrations) in 10 mM Tris-HCl (pH 7.5), 60 mM NH₄Cl, 10 mM Mg(OAc)₂ and 6 mM β-mercaptoethanol at 25°C for 10 minutes prior to measuring anisotropy in a thermostated chamber. Anisotropy values were plotted as a function of 30S subunit concentration, and dissociation constants (Kᵰ) were obtained from fitting the data to the equation: 

\[ A = A_{\text{min}} + (A_{\text{max}} - A_{\text{min}}) \left( \frac{([\text{IF3-AF}]_{\text{input}} + [30S]_{\text{input}} + K_D) - \left( ([\text{IF3-AF}]_{\text{input}} + [30S]_{\text{input}} + K_D)^2 - (4[\text{IF3-AF}]_{\text{input}} [30S]_{\text{input}})^{1/2}\right)}{2[\text{IF3-AF}]_{\text{input}}} \right), \]

where A is the measured anisotropy, Aₘᵢₙₐₜ is the minimum anisotropy, and Aₘₐₓ is the maximum anisotropy.

2.3. Results and discussions

2.3.1. G1338A and A790G increase translation from non-canonical start codons.

To investigate the role of 16S rRNA in start codon selection, we screened a collection of P-site mutations for those that increase initiation from non-canonical start codons. A specialized ribosome system was used to allow translation activity of the
mutant ribosomes to be quantified in vivo without secondary effects on cell growth (1). Prior to the screen, sensitivity was optimized in this system by strengthening the promoter of the chromosomally-encoded lacZ reporter gene and by assaying β-galactosidase activity with the substrate CPRG (see Materials and Methods). We found that two mutations, G1338A and A790G, enhanced translation from ACG, AUC, and CUG, but not from AUG (Table 2.1). Both of these mutations increased spurious initiation modestly (by 30-110%) but significantly (P < 0.05). This phenotype was not attributable to a number of other P-site mutations (e.g., A790C, A790U, G926A, G926C, G926U, m2G966A, m2G966C, m2G966U, G1338C, G1338U, A1339C, A1339G, A1339U, C1400A, C1400G, and C1400U).

Next, we measured the effects of G1338A and A790G in the presence of infC362, a mutant allele of the IF3 gene. In strains expressing the control 16S rRNA, infC362 increased translation specifically from non-canonical start codons by 5- to 9-fold (Table 2.1), consistent with previous studies (164). These data indicate that the alternative SD-ASD helix has no appreciable effect on the ability of IF3 to increase the stringency of start codon selection. When 16S rRNA containing G1338A was expressed in the infC362 background, a further increase of 60-80% was observed. In each case (ACG, AUC, and CUG), the increase attributed to G1338A was significant (P < 0.05). Thus, the effects of G1338A and infC362 are additive, suggesting that they act independently to increase spurious initiation. By contrast, A790G did not confer an increase in spurious initiation in the presence of infC362. Instead, A790G decreased translation from ACG, AUC, and CUG by ~ 2-fold. Similar results were obtained when
these experiments were repeated with \textit{infC561} (data not shown), arguing against allele-specific effects. These data provide evidence that G1338A and A790G decrease the fidelity of initiation in distinct ways.

Interestingly, in the \textit{infC362} background, G1338A increased translation from AUG significantly (P < 0.05; Table 2.1). It has been shown that IF3 acts in concert with the other initiation factors to stimulate initiation complex formation at AUG start codons. Additionally, IF3 negatively regulates the process to enhance accuracy (7,8). It is possible that in the \textit{infC362} background, the efficiency of initiation complex formation at the canonical AUG start codon is compromised and G1338A compensates to restore high-level translation.
Table 2.1. Effects of 16S rRNA mutations G1338A and A790G on translation of mRNA containing non-canonical start codons.

<table>
<thead>
<tr>
<th>Start codon</th>
<th>16S allele(^1)</th>
<th>Wild-type (infC(^+)) background</th>
<th>infC362 background</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(\beta)-gal Activity(^2)</td>
<td>Ratio (16S mutant/control)</td>
<td>(\beta)-gal Activity(^2)</td>
</tr>
<tr>
<td>AUG</td>
<td>Vector only</td>
<td>11 ± 0.7</td>
<td>9.1 ± 2</td>
</tr>
<tr>
<td></td>
<td>control</td>
<td>1500 ± 90</td>
<td>1300 ± 40</td>
</tr>
<tr>
<td></td>
<td>G1338A</td>
<td>1300 ± 70</td>
<td>0.9</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1600 ± 80</td>
<td>1.2</td>
</tr>
<tr>
<td></td>
<td>A790G</td>
<td>390 ± 60</td>
<td>0.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>190 ± 10</td>
<td>0.2</td>
</tr>
<tr>
<td>ACG</td>
<td>vector only</td>
<td>&lt; 0.12</td>
<td>&lt; 0.12</td>
</tr>
<tr>
<td></td>
<td>control</td>
<td>2.6 ± 0.2</td>
<td>24 ± 1</td>
</tr>
<tr>
<td></td>
<td>G1338A</td>
<td>3.8 ± 0.5</td>
<td>1.5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>44 ± 3</td>
<td>1.8</td>
</tr>
<tr>
<td></td>
<td>A790G</td>
<td>5.5 ± 0.8</td>
<td>2.1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>14 ± 1</td>
<td>0.6</td>
</tr>
<tr>
<td>AUC</td>
<td>vector only</td>
<td>&lt; 0.12</td>
<td>&lt; 0.12</td>
</tr>
<tr>
<td></td>
<td>control</td>
<td>5.8 ± 0.3</td>
<td>30 ± 2</td>
</tr>
<tr>
<td></td>
<td>G1338A</td>
<td>7.7 ± 0.4</td>
<td>1.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>50 ± 3</td>
<td>1.7</td>
</tr>
<tr>
<td></td>
<td>A790G</td>
<td>8.4 ± 0.7</td>
<td>1.4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>12 ± 1</td>
<td>0.4</td>
</tr>
<tr>
<td>CUG</td>
<td>vector only</td>
<td>&lt; 0.12</td>
<td>&lt; 0.12</td>
</tr>
<tr>
<td></td>
<td>control</td>
<td>5.3 ± 0.2</td>
<td>30 ± 3</td>
</tr>
<tr>
<td></td>
<td>G1338A</td>
<td>6.7 ± 0.3</td>
<td>1.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>47 ± 4</td>
<td>1.6</td>
</tr>
<tr>
<td></td>
<td>A790G</td>
<td>9.5 ± 0.6</td>
<td>1.8</td>
</tr>
<tr>
<td></td>
<td></td>
<td>14 ± 2</td>
<td>0.5</td>
</tr>
</tbody>
</table>

\(^1\)Each 16S allele contains the specialized ASD sequence to allow specific translation of the lacZ reporter gene.

\(^2\)Assay based on cleavage of CPRG, as described in Materials and Methods. Reported values represent the mean ± SEM.
2.3.2. G1338A increases the affinity of tRNA\textsuperscript{fMet} for the 30S subunit P site.

Previous studies showed that G1338A can suppress phenotypes conferred by other P-site mutations, suggesting that G1338A might stabilize tRNA in the 30S P site (1). To investigate this possibility, we first needed to purify mutant 30S subunits. Plasmids encoding \textit{rrnB} with either G1338A or A790G in the 16S gene (containing a wild-type ASD) were moved into an \textit{E. coli} strain lacking all chromosomal \textit{rrn} operons (\textDelta7 prrn), replacing the resident plasmid containing \textit{rrnC} (see Materials and Methods). Although G1338 and A790 are universally conserved, ribosomes carrying either G1338A or A790G supported robust cell growth (Table 2.2). These mutations decreased the growth rate of the \textDelta7 prrn strain by only 6\% and 18\%, respectively.

From these strains, mutant and control 30S subunits were purified and their affinity for tRNA\textsuperscript{fMet} was compared (Figure 2.1; Table 2.2). Subunits containing G1338A exhibited higher affinity for tRNA\textsuperscript{fMet} than control subunits. Because this difference in affinity was modest (~ 2-fold), mutant and control 30S subunits were re-purified two more times and the binding experiment was repeated. In each independent experiment, subunits with G1338A bound tRNA\textsuperscript{fMet} with ~ 2-fold higher affinity than control subunits. By contrast, A790G decreased the affinity of tRNA\textsuperscript{fMet} by ~ 4-fold (Fig. 2.1; Table 2.2).

Structural studies have shown that A1339 and G1338 dock into the minor groove of the anticodon stem of P-site tRNA(12,80,141). A1339 and G1338 are positioned to form Type I and Type II interactions with tRNA base pairs 30-40 and 29-
41, respectively. Our data suggest that G1338 forms a suboptimal interaction with tRNA^{fMet} in the 30S P site, since tRNA^{fMet} binding is enhanced by mutation G1338A (Fig. 2.1). Study of an analogous Type II interaction showed that replacement of the docking A for G caused a modest decrease in stability (40), in line with our results. We propose that the ability of G1338A to stabilize fMet-tRNA^{fMet} in the 30S P site partially compensates for mismatches in the codon- anticodon helix and thereby allows increased levels of spurious initiation.

Although our data show that G1338 helps increase the stringency of start codon selection, the contribution of G1338 to fidelity appears minor compared that of IF3. Mutations in IF3 generally derepress spurious initiation by 5- to 40-fold (Table 2.1; (115,133,164), while the effects of G1338A are 2-fold or less (Table 2.1). Furthermore, the ability of IF3 to discriminate the start codon does not depend on the Type II "G-
minor" interaction, because the degree of repression by IF3 (infC+ versus infC362) is comparable for ribosomes harboring either G or A at position 1338 (Table 2.1).
Table 2.2. Summary of effects conferred by 16S rRNA mutations G1338A and A790G.

<table>
<thead>
<tr>
<th>16S rRNA allele</th>
<th>Growth rate in Δ7 prrn strain</th>
<th>Equilibrium binding constant ($K_D$; nM) of 30S subunit for indicated ligand</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(doublings /hour)</td>
<td>tRNA$^{f_{Met}}$</td>
</tr>
<tr>
<td>Control</td>
<td>1.7 ± 0.09</td>
<td>310 ± 20</td>
</tr>
<tr>
<td>G1338A</td>
<td>1.6 ± 0.02</td>
<td>190 ± 10</td>
</tr>
<tr>
<td>A790G</td>
<td>1.4 ± 0.04</td>
<td>1100 ± 200</td>
</tr>
</tbody>
</table>

Reported values represent the mean ± SEM.

1. *E. coli* strain lacking all chromosomal *rrn* operons and thus dependent on a plasmid-borne *rrn* operon.
2. Measured in the presence of a model mRNA with an AUG start codon.
3. IF3 labeled with Alexa Fluor 488 (Invitrogen).

Figure 2.1. Mutation G1338A increases the affinity of tRNA$^{f_{Met}}$ for the 30S subunit P site.

Wild-type (○) or mutant 30S subunits harboring G1338A (△) or A790G (□) were incubated at various concentrations with a model mRNA (containing an AUG codon; 2 µM) and 3'-[32P]-tRNA$^{f_{Met}}$. Then, the fraction of 3'-[32P]-tRNA$^{f_{Met}}$ bound was determined by filtration through a bi-layer of nitrocellulose and nylon membranes (see Materials and Methods). Equilibrium binding constants ($K_D$) listed in Table 2 come from at least three independent binding experiments like the one shown here.

2.3.3. A790G decreases the affinity of IF3 for the 30S subunit.
The genetic data suggested that A790G and G1338A differentially perturb the initiation process (Table 2.1). Although both contribute to the 30S P site, A790 and G1338 lie in different domains of the subunit and contact opposite sides of the P-site tRNA (80,141). IF3 interacts predominantly with the platform domain of 16S rRNA, protecting A790 and nearby 16S rRNA nucleotides from chemical probes (35,106). A mutation in this region (G791A) has been shown to reduce the affinity of IF3 for the 30S subunit by 9-fold (166). With these observations in mind, we considered the possibility that decreased fidelity conferred by A790G was due to a defect in IF3 binding.

To study the effects of these 16S rRNA mutations on IF3 binding, we labeled IF3 using an amino-reactive fluorophore (Alexa Fluor® 488; see Materials and Methods). The degree of labeling was estimated to be 0.6 dyes per IF3. In the first binding experiment, fluorescently-labeled IF3 (IF3-AF; 1 μM) was incubated with each preparation of 30S subunits (1 μM), reactions were subjected to sucrose gradient sedimentation, and the relative amount IF3-AF migrating at 30S was assessed (Fig. 2.2A). With wild-type and G1338A subunits, an obvious peak of fluorescence was detected in fractions corresponding to the 30S subunit peak, deduced by absorbance at 260 nm. By contrast, IF3-AF did not co-migrate with 30S subunits containing A790G. Instead, increased fluorescence signal was observed at the top of the gradient, where free IF3-AF was expected to remain. Consistent with this interpretation, control reactions containing 50S subunits or lacking subunits gave identical distributions, with IF3-AF detected only at the top of the gradient.
Next, we estimated the affinity of IF3-AF for each of the 30S subunits using fluorescence anisotropy (Fig. 2.2 B; Table 2.2). In each case, an increase in anisotropy attributable to binary complex formation was observed when 30S subunits were added to IF3-AF. Mutation G1338A did not affect the affinity of IF3-AF for 30S subunits, while A790G decreased the affinity for IF3-AF by 10-fold (Fig. 2.2B; Table 2.2). These data corroborate results from the sucrose gradient sedimentation analysis and suggest that decreased fidelity conferred by A790G is due to decreased affinity of IF3 for the 30S subunit.

How IF3 increases the accuracy of start codon selection remains unclear, but recent studies provide clues about the mechanism (7,8,162). IF3 negatively regulates formation of the 30S pre-initiation complex by destabilizing tRNA (including fMet-tRNA^fMet) in the 30S P site. IF3 also negatively regulates the subsequent step, 50S subunit docking. Based on these observations, reduced IF3 activity in the cell is predicted to both (1) stabilize fMet-tRNA^fMet in the pre-initiation complex and (2) stimulate 50S subunit docking. These effects may allow initiation despite mismatches in the codon-anticodon helix, thus decreasing the stringency of start codon selection. Mutation A790G, which inhibits IF3 binding to the 30S subunit (Fig. 2.2), may stimulate spurious initiation in an analogous way. Even though A790G decreases the affinity of tRNA^fMet by 4-fold in the absence of factors (Fig. 2.1), this mutation may effectively stabilize fMet-tRNA^fMet in the pre-initiation complex because it inhibits IF3 interaction. Alternatively or additionally, reduced IF3 binding conferred by A790G may stimulate 50S docking, thereby increasing spurious initiation.
Figure 2.2 Mutation A790G decreases the affinity of IF3 for the 30S subunit.

(A) Fluorescently-labeled IF3 (IF3-AF) was incubated with mutant or wild-type 30S subunits, with 50S subunits, or in absence of subunits (as indicated), and the reactions were subjected to sucrose gradient sedimentation analysis. To determine the distribution of IF3-AF in each gradient, 0.5 mL fractions were collected and the fluorescence intensity (excitation 495 nm; emission 520 nm) of each fraction was quantified. To determine the distribution of ribosomal subunits in each gradient, A_{260} was monitored during the fractionation process using an in-line UV detector. (B) Fluorescence anisotropy was used to estimate the affinity of IF3-AF for 30S subunits without (○) or with mutation G1338A (△) or A790G (□). Subunits at various concentrations were incubated with 10 nM IF3-AF and fluorescence anisotropy was measured. Equilibrium binding constants (K_D) listed in Table 2 come from at least three independent binding experiments like those shown here.
2.3.4. Mutations ΔC1400 and G529U do not confer fidelity phenotypes in the specialized ribosome system.

Several mutations in the 16S rRNA have been reported to stimulate initiation from non-canonical start codons (116). In those strains, the reporter lacZ gene contained a natural SD sequence and was therefore recognized by both the endogenous (wild-type) ribosomes and the mutant ribosomes derived from the plasmid-encoded rrnB alleles. Among the mutations associated with decreased initiation fidelity were ΔC1400 and G529U. In our specialized ribosome system, base substitutions at position 1400 did not cause spurious initiation (data not shown), which prompted us to compare the effects of ΔC1400 and G529U in the two genetic systems (Table 2.3). In the specialized ribosome system, these mutations decreased translation from AUG by 60 to 70-fold, to near background levels, suggesting that ribosomes harboring ΔC1400 or G529U are largely inactive. These mutations also decreased translation from non-canonical start codons AUC and CUG by at least 40-fold, to the detection limit of the assay. In the infC362 background, translation from AUC and CUG by ribosomes harboring ΔC1400 or G529U was detected, but reduced by >100-fold and ~60-fold, respectively, compared to the control strain.

We considered the possibility that loss of activity conferred by ΔC1400 and G529U in our strains was due to reduced levels of mutant 30S subunits. To test this, we grew our strains in the presence of arabinose to induce expression of the mutant 16S rRNA, fractionated the corresponding lysates by sedimentation through sucrose gradients, and used primer extension to determine the relative amount of mutant 16S
rRNA in the 30S, 70S, and polysome fractions (Fig. 2.3). This analysis was performed in strains where the start codon of the lacZ mRNA was either AUC or AUG, and similar results were obtained in both backgrounds. For ΔC1400, mutant 16S rRNA was present in the 30S fraction (~20% of total) and was also detected in the 70S and polysome fractions, albeit at substantially lower levels (1-3%; Fig. 3A). These data are inconsistent with the idea that reduced activity conferred by ΔC1400 results from decreased levels of 30S subunits. There was no evidence for partially assembled subunits from the A₂₆₀ trace during fractionation (data not shown), further suggesting that ΔC1400 does not alter biogenesis of the 30S subunit. For G1338A and A790G, mutant 16S rRNA was detected in all ribosomal fractions and was more evenly distributed among the fractions (Fig. 3B-C), consistent with the higher activity of these mutant ribosomes (Table 1 and 3). We also tried to analyze the distribution of 16S rRNA containing G529U in various ribosomal fractions (data not shown). However, primer extension in this region resulted in a ladder of bands, regardless of the template, and thus the experiment could not be easily interpreted.

These findings raise the possibility that increased spurious initiation observed in the conventional system (at least in the case of ΔC1400) is conferred indirectly by the mutant 30S subunits. For example, the mutant subunits might bind and sequester IF3, causing the endogenous wild-type ribosomes to initiate with less stringency. It will be of interest to further investigate whether mutant subunits can indirectly alter translation in vivo, as the results may hold relevance for numerous previous studies that employed the conventional system to characterize rRNA mutations.
Table 2.3. Spurious initiation observed upon expression of 16S rRNA containing ΔC1400 and G529U depends on recognition of the reporter mRNA by endogeneous (wild type) ribosomes.

<table>
<thead>
<tr>
<th>Start Codon</th>
<th>16S allele</th>
<th>Specialized genetic system&lt;sup&gt;1&lt;/sup&gt;</th>
<th>Conventional genetic system&lt;sup&gt;2&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>infC&lt;sup&gt;+&lt;/sup&gt; background</td>
<td>infC&lt;sup&gt;+&lt;/sup&gt; background</td>
</tr>
<tr>
<td></td>
<td></td>
<td>infC&lt;sup&gt;362&lt;/sup&gt; background</td>
<td>infC&lt;sup&gt;+&lt;/sup&gt; background</td>
</tr>
<tr>
<td>AUG</td>
<td>vector only</td>
<td>11 ± 0.7</td>
<td>9.1 ± 2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>control</td>
<td>1500 ± 90</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1300 ± 40</td>
</tr>
<tr>
<td></td>
<td></td>
<td>ΔC1400</td>
<td>21 ± 0.4</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>18 ± 0.4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>G529U</td>
<td>24 ± 0.8</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>20 ± 0.6</td>
</tr>
<tr>
<td></td>
<td>control</td>
<td></td>
<td>7300 ± 200</td>
</tr>
<tr>
<td></td>
<td>ΔC1400</td>
<td></td>
<td>8700 ± 500</td>
</tr>
<tr>
<td></td>
<td>G529U</td>
<td></td>
<td>6500 ± 600</td>
</tr>
<tr>
<td></td>
<td>control</td>
<td></td>
<td>1500 ± 90</td>
</tr>
<tr>
<td></td>
<td>ΔC1400</td>
<td></td>
<td>1300 ± 40</td>
</tr>
<tr>
<td></td>
<td>G529U</td>
<td></td>
<td>7300 ± 200</td>
</tr>
</tbody>
</table>

<sup>1</sup>This work. In these strains, the plasmid-encoded 16S rRNA contains an alternative (specialized) ASD sequence to allow specific translation of the lacZ reporter mRNA, which carries the complementary SD element. Values (mean ± SEM) are β-galactosidase units determined from the cleavage of CPRG, as described in Materials and Methods.

<sup>2</sup>From O'Connor <i>et al.</i>, 1997. In these strains, the plasmid-encoded 16S rRNA contains a natural ASD, thus the corresponding ribosomes and the endogenous wild-type ribosomes both recognize the lacZ reporter mRNA. Values (mean ± SEM) are β-galactosidase Miller units.

<sup>3</sup>Not reported.
Figure 2.3. Detection of mutant 16S rRNA in ribosomal fractions separated by sucrose gradient sedimentation.

In each panel, primer extension products corresponding to wild-type and mutant 16S rRNA are indicated (arrowheads). Control reactions include those lacking template (-RNA) and those in which 16S rRNA from the corresponding control strain (i.e., "wild-type" specialized) was used as template [30S(control); lanes 2 and 7]. Ribosomal fractions (as indicated) from strains expressing 16S rRNA containing ΔC1400 (A), G1338A (B), and A790G (C) were analyzed. In these strains, the start codon of the specialized reporter mRNA was either AUC (lanes 2-5) or AUG (lanes 7-10). The percentage of mutant 16S rRNA (% mut) in each fraction is indicated below the lanes. Panels B and C include additional reactions in which 16S rRNA purified from each Δ7 prrn strain was used as template (as indicated).
2.4. Conclusions

In this study, we show that mutations G1338A and A790G of 16S rRNA increase translation from non-canonical start codons in vivo. This phenotype is conferred directly, because the endogenous wild-type ribosomes fail to recognize the reporter mRNA in our strains. Mutation G1338A increases the affinity of the initiator tRNA for the 30S P site, suggesting that G1338 normally forms a suboptimal Type II interaction with fMet-tRNA^{fMet}. We propose that an enhanced Type II interaction conferred by G1338A partially compensates for mismatches in the codon-anticodon helix and thereby increases spurious initiation. Mutation A790G, on the other hand, decreases the affinity of IF3 for the 30S subunit. Reduced IF3 binding may stabilize fMet-tRNA^{fMet} in the pre-initiation complex and/or stimulate 50S subunit docking, thereby increasing spurious initiation. These mutations should prove useful in further elucidating the kinetic pathway of initiation complex formation and the molecular basis for fidelity during the process.
Chapter 3

Control of Translation Initiation Involves a Factor-induced Rearrangement of helix 44 of 16S Ribosomal RNA.

3.1. Introduction

Translation initiation involves recognition of the start codon by the initiator tRNA in the ribosomal P site. In bacteria, this process is kinetically controlled by three initiation factors (IF1, IF2, and IF3) and occurs in two major phases (15,57,58). The first phase is assembly of the 30S initiation complex (30SIC). During this phase, the initiator tRNA (fMet-tRNA^{fMet}) binds the 30S P site and its anticodon pairs with the start codon of mRNA. The second phase includes the docking of the 50S subunit onto the 30SIC, IF2-dependent GTP hydrolysis, and dissociation of the initiation factors. This generates a 70S initiation complex (70SIC) ready for the first round of elongation.

The efficiency of translation initiation is determined by cis-acting elements located in the translation initiation region (TIR) of the mRNA. One important element is the start codon. In bacteria, AUG is the most common start codon, but GUG and UUG are also used quite often (41). Another important element is a purine rich sequence (e.g., AGGA) termed the Shine-Dalgarno (SD) (148), which lies 6-12 nucleotides upstream from the start codon (150). The SD pairs with the 3' end of 16S
rRNA, termed anti-Shine-Dalgarno sequence (ASD), and this mRNA-rRNA interaction facilitates positioning of the start codon in the 30S P site (75,79,160). The sequence of the SD, sequence of the start codon, number of nucleotides between the SD and start codon, and the secondary structure of the TIR influence the efficiency of initiation (37,131,162,170), indicating the importance of each of these determinants in the initiation process.

Initiation at the correct start codon is critical for establishing the translational reading frame. A number of mutations have been isolated in *E. coli* that increase spurious initiation (i.e., translation from non-canonical start codons such as ACG, AUU, AUC, and CUG). These mutations were mapped to *infC*, the gene encoding IF3, strongly implicating the factor in start codon selection (62,133,164). Reducing the concentration of wild type IF3 in the cell also increases spurious initiation, while overexpression of IF3 confers the opposite phenotype (119,133). Indeed, the ability of IF3 to repress translation from non-canonical start codons explains how the factor auto-regulates translation of its own gene, which starts with AUU (18,20,133).

IF3 appears to act at both stages of initiation to ensure accuracy. IF3 destabilizes fMet-tRNAfMet during 30SIC formation, which presumably makes its binding more dependent on cognate codon-anticodon interaction (7,8,84). IF3 also inhibits formation of a stable 70SIC when the mRNA contains a non-canonical start codon such as AUU (54,55,102).

Although IF1 is an essential protein in the cell (33), its role in translation initiation remains elusive. IF1 is a small (~ 8 kDa) β-barrel protein that binds the 30S
A site and enhances the activities of the other two factors (7,26,106,126,145,161), which may stem in part from cooperative binding (106,177,188). Recent biochemical studies suggest that IF1 negatively regulates 70SIC formation to ensure high fidelity (102). However, in vivo evidence that IF1 increases the accuracy of initiation is lacking.

Ribosomal RNA undoubtedly plays a role in initiation, but precisely how 16S rRNA participates in the process remains unclear. Here, we randomly mutagenized 16S rRNA and isolated a number of mutations that increase translation from the non-canonical start codon AUC. These mutations map to distinct regions including the neck, the 790 loop, and helix 44 (h44). The mutations in h44 map near the non-canonical A1413°G1487 pair in a region distorted by either IF1 or IF3. Site-directed mutagenesis in this region supports the idea that this factor-induced rearrangement of h44 helps control translation initiation. We also report that many P-site mutations increase the stringency of start codon selection. These data provide evidence that the affinity of fMet-tRNAfMet for the 30S P site is tuned to balance accuracy and efficiency during initiation.

3.2. Materials and Methods

3.2.1. Bacterial strains

Strain KLF2511 is CSH142 [F- ara Δ(gpt-lac)5] containing λ(Pant-SD*-AUC-lacZ) and was constructed as described previously (129). Strain KLF2672 is KLF2511 containing Δ(recA-srl)306 srl301::Tn10. Other strains specialized ribosome strains
were described previously (129). All *E. coli* Δ7 prn strains were made using strain SQZ10 as described previously (129).

3.2.2. Genetic screen

Mutations were generated in the 16S rRNA gene by propagation of pKF207 in the mutator strain XL1-Red (Stratagene). Twelve transformants were independently grown to saturation iteratively for 7 days. Mutagenized plasmids were then purified from each culture and used to transform the reporter strain KLF2672. The transformed cells were plated on LB plates containing ampicillin (100 µg/mL), kanamycin (30 µg/mL), L-arabinose (5 mM) and X-gal (20 µg/mL), and the resulting colonies were screened for those that exhibited a bluer color than their neighbors. Plasmids were purified from the positive isolates, and the 16S rRNA gene was sequenced.

3.2.3. β-galactosidase assay

Specific β-galactosidase activity was measured as described (129), except that cells were grown overnight on plates rather than in liquid culture.

3.2.4. Purification of 35S-Labeled IF1

IF1 was overexpressed in strain BL21/DE3 (pLysS) from a pET24b-based construct (kindly provided by A. Dallas and H. Noller). The overproducer cells were grown at 37°C in 250 mL M9 medium supplemented with casamino acids (0.05%), thiamine (10 µg/mL), proline, tryptophan, leucine (80 µg/mL each), and kanamycin (30
µg/mL). At an OD₆₀₀ of 0.5, IPTG (1 mM) and [³⁵S] L-methionine (2 µCi/mL, MP Biomedical) were added to the culture. The incubation was continued for 2 h, after which the cells were collected by centrifugation at 2800×g for 15 min, washed with buffer A [20 mM Tris-HCl (pH 7.6), 20 mM NaCl, 10 mM MgCl₂, 6 mM β-ME, 0.5 mM EDTA]. The cells were lysed by freezing and thawing them twice, and the lysate was clarified by centrifugation at 1600×g for 10 min. The supernatant was passed through a 0.22 mm filter and loaded onto a Resource S column (GE Corporation). The bound protein was washed with 10 column volumes of buffer A, a NaCl gradient was applied, and IF1 was collected as a single peak (at ~125 mM NaCl). The IF1 preparation was dialyzed against buffer B [20 mM Tris-HCl (pH 7.5), 100 mM NH₄Cl, 10 mM Mg(OAc)₂, 1 mM DTT, and 10% glycerol] and stored at −70 °C. The purity of IF1 was >95%, as judged by SDS-PAGE.

3.2.5. Binding assays

Control and mutant 30S subunits were purified from D7 prrn strains of E. coli as described previously (129). To determine the affinity of IF1 for the 30S subunit in the absence of IF3, [³⁵S]-IF1 (250 nM) was incubated with activated 30S subunits (various concentration) at room temperature for 20 min in buffer C [10 mM Tris-HCl (pH 7.5), 100 mM NH₄Cl, 5 mM Mg(OAc)₂, 6 mM β-ME and 0.03% Nikkol]. Binding reactions were loaded onto Microcon YM100 filter units (Millipore) and spun at 1600×g for 5 min to separate the bound and free [³⁵S]-IF1. The flow through was collected and the signal was measured by liquid scintillation assay. A no subunit
control was included to estimate the total free $^{35}$S-IF1. The concentration of bound $^{35}$S-IF1, averaged from at least 3 independent experiments, was plotted versus the concentration of 30S subunits. To determine $K_D$, the data were fit to the equation:

$$[30S\cdot IF1]=(30S\text{input}+IF1\text{input}+K_D-((30S\text{input}+IF1\text{input}+K_D)^2-(4\times IF1\text{input}\times 30S\text{input}^{1/2}))/2),$$

using the program KaleidaGraph. The standard error reported in Table 2 derived from the curve-fitting. The same procedure was used to measure $^{35}$S-IF1 binding in the presence of IF3 (350 nM), except that the concentration of $^{35}$S-IF1 was 20 nM.

Assays used to measure binding of 3’-$^{32}$P-tRNA$^{\text{O fashion}}$ and IF3-AF to the 30S subunit were described previously (129).

3.3. Results

3.3.1. Identification of 16S rRNA mutations that increase translation from AUC

In previous work, a specialized ribosome system was used to identify two 16S rRNA mutations, G1338A and A790G, which increase translation from non-canonical start codons in vivo (129). Using the same genetic system, we sought to identify additional mutations with the same phenotype. Plasmid pKF207, encoding specialized 16S rRNA (ASD*: 5’-GGGGU-3’), was mutagenized by propagation in the mutator strain XL1-Red (Stratagene) and used to transform the indicator strain KLF2672. Strain KLF2672 expresses lacZ mRNA with the complementary SD sequence (SD*: 5’-AUCCC-3’) and the start codon AUC. Transformants were screened on X-gal plates for those expressing increased levels of $\beta$-galactosidase. Nineteen independent isolates
carrying one or more mutations in 16S rRNA were identified (Table 3.1). With one exception (U13G), these mutations were transitions, a bias that appears to be due largely to the mutagenesis with XL1-Red (data not shown). Mutation G1539A mapped to the ASD* and is predicted to change the only wobble base pair in the SD*-ASD* helix to a Watson-Crick base pair. This mutation uniformly increased expression from SD*-AUC-lacZ and SD*-AUG-lacZ by ~ 3-fold (data not shown). The other 16S rRNA mutations cluster to distinct regions that overlap with class III protection sites (Figure 3.1). Class III sites were previously implicated in ligand-induced conformational changes based on chemical protection experiments (104). Two mutations, U789C and A790G, map to the IF3-binding site. Four mutations cluster in h44, just "down" from the 30S A site, and surround A1413°G1487 (Figure 3.1B and C), a non-canonical pair conserved in bacteria and perturbed by either IF1 or IF3 (22,26,34,106). Another group of mutations cluster near the neck of the 30S subunit (Figure 3.1D).

To determine which mutations recovered were responsible for the phenotype, each mutation was introduced into pKF207 de novo by site-directed mutagenesis. The resulting plasmids were transformed into an indicator strain containing SD*-AUC-lacZ. Mutations U13G, U789C, A790G, G886A, U1083C, A1092G, G1094A, C1389U, A1410G, C1411U, U1414C, G1486A were each confirmed to increase expression from SD*-AUC-lacZ. To quantify the phenotypes, the plasmids were also moved into strains containing SD*-ACG-lacZ or SD*-AUG-lacZ and β-galactosidase activity of all resulting strains was compared (Table 3.2). Translation from AUC and ACG was
increased specifically and to various extents depending on the mutation. The level of β-galactosidase from SD*-AUC-lacZ or SD*-ACG-lacZ relative to SD*-AUG-lacZ was used to approximate the frequency of spurious initiation (Figure 3.2). Mutations U789C and A790G appeared most defective in initiation fidelity (5- to 6-fold decrease), similar to effects conferred by various infC alleles (93,133,164). The other mutations had modest but significant effects. We also determined the effects of the 16S rRNA mutations on initiation fidelity in an infC362 background. This mutation in the IF3 gene increased spurious initiation by ~ 5-fold. The effects of infC362 and the 16S rRNA mutations were completely additive in some cases (U1083C, A1092G, G1094A, C1389U, C1411U, and U1414C) and partially additive in the others (Figure 3.2).
Table 3.1. Mutations in 16S rRNA recovered in a screen for increased translation from AUC.

<table>
<thead>
<tr>
<th>16S mutation(s)</th>
<th>Independent isolates(^1)</th>
<th>Location</th>
</tr>
</thead>
<tbody>
<tr>
<td>U13G</td>
<td>1</td>
<td>central pseudoknot</td>
</tr>
<tr>
<td>A790G</td>
<td>1</td>
<td>790 loop</td>
</tr>
<tr>
<td>G886A</td>
<td>1</td>
<td>h27</td>
</tr>
<tr>
<td>U1083C</td>
<td>1</td>
<td>neck region</td>
</tr>
<tr>
<td>A1092G</td>
<td>1</td>
<td>neck region</td>
</tr>
<tr>
<td>G1094A</td>
<td>1</td>
<td>neck region</td>
</tr>
<tr>
<td>A1179G(^2)</td>
<td>1</td>
<td>junction of h38-h39-h40</td>
</tr>
<tr>
<td>C1389U</td>
<td>1</td>
<td>neck helix (h28)</td>
</tr>
<tr>
<td>A1410G</td>
<td>4</td>
<td>h44</td>
</tr>
<tr>
<td>C1411U</td>
<td>1</td>
<td>h44</td>
</tr>
<tr>
<td>U1414C(^3)</td>
<td>2</td>
<td>h44</td>
</tr>
<tr>
<td>U1414C(^3), G616A</td>
<td>1</td>
<td>h44; h21</td>
</tr>
<tr>
<td>U789C, U1118C(^2)</td>
<td>1</td>
<td>790 loop; junction of h38-h39-h40</td>
</tr>
<tr>
<td>G1486A(^3), C840U</td>
<td>1</td>
<td>h44; h26</td>
</tr>
<tr>
<td>ASD sequence (GGGGU→GGGAU)</td>
<td>1</td>
<td>3' end</td>
</tr>
</tbody>
</table>

\(^1\)Isolates were only considered independent if they originated from separate preparations of mutagenized pKF207.

\(^2\)A1179 and U1118 form a base-stacking interaction.

\(^3\)U1414 and G1486 form a wobble base pair.
Table 3.2. Effects of 16S rRNA mutations isolated in the genetic screen on translation from ACG, AUC and AUG.

<table>
<thead>
<tr>
<th>16S allele&lt;sup&gt;a&lt;/sup&gt;</th>
<th>β-galactosidase activity&lt;sup&gt;b&lt;/sup&gt;</th>
<th></th>
<th>Relative translation level&lt;sup&gt;c&lt;/sup&gt;</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ACG</td>
<td>AUC</td>
<td>AUG</td>
<td>ACG/AUG</td>
<td>AUC/AUG</td>
</tr>
<tr>
<td>Control</td>
<td>3.1 ± 0.2</td>
<td>8.0 ± 0.4</td>
<td>2300 ± 80</td>
<td>0.0014 ± 0.0001</td>
<td>0.0035 ± 0.0002</td>
</tr>
<tr>
<td>U13G</td>
<td>6.6 ± 0.1</td>
<td>13 ± 0.6</td>
<td>1500 ± 60</td>
<td>0.0043 ± 0.0002</td>
<td>0.0087 ± 0.0005</td>
</tr>
<tr>
<td>U789C</td>
<td>5.0 ± 0.2</td>
<td>9.9 ± 0.1</td>
<td>500 ± 30</td>
<td>0.0092 ± 0.0006</td>
<td>0.018 ± 0.0009</td>
</tr>
<tr>
<td>A790G</td>
<td>7.2 ± 0.1</td>
<td>13 ± 0.1</td>
<td>500 ± 60</td>
<td>0.014 ± 0.002</td>
<td>0.025 ± 0.003</td>
</tr>
<tr>
<td>G886A</td>
<td>8.6 ± 0.1</td>
<td>16 ± 0.1</td>
<td>1800 ± 50</td>
<td>0.0047 ± 0.0002</td>
<td>0.0088 ± 0.0003</td>
</tr>
<tr>
<td>U1083C</td>
<td>4.5 ± 0.1</td>
<td>11 ± 0.2</td>
<td>1700 ± 80</td>
<td>0.0026 ± 0.0001</td>
<td>0.0062 ± 0.0003</td>
</tr>
<tr>
<td>A1092G</td>
<td>4.3 ± 0.1</td>
<td>11 ± 0.2</td>
<td>1800 ± 30</td>
<td>0.0023 ± 0.0001</td>
<td>0.0058 ± 0.0001</td>
</tr>
<tr>
<td>G1094A</td>
<td>4.3 ± 0.1</td>
<td>9.1 ± 0.4</td>
<td>1000 ± 100</td>
<td>0.0043 ± 0.0004</td>
<td>0.0089 ± 0.001</td>
</tr>
<tr>
<td>C1389U</td>
<td>5.0 ± 0.03</td>
<td>12 ± 0.1</td>
<td>1900 ± 90</td>
<td>0.0026 ± 0.0001</td>
<td>0.0059 ± 0.0003</td>
</tr>
<tr>
<td>A1408G</td>
<td>5.0 ± 0.01</td>
<td>12 ± 0.7</td>
<td>1600 ± 100</td>
<td>0.0031 ± 0.0002</td>
<td>0.0075 ± 0.0006</td>
</tr>
<tr>
<td>A1410G</td>
<td>8.5 ± 0.1</td>
<td>15 ± 0.5</td>
<td>1300 ± 30</td>
<td>0.0067 ± 0.0002</td>
<td>0.012 ± 0.0005</td>
</tr>
<tr>
<td>C1411U</td>
<td>4.2 ± 0.03</td>
<td>8.9 ± 0.3</td>
<td>1200 ± 10</td>
<td>0.0034 ± 0.0001</td>
<td>0.0071 ± 0.0003</td>
</tr>
<tr>
<td>U1414C</td>
<td>7.7 ± 0.5</td>
<td>15 ± 0.1</td>
<td>1900 ± 100</td>
<td>0.0041 ± 0.0003</td>
<td>0.0080 ± 0.0004</td>
</tr>
<tr>
<td>G1486A</td>
<td>4.9 ± 0.1</td>
<td>9.7 ± 0.1</td>
<td>2200 ± 20</td>
<td>0.0022 ± 0.0001</td>
<td>0.0045 ± 0.0001</td>
</tr>
</tbody>
</table>

<sup>a</sup>Each 16S allele contains the specialized SD* sequence to allow specific translation of the lacZ reporter gene.

<sup>b</sup>Specialized 16S rRNA (without or with mutations as indicated) was expressed in indicator strains KLF2522, KLF2526 and KLF2528 (harboring SD*-ACG-lacZ, SD*-AUC-lacZ, or SD*-AUG-lacZ, respectively) and specific β-galactosidase activity (based on cleavage of CPRG) was measured. Values represent the mean ± SEM (arbitrary units) of at least 3 independent experiments.

<sup>c</sup>Values represent the quotient of two means ± standard error.
Figure 3.1. Nucleotides identified in this study as important for initiation fidelity cluster to h44, the 790 loop and the neck region.

(A) Mutation sites overlap with class III protection sites. Positions of mutations identified in this study (red triangles); class III protection sites attributed to ligand-induced conformation changes (black squares; Moazed and Noller, 1987).

(B) Positions of mutations in h44 and the 790 loop (Image based on Selmer et al., 2006; PDB 2J02).

(C) Binding of IF1 to the 30S subunit distorts h44 where our mutations map. Blue ribbon, with IF1; Gray ribbon, without IF1 (Image based on Cater et al., 2001; PDB 1HR0 and 1J5E).

(D) Positions of mutations in the neck region (Image based on Selmer et al., 2006; PDB 2J02).
Figure 3.2. Effects of 16S rRNA mutations on the relative level of translation from AUC and ACG.
Ordinate values correspond to the level of β-galactosidase expressed by specialized ribosomes from either SD*-AUC-lacZ (A) or SD*-ACG-lacZ (B) relative to SD*-AUG-lacZ. Gray bars, wild type infC+ background; white bars, infC362 background. Values represent the quotient of two means ± standard error.

3.3.2. Mutant ribosomes can support the growth of E. coli

As a first step toward analyzing the mutant 30S subunits in vitro, we moved a number of mutations (U13G, U789C, A790G, A790U, A790C, G886A, G1094A, G1338A, G1338U, A1408G, A1410G and U1414C) into a Δ7 prrn strain of E. coli, SQZ10. This strain lacks all chromosomal rRNA operons and relies on a single plasmid-borne rRNA operon, and hence a homogeneous population of ribosomes. Each mutation could be readily moved into the Δ7 prrn background, and the resulting strains did not exhibit obvious growth defects (data not shown). All ligand-binding
experiments described below were performed using 30S subunits purified from these Δ7 prrn strains as described (129).

3.3.3. A790G and U789C confer a defect in IF3 binding

A790G was previously shown to decrease the affinity of IF3 for the 30S subunit by ~ 10-fold (129). We presumed that the adjacent mutation U789C similarly inhibited IF3 binding. To test this, we purified subunits containing U789C or each substitution at position 790 and compared their ability to bind IF3 (129). IF3 was labeled with amino-reactive fluorophore (Alexa Fluor 488) and sucrose gradient sedimentation analysis was used to detect the binding of labeled IF3 (IF3-AF) to the mutant 30S subunits. IF3-AF failed to migrate with 30S subunits containing A790G or U789C, indicating that both mutations conferred a substantial defect in IF3 binding. A790U reduced the level of co-migrating IF3-AF while A790C had no appreciable effect on the level of co-migrating IF3-AF (Figure 3.3).
Figure 3.3. Effect of mutations in the 790 loop on the binding of IF3-AF to the 30S subunit.
Fluorescently labeled IF3 (IF3-AF; 1 µM) was incubated with each preparation of 30S subunits (1 µM) at 37°C for 20 min and then subjected to sucrose gradient sedimentation analysis. Fractions (0.5 mL) were collected and the fluorescence intensity (excitation 495 nm; emission 520 nm) of each was quantified to determine the distribution of IF3-AF in the gradient. The position of 30S subunits in the gradient, as identified by A260, is indicated.

3.3.4. A1410G and G886A confer a defect in IF1 binding

A group of mutations mapped to h44 of 16S rRNA (Fig. 1B and C). Structural and biochemical studies have shown that IF1 binds the 30 A site, close to where our h44 mutations cluster (26,34,106). Thus, we considered the possibility that some of our mutations affected IF1 binding. 30S subunits containing U13G, G886A, A1408G, A1410G or U1414C were purified and the affinities of [35S]-IF1 for these subunits were measured using membrane filtration devices (Microcon YM100, MWCO 100K; Millipore Corporation) to separate free and bound [35S]-IF1 (see Materials and Methods). The dissociation constant (K_D) for IF1 binding to control 30S subunits was
estimated at ~ 0.6 μM, consistent with earlier studies using true equilibrium methods (34,188). G886A and A1410G reduced the affinity for IF1 by 2- and 3-fold, respectively, while the other mutations tested did not affect IF1 binding (Figure 3.4A, Table 3.3). 50S subunits failed to bind [35S]-IF1, indicating that interaction between [35S]-IF1 and the 30S subunit was specific. As an additional control, we tested A1408G, a mutation shown previously to confer a defect in IF1 binding (34). A1408G conferred a defect in our hands as well, but of lesser magnitude than that reported previously (Fig. 3.4A, Table 3.3).

Since IF1 and IF3 bind cooperatively to the 30S subunit, we also measured the binding of [35S]-IF1 to mutant 30S subunits in the presence of saturating IF3 (350 nM). IF3 increased the affinity of IF1 for control 30S subunits by ~ 30-fold, consistent with previous studies (188). G886A or A1410G decreased the affinity for IF1 by 2- to 3-fold in the presence of IF3, as was observed in its absence (Table 3.3). The fact that these mutations similarly affected IF1 binding in the absence and presence of IF3 indicates that they do not disrupt the mechanism of cooperativity between the two factors.
Figure 3.4. Effects of 16S rRNA mutations on the binding of IF1 or tRNA\textsuperscript{fMet} to the 30S subunit.

(A) Examples of experiments to estimate the affinity of IF1 for mutant 30S subunits. \(^{35}\text{S}\)-IF1 (0.25 \text{\textmu M}) was incubated with 30S subunits (various concentration) in buffer C [10 mM Tris -HCl (pH 7.5), 100 mM NH\(_4\)Cl, 5 mM Mg(OAc)\(_2\), 6 mM \(\beta\)-ME and 0.03\% Nikkol] at room temperature for 20 min. Microfiltration devices were used to separate free and bound \(^{35}\text{S}\)-IF1. The fraction of \(^{35}\text{S}\)-IF1 bound, averaged from at least 3 independent experiments, was plotted as a function of the concentration of 30S subunits, and the data were fit using KaleidaGraph.

(B) Examples of experiments to estimate the affinity of tRNA\textsuperscript{fMet} for mutant 30S subunits. 3'\(^{32}\text{P}\)-tRNA\textsuperscript{fMet} (5 nM) was incubated with 30S subunits (various concentration) with a model mRNA containing an AUG start codon (m701; 3 \text{\textmu M}) in buffer D [50 mM Tris-HCl (pH 7.5), 100 mM NH\(_4\)Cl, 20 mM MgCl\(_2\), 6 mM \(\beta\)-ME] for 20 min at 37\(^\circ\)C. The free and bound 3'\(^{32}\text{P}\)-tRNA\textsuperscript{fMet} were separated by filtration through a bilayer of nitrocellulose and nylon membranes. The fraction of 3'\(^{32}\text{P}\)-tRNA\textsuperscript{fMet} bound was plotted as a function of 30S subunit concentration, and the data were fit using KladeidaGraph. Equilibrium binding constants (\(K_d\)) listed in Table 2 come from at least three independent binding experiments like the ones shown here.
Table 3.3. Effects of 16S rRNA mutations on ligand binding.

<table>
<thead>
<tr>
<th>16S allele&lt;sup&gt;a&lt;/sup&gt;</th>
<th>IF1 binding&lt;sup&gt;b&lt;/sup&gt; - IF3</th>
<th>+ IF3</th>
<th>tRNA binding&lt;sup&gt;c&lt;/sup&gt;</th>
<th>IF3 binding&lt;sup&gt;d&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control</td>
<td>600 ± 80</td>
<td>30 ± 6</td>
<td>240 ± 10</td>
<td>++</td>
</tr>
<tr>
<td>U13G</td>
<td>700 ± 100</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
</tr>
<tr>
<td>U789C</td>
<td>ND</td>
<td>ND</td>
<td>340 ± 10</td>
<td>-</td>
</tr>
<tr>
<td>A790G</td>
<td>700 ± 100</td>
<td>ND</td>
<td>300 ± 10</td>
<td>-</td>
</tr>
<tr>
<td>A790U</td>
<td>ND</td>
<td>ND</td>
<td>510 ± 30</td>
<td>+</td>
</tr>
<tr>
<td>A790C</td>
<td>ND</td>
<td>ND</td>
<td>520 ± 10</td>
<td>++</td>
</tr>
<tr>
<td>G886A</td>
<td>1100 ± 200</td>
<td>60 ± 9</td>
<td>240 ± 10</td>
<td>++</td>
</tr>
<tr>
<td>G1094A</td>
<td>ND</td>
<td>ND</td>
<td>350 ± 30</td>
<td>ND</td>
</tr>
<tr>
<td>G1338A</td>
<td>700 ± 200</td>
<td>ND</td>
<td>170 ± 10</td>
<td>++</td>
</tr>
<tr>
<td>G1338U</td>
<td>ND</td>
<td>ND</td>
<td>540 ± 50</td>
<td>ND</td>
</tr>
<tr>
<td>A1408G</td>
<td>2100 ± 600</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
</tr>
<tr>
<td>A1410G</td>
<td>2200 ± 500</td>
<td>90 ± 20</td>
<td>210 ± 20</td>
<td>++</td>
</tr>
<tr>
<td>U1414C</td>
<td>600 ± 100</td>
<td>40 ± 20</td>
<td>240 ± 10</td>
<td>++</td>
</tr>
</tbody>
</table>

ND, not determined.

<sup>a</sup> Mutant 30S subunits were purified from a Δ7 prrn strain of *E. coli* (See text).

<sup>b</sup> Values correspond to the estimated $K_D$ ± standard error (nM) derived from curve fitting.

<sup>c</sup> Values correspond to $K_D$ ± SEM (nM) from 3 independent experiments.

<sup>d</sup> Binding of IF3-AF to 30S was assessed by the level of co-migration after sucrose gradient sedimentation assay. ++, extent of co-migration indistinguishable from control; +, extent of co-migration substantially reduced; -, little or no co-migration.
3.3.5. A1408G causes spurious initiation in vivo

A1408G was not isolated in the genetic screen but conferred a defect in IF1 binding. Thus, it was of interest to determine whether A1408G also increased spurious initiation. We engineered A1408G into pKF207, transformed the appropriate indicator strains, and found that, indeed, A1408G increased translation from AUC and ACG to the same degree as other h44 mutations (Figure 3.2, Table 3.2). These data further implicate IF1 in the mechanism of initiation fidelity.

3.3.6. Watson-Crick base pairs at 1413-1487 and 1414-1486 decrease initiation fidelity

In contrast to mutation A1410G, U1414C did not affect IF1 binding despite that it conferred a similar decrease in initiation fidelity (Figure 3.2). Structural studies have shown that bound IF1 locally displaces one strand of h44 along the helical axis, disrupting several base pairs including A1413ºG1487 and U1414•G1486 (26). Both U1414C and G1486A were isolated and are predicted to change the wobble U1414•G1486 base pair to C-G and U-A respectively, suggesting that they may decrease initiation fidelity by stabilizing the paired conformation of h44. With this in mind, we engineered mutations to generate alternative base pairs at positions 1413-1487 and 1414-1486 and tested their effects on initiation fidelity. Substitutions predicted to stabilize base-pairing at either position conferred increased levels of spurious initiation. In contrast, mismatch A1413×A1487 decreased spurious initiation, while G1413×A1487 and U1414×C1486 had no appreciable effect. The only
exceptional case was G1414×G1486, which increased spurious initiation somewhat despite that it would be predicted to destabilize the paired conformation of h44 (Figure 3.5, Table 3.4).

A1410 forms a Watson-Crick base pair with U1490, which remains paired upon IF1 binding (26). A1410G was isolated and is predicted to change the Watson-Crick base pair to a wobble G•U pair. We engineered alternative base pairs at this position and found that U1490C increased spurious initiation by ~4-fold. A1410G-U1490C double mutation, which generates an alternative Watson-Crick base pair at 1410-1490, restored initiation fidelity (Figure 3.5, Table 3.4). These data suggest that a Watson-Crick base pair at 1410-1490 is important for initiation fidelity, consistent with 95% conservation of a Watson-Crick pair at this position (22).
Table 3.4. Effects of engineered mutations in h44 on translation of \textit{lacZ} from ACG, AUC and AUG.

<table>
<thead>
<tr>
<th>16S allele(^a)</th>
<th>(\beta)-galactosidase activity(^b)</th>
<th>Relative translation level(^c)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ACG</td>
<td>AUC</td>
</tr>
<tr>
<td>Control</td>
<td>3.1 ± 0.2</td>
<td>8.0 ± 0.4</td>
</tr>
<tr>
<td>A1413C</td>
<td>8.6 ± 0.7</td>
<td>21 ± 0.5</td>
</tr>
<tr>
<td>A1413G</td>
<td>1.7 ± 0.6</td>
<td>7.0 ± 0.7</td>
</tr>
<tr>
<td>U1414G</td>
<td>1.9 ± 0.5</td>
<td>5.2 ± 0.6</td>
</tr>
<tr>
<td>G1486C</td>
<td>1.5 ± 0.03</td>
<td>3.3 ± 0.5</td>
</tr>
<tr>
<td>G1487U</td>
<td>5.5 ± 0.6</td>
<td>12 ± 0.2</td>
</tr>
<tr>
<td>G1487A</td>
<td>0.7 ± 0.2</td>
<td>1.7 ± 0.2</td>
</tr>
<tr>
<td>U1490C</td>
<td>4.7 ± 0.5</td>
<td>9.9 ± 0.06</td>
</tr>
<tr>
<td>U1414G&amp;G1486C</td>
<td>5.0 ± 0.5</td>
<td>14 ± 0.6</td>
</tr>
<tr>
<td>A1410G&amp;G1490C</td>
<td>2.6 ± 0.6</td>
<td>6.1 ± 0.4</td>
</tr>
</tbody>
</table>

\(^a\)Each 16S allele contains the specialized SD* sequence to allow specific translation of the \textit{lacZ} reporter gene.

\(^b\)Specialized 16S rRNA (without or with mutations as indicated) was expressed in indicator strains KLF2522, KLF2526 and KLF2528 (harboring SD*-ACG-\textit{lacZ}, SD*-AUC-\textit{lacZ}, or SD*-AUG-\textit{lacZ}, respectively) and specific \(\beta\)-gal activity (based on cleavage of CPRG) was measured. Values represent the mean ± SEM (arbitrary units) of at least 3 independent experiments.

\(^c\)Values represent the quotient of two means ± standard error.
3.3.7. Most 30S P-site mutations increase the stringency of start codon selection

It has been proposed that IF3 acts to increase fidelity in part by destabilizing fMet-tRNA\(^{\text{fMet}}\) in the 30S P site (7,8,84). To explore this possibility, we analyzed a number of P-site mutations for their effects on initiation fidelity. Most P-site mutations decreased translation from both SD*-AUC-lacZ and SD*-AUG-lacZ, but decreased translation from AUC to a larger degree (Table 3.5). In other words, these P-site mutations effectively *increase* the stringency of start codon selection.
Table 3.5. Effects of P-site mutations on translation from AUC and AUG.

<table>
<thead>
<tr>
<th>16S allele&lt;sup&gt;a&lt;/sup&gt;</th>
<th>β-galactosidase activity&lt;sup&gt;b&lt;/sup&gt;</th>
<th>Relative translation level&lt;sup&gt;c&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>AUC</td>
<td>AUG</td>
</tr>
<tr>
<td>Control</td>
<td>8.0 ± 0.4</td>
<td>2300 ± 80</td>
</tr>
<tr>
<td>G926A</td>
<td>&lt;0.12</td>
<td>490 ± 90</td>
</tr>
<tr>
<td>G926U</td>
<td>&lt;0.12</td>
<td>470 ± 100</td>
</tr>
<tr>
<td>G926C</td>
<td>&lt;0.12</td>
<td>330 ± 90</td>
</tr>
<tr>
<td>G1338A</td>
<td>11 ± 0.3</td>
<td>2400 ± 60</td>
</tr>
<tr>
<td>G1338U</td>
<td>0.5 ± 0.1</td>
<td>510 ± 30</td>
</tr>
<tr>
<td>G1338C</td>
<td>&lt;0.12</td>
<td>110 ± 20</td>
</tr>
<tr>
<td>C1400A</td>
<td>1.5 ± 0.5</td>
<td>450 ± 20</td>
</tr>
<tr>
<td>C1400U</td>
<td>1.9 ± 0.4</td>
<td>1200 ± 300</td>
</tr>
<tr>
<td>C1400G</td>
<td>0.2 ± 0.05</td>
<td>190 ± 30</td>
</tr>
<tr>
<td>A790G</td>
<td>13 ± 0.1</td>
<td>500 ± 60</td>
</tr>
<tr>
<td>A790C</td>
<td>1.5 ± 0.2</td>
<td>1300 ± 300</td>
</tr>
<tr>
<td>A790U</td>
<td>4.3 ± 0.6</td>
<td>1500 ± 200</td>
</tr>
<tr>
<td>G966A</td>
<td>1.6 ± 0.3</td>
<td>800 ± 100</td>
</tr>
<tr>
<td>G966U</td>
<td>0.5 ± 0.1</td>
<td>700 ± 80</td>
</tr>
<tr>
<td>G966C</td>
<td>0.2 ± 0.01</td>
<td>500 ± 50</td>
</tr>
</tbody>
</table>

<sup>a</sup> Each 16S allele contains the specialized SD* sequence to allow specific translation of lacZ reporter gene.

<sup>b</sup> Specialized 16S rRNA (without or with mutations as indicated) was expressed in indicator strains KLF2526 and KLF2528 (harboring SD*-AUC-lacZ or SD*-AUG-lacZ, respectively) and specific β-galactosidase activity (based on cleavage of chlorophenol red-β-D-galactopyranoside (CPRG)) was measured. Values represent the mean ± SEM (arbitrary units) of at least 3 independent experiments.

<sup>c</sup> Values represent the quotient of two means ± standard error.
3.3.8. Effect of 30S mutations on tRNA\textsuperscript{fMet} binding

Most P-site mutations would be expected to destabilize fMet-tRNA\textsuperscript{fMet} in the 30SIC. A notable exception is G1338A, which stabilizes tRNA\textsuperscript{fMet} in the P site (129). To determine how other 16S rRNA mutations affected tRNA\textsuperscript{fMet} binding, we estimated the affinity of 3'-[\textsuperscript{32}P]-tRNA\textsuperscript{fMet} for the 30S P site in various mutant 30S subunits using filter binding as described (129) (Figure 3.4B, Table 3.3). Of the mutations tested, G1338U, A790C, and A790U were the most defective in tRNA\textsuperscript{fMet} binding. A790G, U789C, and G1094A conferred a modest defect, while A1410G, U1414C and G886A showed no apparent defect.
3.4. Discussion

In this study, a number of mutations in 16S rRNA that increase translation from AUC were isolated. These mutations cluster to distinct regions that overlap remarkably well with previously identified class III protection sites. Class III sites were defined as those nucleotides protected from chemical probes by multiple ligands (e.g., 50S subunits, tRNA, and various antibiotics). Because these ligands can bind simultaneously to the 30S subunit, these class III protections were attributed to ligand-induced conformational changes, rather than to direct protections (104). The fact that many of our mutations map near class III sites suggests that they perturb rRNA rearrangements important for initiation.

Nucleotides 783-793 constitute a major portion of the IF3-binding site (35,42,106). Two mutations (U789C and A790G) lie in this region and confer the strongest phenotypes (Fig. 2). A790G and the adjacent G791A were shown previously to decrease the affinity of IF3 for the 30S subunit by ~ 10-fold (129,166), and U789C appears similarly defective in IF3 binding (Fig. 3). It is well known that IF3 regulates initiation complex formation to prevent initiation from non-canonical start codons (62,133,164). Chromosomal mutations that increase spurious initiation events have been mapped almost exclusively to the IF3 gene (infC), in line with our observation that 16S rRNA mutations that target the IF3-binding site are the most defective in initiation fidelity. When mutations in IF3 (infC362) and its binding site (A790G) are present in the same cell, AUG is discriminated from AUC by a factor of 10 rather than 300 (Fig. 2), underscoring the importance of IF3 in regulating initiation.
A second cluster of mutations surrounds the conserved A1413-G1487 base pair in h44. Structural studies reveal that IF1 binds the 30S A site and induces a distortion of h44, involving displacement of one strand away from the helical axis and consequential disruption of base pairs A1413-G1487 and U1414-G1486 (26). This rearrangement of h44 is centered right where our mutations cluster, strongly implicating IF1 in the mechanism of initiation fidelity. While certain mutations (i.e., A1410G, A1408G, G886A) decrease the affinity of IF1 for the subunit, U1414C does not. These data suggest that defects in fidelity might stem from inhibition of IF1 binding, the factor-induced rRNA rearrangement, or both. With this in mind, we tested a number of substitutions predicted to stabilize base pairs 1413-1487 and 1414-1486. In each case, an increase in spurious initiation was observed, consistent with the idea that initiation is regulated by the factor-induced rearrangement of h44. By contrast, engineered mismatches G1413-G1487, A1413-A1487, and U1414-C1486 did not decrease initiation fidelity. The only anomalous result came from the G×G mismatch at 1414-1486, which conferred a slight increase in spurious initiation. It is possible this purine-purine pair generates a unique conformation responsible for its phenotype, which fails to follow the trend.

A recent biochemical study suggests that IF1 helps IF3 negatively regulate 70SIC formation to enhance start codon selection (102). Changing the start codon from AUG to AUU slowed both IF3 dissociation and subunit joining, and omission of IF1 from the reaction increased the rates of these events in the presence of AUU. Interestingly, the ability of IF1 to inhibit IF3 release and subunit joining was lost in the
presence of streptomycin. It was proposed that IF1 induces an "initiation-unfavorable" conformation of the 30S subunit, which could be reversed by cognate codon-anticodon pairing in the 30SIC or by streptomycin. Our data are consistent with this model and further suggest that the structural corollary of the "initiation-unfavorable" conformation is the distorted state of h44, in which base pairs 1413-1487 and 1414-1486 are unpaired, as observed in the co-crystal structure of 30S•IF1 (26). Binding of either IF1 or IF3 to the 30S subunit causes G1487 to become hyper-reactive to kethoxal, suggesting that both factors stabilize the unpaired state of h44 (34,106). Streptomycin, on the other hand, protects G1487 from kethoxal modification, as does tRNA (presumably P-tRNA) (104). These ligands do not directly interact with G1487; instead, they most probably stabilize the paired state of h44. Clues of how the unpaired state of h44 might inhibit 70SIC formation come from structural studies. The center of the IF1-induced distortion lies between inter-subunit bridges B2a and B3 (25,80,141), suggesting that docking of the 50S subunit might be inhibited by displacement of the 16S rRNA nucleotides that contribute to these bridges. The idea that IF1 helps regulate 50S docking is consistent with evidence that IF1 and IF3 work together to facilitate subunit splitting (53,111,121).

Our study provides evidence that IF1 contributes to initiation fidelity *in vivo*. A previous report showed that several mutations in *infA* failed to increase translation from GUG and UUG (32), and those data have been taken by some investigators as evidence that IF1 plays no role in start codon selection. However, mutations in *infC* also fail to increase translation from GUG and UUG (133,164), indicating that IF3 does not discriminate against these naturally-occurring non-AUG start codons. In our view,
there is no reason to suspect that IF1 would discriminate against these start codons either. For technical reasons, Croitoru et al. were unable to measure the effects of infA alleles on spurious initiation (i.e., translation from Class IIa codons such as AUU, AUC, and ACG). Whether these infA alleles increase spurious initiation remains an open question worth addressing.

One mutation, U13G, maps to the central pseudoknot and has been isolated previously in selections for streptomycin resistance (63). Streptomycin binds adjacent to U13 (25), and other base substitutions at this position have been shown to inhibit streptomycin binding (124). Although not yet understood, this connection between streptomycin and initiation fidelity in vivo is noteworthy, particularly in light of the effects of streptomycin on the ability of IF1 to regulate initiation in vitro (102).

A third cluster of mutations localizes to the neck region, and these mutations may act by altering the dynamics of the head domain. Structural and biochemical studies have revealed that the head can assume various positions, and its orientation depends on which ligands (e.g., 50S subunit, translation factors, tRNA, antibiotics) are bound to the 30S subunit (12,26,67,118,140). Because P-site tRNA influences the orientation of the head domain (12), we tested whether G1094A increases the intrinsic stability of the initiator tRNA in the 30S subunit. On the contrary, G1094A decreases the affinity of tRNA^{fMet} for the 30S P site, arguing against the idea that G1094A acts in a manner analogous to G1338A. Upon subunit joining, the head domain of the 30S subunit tilts toward the 50S subunit (67), which raises the possibility that these neck mutations might promote premature 50S docking. While we have not yet investigated
the effects of any of our mutations on subunit association (in the absence or presence of factors), such experiments may prove quite informative.

A number of P-site mutations (e.g., G926A, G926C, G926U, m\(^2\)G966C, G1338U) increase the stringency of start codon selection (Table 3). These mutations are likely to destabilize fMet-tRNA\(^{\text{fMet}}\) in the 30SIC, consistent with the fact that G1338U destabilizes tRNA\(^{\text{fMet}}\) in the 30S P site (Fig. 4B, Table 2). A notable exception is G1338A, which decreases the stringency of start codon selection and stabilizes tRNA\(^{\text{fMet}}\) in the 30S P site (129). These data support a model in which discrimination against non-canonical start codons relies, at least in part, on the mutual stabilization of mRNA and fMet-tRNA\(^{\text{fMet}}\) in the 30SIC that comes from codon-anticodon interaction (7,84). In G1338U subunits, fMet-tRNA\(^{\text{fMet}}\) is less stable in the 30SIC, and consequently the probability of productive initiation depends more heavily on cognate codon-anticodon pairing. In G1338A subunits, the lifetime of 30SIC with fMet-tRNA\(^{\text{fMet}}\) paired to AUC is lengthened, allowing a higher level of spurious initiation.

Indeed, our data are consistent with the proposal that IF3 can repress translation from non-canonical start codons by simply destabilizing fMet-tRNA\(^{\text{fMet}}\) in the 30SIC (7,84).

A790 interacts with both IF3 and fMet-tRNA\(^{\text{fMet}}\) (35,106,141), making it potentially more difficult to predict the effects of mutations at this position on 30SIC stability. However, of the three substitutions, A790G destabilizes IF3 the most and tRNA\(^{\text{fMet}}\) the least, and it is the only mutation that increases spurious initiation. A790C has no apparent effect on IF3 binding, destabilizes tRNA\(^{\text{fMet}}\) to the greatest degree, and decreases spurious initiation. A790U has an intermediate effect on IF3 binding.
destabilizes tRNA^{fMet} to the same degree as A790C, and confers the least effect on spurious initiation. All these results are in line with the model proposed by Ehrenberg in which IF3 and fMet-tRNA^{fMet} bind the 30S subunit in an anti-cooperative manner, and their interplay during initiation determines the stringency of start codon selection (7). Clearly, it will be of interest in the future to directly test the effects of these mutations on fMet-tRNA^{fMet} stability in the 30SIC.

Although ribosomes with certain P-site mutations can discriminate AUG from AUC very well, these mutant ribosomes translate from AUG with reduced efficiency (Table 3). Thus, it seems that the affinity of fMet-tRNA^{fMet} for the 30S P site is tuned to balance efficiency and accuracy during initiation.
Chapter 4

Role of the 30S Subunit in Start Codon Selection

4.1. Introduction

In bacteria, translation initiation involves recognition of the start codon by the initiator tRNA (fMet-tRNA) in the ribosome. This process occurs via two major steps and is kinetically controlled by three initiation factors (IF1, 2 and 3) (15,57). In the first phase, fMet-tRNA binds the 30S P site and its anticodon pairs with the start codon, which forms the 30S initiation complex (30SIC). In the second phase, the 50S subunit associates with the 30SIC, resulting in GTP hydrolysis and factor release, to generate 70S initiation complex (70SIC), ready for first round of elongation.

Faithful selection of the start codon is essential for establishing the correct translational frame. Base complementarity in the codon-anticodon helix is the primary way in which fMet-tRNA recognizes the start codon. While AUG is the predominant start codon, UUG and GUG are quite commonly used (41). This implies wobble base pairs are tolerated at the first position of the start codon. In contrast, single substitutions at other positions (e.g., AUU, AUC) are disfavored, typically reducing the level of gene expression by 100-fold (129,130,133,164). In mRNA, other cis elements are involved in initiation to “find” the start codon. One example is a purine-rich
sequence, termed the Shine-Dalgarno (SD) (148), which is located 6-12 nucleotides upstream from the start codon (150). The SD pairs with the 3’end of 16S ribosomal RNA (rRNA) to facilitate positioning of the start codon in the P site (21,79,160). Based on current models of 30SIC formation, mRNA and fMet-tRNA bind to the 30S subunit in a random order, followed by a rate-limiting rearrangement step leading to start codon recognition (57). Implicit in these models is that each ligand (fMet-tRNA and mRNA) interacts initially in a kinetically labile manner, and then both ligands are mutually stabilized by codon-anticodon pairing in the P site. It can also be inferred that the rearrangement step is codon dependent, in a way favoring canonical start codon.

Mutations that increase translation from non-canonical start codons have been mapped to infC, the gene encoding IF3 (133,164). While these studies implicate IF3 in initiation fidelity, the mechanism by which IF3 enhances the start codon selection is not clear. It has been proposed that IF3 somehow interacts with the codon-anticodon helix, favoring complementarity in relation to mismatches (99). A recent biochemical study showed that IF3 slowed down the 50S docking and destabilized 30S·mRNA·tRNA complex programmed with AUG (30S·AUG·fMet-tRNA). Thus, an alternative mechanism was suggested in which the main contribution of IF3 is to tune the kinetics of initiation in a way that the intrinsic difference in stability between 30S·AUG·fMet-tRNA and 30S·AUU·fMet-tRNA can be efficiently utilized for accuracy (7).

Recent studies indicate that the subunit-joining step of the process is negatively regulated by not only IF3 but also IF1. IF1 slowed 50S docking when the start codon is
changed from AUG to AUU and strongly inhibited docking with m002, a model mRNA containing a strong SD closely juxtaposed to the start codon (102). These inhibitory effects of IF1 were eliminated with the antibiotic streptomycin. It was proposed that IF1 negatively regulates initiation by inducing a docking-unfavorable conformation in the 30SIC, which can be shifted to a docking-favorable state by canonical codon-anticodon in the P site or by streptomycin. Here, we provide evidence for such a conformational switch in h44, regulated oppositely by IF1 and start codon recognition codon-anticodon pairing in the P site. We also show that IF3 increases the accuracy of initiation by uniformly destabilizing fMet-tRNA rather than by enhancing the intrinsic selectivity of the programmed P site.

4.2. Materials and Methods

4.2.1. Bacterial strains

\textit{E. coli} Δ7 strains were made using SQZ10 as described previously (130).

CSH142 (\(F-\ ara \Delta(gpt-lac)5\))

4.2.2. Preparation of fMet-tRNA\textsuperscript{fMet} and mRNA

Aminoacylation and formylation of tRNA\textsuperscript{fMet} was performed as described (175). tRNA\textsuperscript{fMet} (10 \(\mu\)M, Chemical Block, Russia) was incubated with methionine (100 \(\mu\)M), neutralized N\(^{10}\)-formyl-tetrahydrofolate (300 \(\mu\)M), MetRS (1 \(\mu\)M), and MTF (1 \(\mu\)M) in charging buffer (100 mM HEPES-KOH, pH7.6; 10 mM ATP; 10mM KCl; 20 mM MgCl\(_2\), 1 mM DTT) at 37 °C for 20 min. The reaction was extracted with phenol and
CHCl₃, and tRNA was ethanol precipitated. The product was dissolved with NaOAc buffer (2 mM, pH 5.2). The aminoacylation and formylation efficiency was > 95% according to acid gel and TLC results.

The mRNAs, which are derived from gene 32, were transcribed *in vitro* using T7 polymerase and single-cut plasmid templates. The product was purified as described (44). The initiation regions corresponding to each mRNA are shown below:

mRNA(AUG): AAAGGAUAAAAAAUGGUUACUUUAAACUCU
mRNA(AUC): AAAGGAUAAAAAUCGUUACUUUAAACUCU

The SD is shadowed and start codon is underlined.

### 4.2.3. Initiation factors

Each initiation factor (IF1, 2 or 3) was overexpressed in strain BL21/DE3 from pET24b-based constructs (kindly provided by Dallas A. and Noller H.). IF1 was purified by resource S column as described (130). His-tagged IF2 (α form) and IF3 were purified by Ni-NTA agarose (Qiagen) method as described (35). The protein concentrations were estimated by Bradford assay. The purities were > 95% judged by Coomassie blue-stained SDS-PAGE gel.

### 4.2.4. Ribosomal subunits
30S subunits (control or mutants) were purified from *E. coli* Δ7 strains as described (129). 50S subunits were purified from *E. coli* CSH142 strain followed the same procedure as in 30S subunit preparation, except that 50S peak was collected.

### 4.2.5. Light scattering to monitor 50S docking

Light scattering experiments were performed as described (102). Typically, 30SIC was formed by incubating heat-activated 30S (0.08 µM) with fMet-tRNA<sup>Met</sup> (0.2 µM), mRNA (0.5 µM), initiation factors (0.2 µM each), GTP (0.1 mM) in buffer A (50 mM Tris-HCl, pH 7.5; 30 mM KCl; 70 mM NH₄Cl; 7 mM MgCl₂ and 1 mM DTT) at 37 °C for 30 min in 1 mL volume. For each mixing experiments, ~60 µL of 30SIC was rapidly mixed with equal volume of 50S (concentration as indicated) in buffer A, using a SX20 stopped flow apparatus (Applied Photophysics). Exciting light at 430 nm was applied to the mixing chamber and scattered light was monitored without a cutoff filter at an angle perpendicular to the excitation beam as a function of time. The light scattering signal was plotted against time and the data were fit to either a single exponential function: $F = F_\infty + A \times \exp(-k_{app} \times t)$ or double exponential function $F = F_\infty + A_1 \times \exp(-k_{app1} \times t) + A_2 \times \exp(-k_{app2} \times t)$ depending on the nature of the curve ($F_\infty$, the final signal; $A$, the amplitude of signal change; $k_{app}$, the apparent rate; t, time).

### 4.2.6. Chemical modification of 30SIC

Chemical modification experiments were performed as described (100). Briefly, 100 µL of 30SIC was formed by incubating 30S (0.25 µM) with fMet-tRNA<sup>Met</sup> (0.5 µM), mRNA (0.5 µM), initiation factors (0.2 µM each), GTP (0.1 mM) in buffer A (50 mM Tris-HCl, pH 7.5; 30 mM KCl; 70 mM NH₄Cl; 7 mM MgCl₂ and 1 mM DTT) at 37 °C for 30 min in 1 mL volume. For each mixing experiments, ~60 µL of 30SIC was rapidly mixed with equal volume of 50S (concentration as indicated) in buffer A, using a SX20 stopped flow apparatus (Applied Photophysics). Exciting light at 430 nm was applied to the mixing chamber and scattered light was monitored without a cutoff filter at an angle perpendicular to the excitation beam as a function of time. The light scattering signal was plotted against time and the data were fit to either a single exponential function: $F = F_\infty + A \times \exp(-k_{app} \times t)$ or double exponential function $F = F_\infty + A_1 \times \exp(-k_{app1} \times t) + A_2 \times \exp(-k_{app2} \times t)$ depending on the nature of the curve ($F_\infty$, the final signal; $A$, the amplitude of signal change; $k_{app}$, the apparent rate; t, time).
µM), mRNA (1 µM), initiation factors (0.5 µM each), GTP (0.1 mM) in buffer B (80 mM K-cacodylate, pH 7.5; 30 mM KCl; 70 mM NH₄Cl; 7 mM MgCl₂ and 1 mM DTT) at 37 °C for 10 min and then at room temperature for 30 min in a volume of 50 µL. 4 µL of diluted DMS (1 µL stock added in 24 µL mixture of ethanol and buffer B (1:1)) was added to the reaction at 37 °C for ~ 8 min. Then, buffer D (1 M Tris, pH 7.5; 200 mM EDTA; 100 mM 2-mercaptoethanol) was added to stop the reaction. Then, the reaction was precipitated in ethanol and resuspended in buffer E (300 mM NaOAC, pH 6.5; 5 mM EDTA; 0.5% SDS). After extraction three times in phenol and two times in chloroform, RNA was precipitated with ethanol and dissolved with ddH₂O to a concentration of ~0.2 µM. Kethoxal modification was performed essentially the same except that K-borate (25 mM, pH 7.0) was added to the reaction throughout after modification to stabilize the conjugates. In NMIA experiments, K-cacodylate was replaced by HEPES (pH 7.9, 80 mM) and modification was incubated for 45 min instead of 8 min. We observed similar results of kethoxal modification under HEPES (pH 7.9, 80 mM) condition. In hydroxyl radicals probing experiments, hydroxyl radicals were generated by adding Fe(II)-EDTA (2mM, preformed with (NH₄)₂Fe(SO₄)₂ and Na₄EDTA), 0.1% H₂O₂ and ascorbic acid (10 mM) to the 30SIC reactions in buffer B. Modifications of fMet-tRNA by DMS and NMIA were preformed in a similar way as in modification of 16S rRNA, except that stoichiometric concentration of fMet-tRNA (0.25 µM) was used in 30SIC formation.

4.2.7. Primer extension
Primer extension of 16S rRNA was performed as described (100). Primers (1495, 1394, 1199, 1046, 972, 880, 794, 680, 530, 495, 342, 172 and 45) were designed to cover 90% of 16S rRNA region. Primers were typically ~20 nt long, and each primer was named by the position of first cDNA nt incorporated. A primer annealing to the 3’ end of fMet-tRNA is used to detect the modifications of anticodon of fMet-tRNA.

4.2.8. Toeprinting

30SIC was formed by incubating 30S (1 µM) with fMet-tRNA\textsuperscript{*Met} (1.5 µM), mRNA (0.1 µM), IFs (1.5 µM each), GTP (0.1 mM) at 37 °C for 3 min in buffer A. \textsuperscript{32}P-labeled primer was pre-annealed to the mRNA and extension was performed as described (149).

The equilibrium constant for 30S·mRNA·fMet-tRNA ternary complex formation (K\textsubscript{TC}) can be expressed as K\textsubscript{TC}=[ABC]/[A][B][C], where [ABC] is the concentration of ternary complex, [A] is the concentration of free mRNA, [B] is the concentration of free tRNA, and [C] is the concentration of free fMet-tRNA. Under the conditions employed, the total concentrations of tRNA and 30S exceed that of mRNA by >10-fold. Hence, [B] and [C] can be considered as constants b and c, and K\textsubscript{TC}=[ABC]/[A]·b·c, A relative equilibrium constant (K\textsubscript{rel}) can be defined, K\textsubscript{rel}=K\textsubscript{TC}·b·c, giving K\textsubscript{rel}=[ABC]/[A] (equation 1). In Figure 4.8, the fraction of signal corresponding to the toeprint is defined by F=[ABC]/([A]+[ABC]) (equation 2). Combining equation 1 and 2, gives K\textsubscript{rel}=1/(1/F-1).
4.2.9. Micro-con assays

30SIC was formed by incubating 30S (0.25 µM) with [35S]fMet-tRNAfMet (0.5 µM), mRNA (1 µM), IFs (0.375 µM each), GTP (0.1 mM) at 37 °C for 10 min in buffer A. The reaction was loaded onto micro-con YM 100 filter units (Millipore) and spun for 1600 g for 5 min to separate the free and bound [35S]fMet-tRNAfMet. The fractions of the retained and flow through portion were collected and the signal was measured in liquid scintillation assay. The fraction bound is defined as $F_{\text{retained}} / (F_{\text{retained}} + F_{\text{flow through}})^{1/2}$.

4.3. Results

4.3.1. Mutations in h44 stimulate 50S docking

Several mutations that decrease the fidelity of start codon selection have been isolated in h44 (130). These mutations map near the IF1-binding site, between nucleotides that contribute to intersubunit bridges B2a and B3 (Figure 4.1) (128,184). To investigate the effects of these mutations on the subunit-joining step of initiation, control (WT) and mutant 30S subunits were purified and rates of docking of 50S subunits to 30S initiation complexes containing a cognate [30SIC(AUG)] or near-cognate [30SIC(AUC)] start codon were measured by stopped-flow spectroscopy. Mixing of 50S subunits (0.11 µM) with 0.04 µM of the WT 30SIC(AUG) resulted in a biphasic increase in light scattering (LS), in which the fast (4 s⁻¹) and slow (0.8 s⁻¹) phases accounted for ~40% and ~60% of the total signal increase, respectively (Figure 4.1).
4.2, Table 4.1). Biphasic increases in LS with similar apparent rates have been reported previously, using 30SICs formed with various mRNAs (102,120). Although the basis of this phenomenon remains unclear, the slow phase may reflect a subpopulation of docking-incompatible complexes that require, for example, an additional component or conformational change before rapid docking can occur. No LS increase was seen in the absence of fMet-tRNA, indicating that 50S docking is completely tRNA-dependent under these conditions. When the start codon was changed to AUC and the experiment repeated, the rate of 50S docking decreased substantially and the curve became monophasic ($k_{app} = 0.03 \text{ s}^{-1}$) (Figure 4.2, Table 4.1). This strong inhibition of 50S docking in the presence of a near-cognate start codon is consistent with previous studies (55,102). Mutation A1413C stimulated 50S docking, an effect most pronounced with the near-cognate start codon AUC. For A1413C 30SIC(AUC), the LS increase was biphasic, with $k_{app1}$ 20-fold faster than $k_{app}$ for WT 30SIC(AUC). Values for $k_{app2}$ and its amplitude ($A_2$) were very similar to the rate and amplitude of the control reaction. In the presence of the cognate start codon AUG, mutation A1413C stimulated 50S docking somewhat, increasing $A_1$ by 70% and by modestly increasing both $k_{app1}$ and $k_{app2}$. We tested the effects of two other h44 mutations, A1410G and U1414C. Like A1413C, these mutations stimulated 50S docking, particularly in the presence of a near-cognate start codon (data not shown). These data are consistent with the idea that these mutations decrease the accuracy of start codon selection in vivo by affecting 50S docking during initiation.
It was shown recently that IF1 inhibits 50S docking to 30SICs containing the near-cognate start codon AUU (102), suggesting a role for IF1 in start codon selection. This raised the possibility that mutation A1413C stimulated 50S docking indirectly by destabilizing IF1 in the 30SIC. However, the effects of A1413C on docking were not suppressed by adding a 12-fold higher concentration of the factor (1 µM; 25-fold excess over 30S) (data not shown), arguing against the idea that A1413C acts by destabilizing IF1. In our experimental system, omission of IF1 from the WT 30SIC(AUC) increased the apparent rate of 50S docking, consistent with the earlier work (102). Notably, the stimulatory effect of IF1 omission was less than that of A1413C, suggesting that the mutation does more than reverse the inhibitory effects of IF1.
Figure 4.1. Nucleotides showing different chemical reactivity in 30SIC (AUG) and 30SIC(AUC) locate in h44, P site and IF3 binding site.

The 30S subunit (only 16S rRNA is shown, colored gray) is viewed from the subunit interface. IF1 is colored wheat. P-tRNA and mRNA are colored blue and green, respectively. Mutations identified previously (130) that increase translation from non-canonical start codon are highlighted in magenta. Nucleotides identified which show different chemical reactivity in the 30SIC(AUG) and 30SIC(AUC) reactions are highlighted in red. The intersubunit bridges according to (184) are colored cyan. The IF3 binding site is roughly defined according to (35). Image is generated based on (26,141) (PDB 2J02 and 1HR0).
Figure 4.2. 16S rRNA mutation A1413C stimulates association of the 50S subunit to the 30SIC(AUC).

30SIC was formed by incubating 30S subunits (WT or mutation as indicated, 0.075 µM) with IFs (IF1, 2, 3; 0.2 µM each or as indicated), GTP (0.1 mM), fMet-tRNA (0.2 µM) and mRNA (containing either AUC or AUG as indicated, 0.5 µM) at 37 °C for 30 min in buffer containing Tris-HCl pH 7.5 (50 mM), KCl (30 mM), NH₄Cl (70 mM), MgCl₂ (7 mM) and DDT (1 mM). 30SIC was rapidly mixed with equal volume of 50S subunits (0.225 µM) using an SX-20 stopped flow apparatus and scattered light signals were measured. Ordinate values correspond to the relative light scattering units.

Table 4.1. Summary of the apparent rates of 50S docking.

<table>
<thead>
<tr>
<th>Start Codon</th>
<th>30S subunit</th>
<th>$k_{app1}$ (s⁻¹)</th>
<th>$A_1$</th>
<th>$k_{app2}$ (s⁻¹)</th>
<th>$A_2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>AUC</td>
<td>WT</td>
<td>0.034 ± 0.0007</td>
<td>0.017</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>A1413C</td>
<td>0.58 ± 0.01</td>
<td>0.01</td>
<td>0.052 ± 0.005</td>
<td>0.017</td>
</tr>
<tr>
<td>AUG</td>
<td>WT</td>
<td>4.0 ± 0.6</td>
<td>0.01</td>
<td>0.68 ± 0.06</td>
<td>0.017</td>
</tr>
<tr>
<td></td>
<td>A1413C</td>
<td>5.0 ± 0.3</td>
<td>0.017</td>
<td>1.0 ± 0.1</td>
<td>0.014</td>
</tr>
</tbody>
</table>

For each independent experiment, at least 4 mixing shots were performed. Data were evaluated by fitting into either a single or double exponential function, yielding the apparent rates and amplitudes. Rates from all mixing experiments were averaged to represent the apparent rates for that independent experiment. Three independent experiments were performed to generate the parameters shown (mean ± SEM).
4.3.2. Structural differences in 30S initiation complexes containing the start codon AUG versus AUC

The experiments described above provided evidence that the rate of 50S docking is influenced by the conformational state of h44. However, whether a conformational change in h44 normally occurs in response to cognate codon-anticodon pairing to help regulate 50S docking remained unclear. To address this question, we compared the chemical reactivity of 16S rRNA in 30SIC(AUG) versus 30SIC(AUC), using dimethyl sulfate (DMS), kethoxal (KE), and N-methylisatoic anhydride (NMIA) as probes. DMS primarily targets N1 of A, KE targets N1 and N2 of G, and NMIA targets the 2’ OH of ribose (100,178). The former two probes have been employed quite extensively to analyze 30S-ligand interactions; however, those studies were largely performed in the context of binary or ternary complexes as opposed to the complete 30SIC (103,105). We performed initial experiments in which the 30S subunit, 30SIC(AUG), and 30SIC(AUC) were probed in parallel, rRNA was extracted, and modifications in >90% of the 16S rRNA molecule were identified and quantified by primer extension. Of ~110 nucleotides targeted by these probes, 6 were differentially modified in the 30SIC(AUG) and 30SIC(AUC) reactions, data that were confirmed in subsequent experiments (see below). One of the nucleotides identified was A1408, which contributes to intersubunit bridge B2a (Figure 4.1).

4.3.3. A1408 exhibits reduced reactivity in the presence of AUG
Previous DMS and KE footprinting studies showed that binding of IF1 to the 30S subunit protects several nucleotides (e.g., G530, A1492, and A1493) and enhances the reactivity of others (e.g., A908, A909, A1408, A1413, and G1487) (34,106). These data can be rationalized in light of the co-crystal structure of the IF1-bound 30S subunit (26). IF1 forms specific contacts to G530, A1492, and A1493. These interactions require that the latter two nucleotides flip out from h44, which increases the exposure of A1408, lying on the opposite strand. IF1 also distorts h44 further “down,” which alters backbone contacts to A908-A909 and disrupts the noncanonical pair A1413-G1487, explaining the chemical enhancements of these bases.

In the 30SIC(AUG), A1408 was found to be hyperreactive to DMS, but the level of modification was reduced by about 50% compared to that in 30SIC(AUC) or in complexes formed in the absence of fMet-tRNA (Table 4.2, Figure 4.3A top). Several observations indicate that the reduced reactivity of A1408 is due to a conformational change rather than a decrease in the level of IF1 binding. First, the IF1-dependent enhancement at A908 was indistinguishable among the complexes (Table 4.2, Figure 4.3A bottom). Second, no loss of protection was seen at G530 in 30SIC(AUG). While some degree of G530 protection in the 30SIC is probably due to P-tRNA binding (Table 4.2, Figure 4.4A and B), in line with previous observations (105), IF1 is presumably responsible for much of the observed protection. We also tried to compare protections at A1492 and A1493, but obtaining quantitative data was difficult due to a strong inhibition of primer extension caused by the naturally modified nucleotide $m^3$G1497. Third, when a 20-fold higher concentration of IF1 (10 µM) was
used, the same modification pattern was obtained (Figure 4.3A, lanes 11 and 12). Together these observations suggest that cognate start codon recognition alters the conformation of A1408 in the 30SIC. Excess fMet-tRNA (3 μM) also failed to change the pattern of reactivity at A1408, suggesting that the effects of a mismatch between codon and anticodon cannot be suppressed by promoting fMet-tRNA binding (Figure 4.3A, lanes 9 and 10).
Table 4.2. Quantification of chemical reactivity of 16S rRNA nucleotides in various 30S complexes

<table>
<thead>
<tr>
<th>Nucleotide</th>
<th>Modification chemical</th>
<th>30S(IFs)</th>
<th>30S(IFs, mRNA)</th>
<th>30SIC(AUG)</th>
<th>30SIC(AUC)</th>
</tr>
</thead>
<tbody>
<tr>
<td>G369</td>
<td>Hydroxyl radicals</td>
<td>0.47 ± 0.06</td>
<td>0.52</td>
<td>0.41 ± 0.04</td>
<td>0.63 ± 0.04</td>
</tr>
<tr>
<td>G530</td>
<td>Kethoxal</td>
<td>0.39 ± 0.1</td>
<td>0.29 ± 0.05</td>
<td>0.24 ± 0.06</td>
<td>0.29 ± 0.06</td>
</tr>
<tr>
<td>A532</td>
<td>DMS</td>
<td>0.85 ± 0.05</td>
<td>1.1</td>
<td>0.86 ± 0.07</td>
<td>1.3 ± 0.07</td>
</tr>
<tr>
<td>G700</td>
<td>Kethoxal</td>
<td>0.44 ± 0.04</td>
<td>0.43 ± 0.04</td>
<td>0.52 ± 0.1</td>
<td>0.45 ± 0.09</td>
</tr>
<tr>
<td>U701</td>
<td>NMIA</td>
<td>0.28</td>
<td>0.18</td>
<td>0.59</td>
<td>0.26</td>
</tr>
<tr>
<td>A794</td>
<td>DMS</td>
<td>1.1 ± 0.1</td>
<td>0.95</td>
<td>0.32 ± 0.08</td>
<td>0.59 ± 0.07</td>
</tr>
<tr>
<td>A908</td>
<td>DMS</td>
<td>2.2 ± 0.3</td>
<td>2.7 ± 0.2</td>
<td>2.2 ± 0.5</td>
<td>2.6 ± 0.4</td>
</tr>
<tr>
<td>A909</td>
<td>DMS</td>
<td>2.8 ± 0.3</td>
<td>4.3 ± 0.5</td>
<td>3.1 ± 0.7</td>
<td>3.9 ± 0.6</td>
</tr>
<tr>
<td>G926</td>
<td>Kethoxal</td>
<td>1.1 ± 0.1</td>
<td>0.48 ± 0.04</td>
<td>0.24 ± 0.05</td>
<td>0.40 ± 0.05</td>
</tr>
<tr>
<td>G1338</td>
<td>Kethoxal</td>
<td>0.92 ± 0.05</td>
<td>0.68 ± 0.08</td>
<td>0.5 ± 0.1</td>
<td>0.8 ± 0.1</td>
</tr>
<tr>
<td>A1408</td>
<td>DMS</td>
<td>5.4 ± 0.6</td>
<td>6.2 ± 0.3</td>
<td>3.6 ± 0.4</td>
<td>5.5 ± 0.5</td>
</tr>
<tr>
<td>A1413</td>
<td>DMS</td>
<td>4.4 ± 0.6</td>
<td>5.1 ± 0.5</td>
<td>3.6 ± 0.5</td>
<td>4.6 ± 0.4</td>
</tr>
</tbody>
</table>

Chemical probing was done as described in Methods. Intensity of a band at a specific position was quantified relative to two nearby natural stops (K bands), above and below the band of interest. Data shown are normalized to the 30S only control (set to 1.0). Values > 1 indicate ligand-induced increases in reactivity; values <1 indicate ligand-induced decreases in reactivity. Values represent at mean ± SEM of at least 3 independent experiments. Numbers shown in bold indicate that the difference between 30SIC(AUG) and 30SIC(AUC) was verified statistically at the 95% confidence level.
4.3.4. A1408 exhibits reduced reactivity in the presence of mutation A1413C

Mutation A1413C alters h44 in a way that stimulates 50S docking (see above). To determine whether A1413C and cognate start codon recognition similarly influence the structure of h44, DMS probing experiments were performed with 30SICs in the presence and absence of the mutation (Figure 4.3B, right panel). Indeed, in complexes formed with A1413C subunits, the reactivity of A1408 was markedly reduced, regardless of the start codon sequence. (Figure 4.3B, line 5 and 6). Mutation A1410G, which also promotes 50S docking, shows a similar reduction in A1408 reactivity (data not shown). These data suggest that A1413C and A1410G cause h44 to adopt a conformation that resembles a state normally induced by cognate start codon recognition.

4.3.5. A1408 exhibits reduced reactivity in the presence of streptomycin

Milon et al. proposed that IF1 negatively regulates initiation by inducing a docking-unfavorable state of the 30SIC, which can be shifted to a docking-favorable state by cognate codon-anticodon pairing in the P site or by streptomycin (102). Our finding that cognate start codon recognition reduces the IF1-dependent enhancement of A1408 prompted us to investigate the effects of streptomycin on the structure of h44. We probed 30SIC(AUG) and 30SIC(AUC) with DMS in the presence and absence of streptomycin. Because streptomycin decreases the affinity of IF1 by ~two-fold (i.e., 2-fold) (102), we used a 20-fold excess of IF1 (5 µM) in this experiment to ensure high occupancy of its binding site. In the presence of streptomycin, A1408 reactivity was
substantially reduced in 30SIC(AUG) and 30SIC(AUC), effects strikingly similar to those of A1413C (Figure 4.3C, lanes 7 and 8). Nucleotides forming direct contacts with the antibiotic (e.g., 913-915) were also protected, as were A908 and A909, which lie near the A1413-G1487 region of h44. Mutation A1413C similarly caused protection of A908 and A909 (data not shown). Together these data support to the idea that a conformational change in the 1408 region of h44 helps control 50S docking in response to start codon recognition.
Figure 4.3. Effects of canonical base pairing in the P site on the chemical reactivity of A1408.

(A-C) Primer extension analysis of 16S rRNA from unmodified 30S subunits (K) or DMS-modified 30S subunits (30S) in the absence or presence of various ligands (IFs, initiation factors; AUG, mRNA with AUG start codon; AUC, mRNA with AUC start codon; tRNA, fMet-tRNA) as indicated. Concentrations were as follows (unless otherwise stated): 30S, 0.25 μM; tRNA, 0.5 μM; mRNA, 1.0 μM; IFs, 0.5 μM each).

(A) Analysis of the 1408 and 908 regions in wild type 30S subunits with various ligands.

(B) Analysis of the 1408 region in wild type (WT) and A143C mutant 30S subunits.

(C) Analysis of the 1408 region in the presence of streptomycin (Str).
Figure 4.4. Effects of canonical base paring in the P site on the chemical reactivity of G530 and A532.

(A-C) Primer extension analysis of 16S rRNA from unmodified 30S subunits (K) or chemical modified 30S subunits (30S) (A, B, kethoxal (KE); C, DMS) in the absence or presence of various ligands. Animations are similar as in Figure 4.3.

(A-C) Analysis of the 530 region in wild type 30S subunits with various ligands.

4.3.6. Protections attributed to P-site tRNA depend on AUG

Several of the nucleotides that showed distinct reactivity in 30SIC(AUG) versus 30SIC(AUC) lie in or near the P site (Figure 4.1). One of these is G1338, which forms a Type II A-minor interaction with nucleotide 41 of the anticodon stem (12). Protection of G1338 by fMet-tRNA was evident in 30SIC(AUG), but no tRNA-dependent protection was seen in the case of 30SIC(AUC) (Table 4.2, Figure 4.5A top).

Another differentially modified nucleotide was G926. Early footprinting studies suggested that P-site tRNA directly protects G926 from KE modification. However, more recent structural studies have revealed that N1 and N2 of G926 donate H-bonds
to the phosphate group of the first nucleotide of the P codon. In our experiments, tRNA-independent protection of G926 was observed in the presence of both mRNA and IF3, but was not seen when either mRNA or IF3 was omitted (Figure 4.6). We infer that mRNA directly protects G926, and IF3 stabilizes the P codon in some way. This interpretation is consistent with evidence that IF3 can reposition mRNA from a “standby” site to a “decoding” site without affecting its overall binding affinity (24,82).

In 30SIC(AUG), G926 was found to be further protected compared to 30S·IFs·mRNA complexes, presumably due to codon-anticodon pairing in the P site. In the case of 30SIC(AUC), no further protection was detected (Table 4.2, Figure 4.5A bottom).

Another nucleotide more protected in 30SIC(AUG) than in 30SIC(AUC) was A794. This protection depended completely on fMet-tRNA, consistent with its previous assignment as a P-tRNA protection. However, this protection is probably indirect, since A794 lies > 10 Å from P-site tRNA. In 30SIC(AUG), A794 was strongly protected from DMS. In 30SIC(AUC), protection was still observed but lessened substantially (Table 4.2, Figure 4.5B).

The reactivity of G532 was also found to differ depending on the start codon. Earlier footprinting studies showed protection of G532 by P-site tRNA (103). This probably corresponds to a tRNA-induced conformational change, since G532 lies away from the P site, on the distal side of the A site. We saw protection of G532 in the 30S·IFs complex, which was lost when mRNA was added. Protection was seen again in the context of the 30SIC, but only in the presence of the cognate AUG start codon (Table 4.2, Figure 4.4C).
Figure 4.5. P-site nucleotides (G1338, G926 and A794) and anticodon of the initiator tRNA are efficiently protected only in 30SIC(AUG).

(A-C) primer extension analysis of 16S rRNA from unmodified 30S subunits (K) or chemical modified 30S subunits (30S) (A, KE; B, DMS; C, hydroxal radicals).

(D) primer extension analysis of fMet-tRNA from unmodified fMet-tRNA (K) or chemical modified fMet-tRNA (tRNA) in the absence or presence of various ligands. Animations are similar as in Figure 4.3.

(A) Analysis of the 926 and 1338 region in wild type 30S subunits with various ligands.

(B) Analysis of the 794 region in wild type 30S subunits with various ligands.

(C) Analysis of the 369 region in wild type 30S subunits with various ligands.

(D) Analysis of the anticodon region (CAU) in fMet-tRNA with various ligands.
Figure 4.6. Protection of G926 depends on both IF3 and mRNA.
Primer extension analysis of 16S rRNA from unmodified 30S subunits (K) or KE-modified 30S subunits (30S) in the absence or presence of various ligands. Animations are similar as in Figure 4.3.
Analysis of the 926 region in wild type 30S subunits with various ligands.

4.3.7. Protection of G369 from hydroxyl radicals is increased in the presence of AUG

Earlier work showed that IF2 leaves no DMS or KE footprint on the 30S subunit (106). In order to compare 30SIC(AUG) and 30SIC(AUC) with respect to IF2 binding, we probed the complexes with hydroxyl radicals. Structural studies predict that IF2 interacts with the h5 region of the 30S shoulder (3,110), thus we focused on this vicinity of the 16S rRNA. IF2-dependent protections of several nucleotides in h5 and h15 were observed the 30S-IFs complex. These nucleotides, such as G369, were further protected in 30SIC(AUG) but not in 30SIC(AUC) (Figure 4.5C). These data are in line with the conclusion that the P site is largely unoccupied in the case of 30SIC(AUC). As cooperativity stems from pairwise interactions between fMet-tRNA
and IF2, IF2 and 30S, and 30S and fMet-tRNA, weakening of the latter interaction would be predicted to result in a corresponding decrease in the cooperative binding of IF2.

4.3.8. U701 is less protected from NIMA in the presence of AUG

One target of NIMA, U701, was more reactive in 30SIC(AUG) than in 30SIC(AUC) (Figure 4.7 top). Previous work has shown that IF3 protects U701 from CMCT modification and from hydroxyl radical cleavage, suggesting that U701 contributes to the IF3-binding site (109). This raised the possibility that the enhanced reactivity at U701 was due to loss of IF3 in the case of 30SIC(AUG). However, the adjacent G700 was similarly protected from KE in both complexes (Figure 4.7 bottom), which argued against a difference in the level of bound IF3. To shed light on these seemingly contradictory observations, we monitored NIMA-dependent reactivity in complexes reactions containing higher concentrations of IF3 (10 µM) (data not shown). Under these conditions, complete protection of U701 was observed in both complexes, suggesting that the difference observed initially reflects a difference in IF3 binding affinity, as opposed to a qualitative difference between the complexes (e.g., conformational changes in the platform). This interpretation is consistent with earlier evidence that cognate codon-anticodon pairing destabilizes IF3 in the 30SIC (102). The reason differences in reactivity were observed with NIMA but not KE remains unclear, but we suspect that it has to do with the probes themselves. NIMA is a good general electrophile predicted to modify not only rRNA but also IF3. Hence NIMA may have
decreased the concentration of active IF3 during the course of the reaction and amplified differences the protection pattern.

Figure 4.7. Canonical base paring in the P site de-protects U701.
Primer extension analysis of 16S rRNA from unmodified 30S subunits (K) or chemical modified 30S subunits (30S) (top, NMIA; bottom, KE) in the absence or presence of various ligands. Animations are similar as in Figure 4.3. Analysis of the 700 region in wild type 30S subunits with various ligands.

4.3.9. Protection of the anticodon of fMet-tRNA is only evident in 30S initiation complexes containing AUG

The fact that G1338 and G926 were similarly reactive in 30SIC(AUC) and 30S·IFs·mRNA complexes suggested that the P site of 30SIC(AUC) was largely
unoccupied. To investigate the extent of codon-anticodon pairing under similar conditions, we assembled 30SIC(AUG) and 30SIC(AUC) with a stoichiometric (0.25 µM) concentration of fMet-tRNA and looked at exposure of the anticodon to DMS and NMIA (Figure 4.5D). Clear protection of the anticodon from each probe was observed in 30SIC(AUG). In contrast, the anticodon nucleotides were essentially as reactive in 30SIC(AUC) as in the free fMet-tRNA control. These data suggest that codon and anticodon are mostly unpaired in the presence of the near-cognate AUC.

4.3.10. Comparing the extent of fMet-tRNA binding in both complexes

The footprinting data described above suggest that the P site is largely unoccupied in 30SIC(AUC). To compare the level of fMet-tRNA binding in 30SIC(AUG) and 30SIC(AUC), we assembled the two complexes in the presence of fMet[^35]S-tRNA and used microfiltration devices (microcon YM100; MWCO 100 kDa) to separate free and bound tRNA. Under conditions essentially identical to those used in the probing experiments, substantial levels of bound fMet-tRNA were detected in both cognate and near-cognate complexes (70% and 50%, respectively) (Figure 4.8A). One caveat here is the IF2-fMet-tRNA binary complex is predicted to be slightly larger (~120 kDa) than the pore size of the filtration devices. Omission of the 30S subunit in analogous binding experiments showed that ~20% of the total fMet[^35]S-tRNA was bound, presumably by IF2 (data not shown). Hence, the difference in the extent of fMet-tRNA binding in 30SIC(AUG) versus 30SIC(AUC) could be as much as two- to three-fold.
Figure 4.8. The initiator tRNA associates with the 30SIC(AUC) to a similar extent, but in a less stable manner compared to that in 30SIC(AUG).

(A) Fraction bound of fMet-tRNA in 30SIC(AUG) and 30SIC(AUC) complexes. [35S]fMet-tRNA^{fMet} (0.5 μM) was incubated with the 30S subunit (0.25 μM) in the presence of mRNA (1 μM, either AUG or AUC), initiation factors (0.375 μM) and GTP (0.1 mM). The reactions were passed through microncon YM100 filter units to separate the free and bound [35S]fMet-tRNA^{fMet}. The signal from the retained and flow through portions was measured in scintillation counter (Unit: cpm). The fraction bound of fMet-tRNA^{fMet} was calculated using equation \( \frac{Cpm_{\text{retained}}}{Cpm_{\text{retained}} + Cpm_{\text{flow through}}} \). Mean ± SEM from three independent experiments was shown.

(B) The fraction of 30S·tRNA·mRNA complexes formed was measured by toeprinting. The P site is filled by adding fMet-tRNA^{fMet} (1.5 μM) to 30S (1 μM) with mRNA (0.1 μM, either AUG or AUC) in the presence of various combinations of initiation factors (IF1, IF2, 1.5 μM; IF3, as indicated). The 30S subunit assembly on mRNA (e.g., formation of 30S·mRNA·fMet-tRNA complex) will stop reverse transcriptase, resulting in a toeprint, intensity of which indicates the level of 30S·mRNA·fMet-tRNA formation. When mRNA is free, reverse transcriptase extends to the 3’ end of mRNA, resulting in a runoff band, intensity of which indicates the level of free mRNA.
4.3.11. Stability of the P codon in 30S initiation complexes containing AUG versus AUC

Next, we used toeprinting to compare the stability of the P codon in various complexes. Under the conditions employed, a relative equilibrium association constant, $K_{rel}$, could be approximated based on the fraction of 30S·mRNA·fMet-tRNA ternary complex (Table 4.3) (See Materials and Methods). With the AUG start codon, the 30S·mRNA·fMet-tRNA ternary complex gave a strong toeprint (fraction of total signal = 0.98 ± 0.005), indicating a stable ternary complex ($K_{rel}$ = 50 ± 10). The stability was slightly enhanced by IF1 and IF2 ($K_{rel}$ = 110 ± 3) but reduced in the presence of all three factors ($K_{rel}$ = 1.7 ± 0.2). Higher concentrations IF3 reduced the toeprint signal further (Table 4.3, Figure 4.8B). These data agree well with earlier evidence that IF2 stabilizes and IF3 destabilizes fMet-tRNA in the 30SIC (7,86). In the presence of AUC and absence of initiation factors, the toeprint corresponding to the 30S·mRNA·fMet-tRNA complex was considerably weaker ($K_{rel}$ = 0.48 ± 0.03) than with AUG ($K_{rel}$ = 50 ± 10). The selectivity ($K_{rel}(\text{AUG})$ versus $K_{rel}(\text{AUC})$) is large, ~100 (Table 4.3). The presence of IF1 and IF2 increased the stability of 30S·AUC·fMet-tRNA more dramatically ($K_{rel}$ = 29 ± 2) in relation to that with AUG. Under this condition, the selectivity goes down to ~4. The further addition of IF3 reduced the toeprint signal of 30S·AUC·fMet-tRNA substantially (fraction = ~0.017, $K_{rel}$ = 0.017 ± 0.0007). Under this condition, the selectivity goes up again to ~100, a level similar in the case of no initiation factors. The weak toeprint in 30S·AUC·fMet-tRNA complex in the presence
of all initiation factors is consistent with the conclusion drawn from the chemical probing data that codon and anticodon are mainly unpaired in 30SIC(AUC).

These data provide evidence that: AUG is selected over AUC even in the absence of initiation factors; the presence of IF1 and IF2 enhances the stability of both 30S·AUG·fMet-tRNA and 30S·AUC·fMet-tRNA complex and largely diminishes the difference of AUG and AUC; further addition of IF3 destabilize both complexes in a way that AUG is favored.

Table 4.3. Summary of stability of 30S complexes containing AUG versus AUC

<table>
<thead>
<tr>
<th></th>
<th>AUG</th>
<th>AUC</th>
<th>Selectivity</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Fraction</td>
<td>$K_{rel}$</td>
<td>Fraction</td>
</tr>
<tr>
<td>no IFs</td>
<td>0.98 ± 0.005</td>
<td>50</td>
<td>0.32 ± 0.01</td>
</tr>
<tr>
<td>IF1,2</td>
<td>0.99 ± 0.0002</td>
<td>110</td>
<td>0.97 ± 0.001</td>
</tr>
<tr>
<td>IF1,2 and IF3(1.5)$^a$</td>
<td>0.63 ± 0.02</td>
<td>1.7</td>
<td>0.017 ± 0.0007</td>
</tr>
<tr>
<td>IF1,2 and IF3(2.5)$^a$</td>
<td>0.58 ± 0.04</td>
<td>1.4</td>
<td>0.016 ± 0.0005</td>
</tr>
<tr>
<td>IF1,2 and IF3(4.0)$^a$</td>
<td>0.57 ± 0.003</td>
<td>1.3</td>
<td>0.014 ± 0.0002</td>
</tr>
</tbody>
</table>

$^a$ The number in the parenthesis indicates the concentration of IF3 (µM);
The experiments were performed in a condition that the concentrations of 30S (1 µM) and fMet-tRNA (1.5 µM) are in 10-fold excess over mRNA (0.1 µM), thus [30S] and [tRNA] can be considered constant;
Fraction is defined as $[30S\cdot tRNA\cdot mRNA]/([mRNA]+[30S\cdot tRNA\cdot mRNA]);$
$K_{rel}$ is the estimate of $K_a$, calculated from fraction as described in the methods;
Selectivity is defined as $K_{rel}(AUG)/K_{rel}(AUC)$
Values represent the mean ± SEM of three independent experiments.
4.4. Discussion

4.4.1. A conformational switch in helix 44 helps control 50S docking

Among the most critical determinants of accurate initiation of translation is the start codon itself. Single base substitutions of the start codon severely reduce the rate of translation \textit{in vivo} and the rate of stable 70SIC formation \textit{in vitro} (55,102,133,164). IF3 and IF1 are known to negatively regulate the initiation process to enhance start codon selection, but the molecular mechanisms through which these factors act and the role of 16S rRNA in the process remains unclear. Here, we provide evidence that initiation is controlled in part by a conformational switch in the 1408 region of h44. First, comparison of 30SIC(AUG) and 30SIC(AUC) by chemical probing shows that an IF1-dependent enhancement at A1408 is reduced in the presence of AUG. This change in reactivity is due to a conformation change rather than loss of IF1 because other portions of the IF1 footprint are unchanged and high concentrations of IF1 fail to diminish the reactivity difference seen at A1408. Second, mutations in h44 (e.g., A1413C, A1410G) increase spurious initiation \textit{in vivo}, stimulate 50S docking \textit{in vitro}, and cause reduced reactivity at A1408. Third, streptomycin, which has been shown stimulate 50S docking by reversing the inhibitory effects of IF1 (102), also causes reduced reactivity at A1408. Collectively, these data support a model in which IF1 alters the local structure of h44 in a way that makes 50S docking unfavorable, and cognate codon-anticodon pairing in the P site partially reverses the structural changes induced by IF1 to promote 50S docking. This model builds on one proposed by
Rodnina and coworkers, based on their studies of IF1. Our data indicate that the key conformational change hypothesized by Milon et al. occurs in the A1408 region of h44.

The P codon forms contacts to C1400, C1402, C1403, and U1498 of h44 but lies ~20 Å from A1408 (12) (141) (80). Hence, cognate codon-anticodon pairing in the P site must indirectly influence the structure of h44 remains unclear. Previous studies indicate that P-site tRNA causes structural changes in h44 in the absence of initiation factors, suggesting that the signaling mechanism is inherent to the 30S subunit and tuned by IF1 (104).

A1408 contributes to intersubunit bridge B2a (128,184), hence conformational changes in the A1408 region may directly regulate formation of B2a and thereby influence 50S docking. Cryo-EM studies predict that the A1408 region should be exposed in the 30SIC and that H69 of the 50S subunit (which forms the other half of B2a) should have relatively unhindered access to the A1408 region (152). These observations raise the possibility that B2a is one of the first bridges to form during 70SIC formation, as appears to be the case during association of the subunits in the absence of other ligands (66).

Earlier chemical probing experiments predicted conformational signaling between the P site and h44. Moazed and Noller (1987) identified six nucleotides in the 30S subunit (A790, G791, A909, A1394, A1413, and A1487) protected from chemical probes by multiple ligands (e.g., P-tRNA, various antibiotics, 50S subunit) (104). These nucleotides were termed Class III sites and the protections were attributed to ligand-induced conformational changes. Interestingly, three of these nucleotides (A909,
A1413, and A1487) localize in or near h44, are protected by either tRNA or streptomycin, and become hyper-reactive in the presence of IF1. These data raise the possibility that the conformational change induced by start codon recognition encompasses the 1413-1487 portion of h44 as well as the 1408 region. Consistent with this idea, mutations in h44 affecting initiation fidelity cluster right in the vicinity of 1413-1487 (Figure 4.1). As this region lies between nucleotides contributing to bridge B2a and B3, formation of both bridges may be regulated by codon-anticodon pairing in the P site. In the presence of all components of the 30SIC, A1413 appeared slightly more protected in the presence of the cognate AUG start codon. While this difference was not verified statistically at the 95% confidence level, each independent experiment showed the same trend in which A1413 was less reactive in 30SIC(AUG) compared to 30SIC(AUC), as would be predicted from the earlier probing studies. Hence, we suspect that codon-anticodon pairing in the P site perturbs h44 in a similar way in the absence and presence of initiation factors but that the resulting decrease in A1413 reactivity is smaller in the latter case.

4.4.2. The P site is largely unoccupied in the 30SIC(AUC)

A number of observations indicate that occupancy of the P site in the 30SIC depends on cognate codon-anticodon pairing. First, chemical protections of 16S rRNA attributed to P-site tRNA were seen with AUG but were absent or substantially diminished with AUC. Second, protections of the anticodon bases of fMet-tRNA were obvious only in the presence of AUG, suggesting that the anticodon was primarily
unpaired in the AUC case. Third, evidence for cooperative binding between fMet-tRNA and IF2 was observed only in the presence of AUG. A simple explanation for this result is that fMet-tRNA provides an additional anchor for IF2 on the 30S subunit only when the anticodon is bound to the P site. Finally, toeprinting analysis of 30SICs suggested that the stability of tRNA in the P site decreases by about 100-fold when the start codon is changed from AUG to AUC, a difference that depends on IF3. Collectively these data suggest that 30SIC formation is highly dependent on a canonical start codon, with the equilibrium shifted strongly leftward (toward disassembly) in the presence of AUC. This may be responsible at least in part for the very slow rate of 70SIC formation observed in the presence of AUC (Figure 4.2).

While the dramatic effect of the start codon mutation on 30SIC formation is consistent with earlier work (64,84) and recent smFRET studies (R. Gonzalez, personal communication), it is at variance with results from a study by Cooperman and colleagues (55). In their experimental system, no differences in the rate or extent of fMet-tRNA binding were seen during 30SIC formation when AUG and AUU start codons were compared, although large effects of the start codon mutation were seen in later steps of the initiation process. In monitoring subunit joining by LS, they observed a fast codon-independent phase (LS1) followed by a slower codon-dependent phase (LS2). LS1 was attributed to accumulation of a loose 70S intermediate, while LS2 was attributed to stable 70SIC formation. We and others (7,8,102,120) fail to see a rapid codon-independent phase analogous to LS1. Instead, the LS increase we see is completely codon-dependent and occurs at a rate (4 s⁻¹) similar to Cooperman’s LS2
(6.5 s\(^{-1}\)), suggesting that the loose 70S intermediate is too unstable under our conditions to detect. Given the importance of fMet-tRNA in 50S docking (120), we speculate that the differences in the LS data are related to the differences in fMet-tRNA binding in the two experimental systems.

Using a micron-based assay, we estimated that the percentage of fMet-tRNA bound in 30SIC(AUC) is substantial (30-50%), not much different than in 30SIC(AUG) (50-70%). Yet the chemical probing and toeprinting data indicate that the extent of tRNA binding to the P site in 30SIC(AUC) is very low. One explanation for these data is that, in the presence of AUC, fMet-tRNA is bound in a different way, without the anticodon engaged in the P site. Current models of 30SIC formation depict random-order binding of mRNA and fMet-tRNA, followed by start codon recognition (57). Implicit in these models is that each ligand (fMet-tRNA and mRNA) interacts initially in a kinetically labile manner, and then both ligands are mutually stabilized by codon-anticodon pairing in the P site. Indeed, the 30SIC(AUC) complex characterized here may resemble an intermediate in the kinetic pathway akin to the “standby” state (37,162).

**4.4.3. IF3 kinetically counteracts the effects of IF2 such that the intrinsic specificity of the programmed P site is exploited.**

Although the importance of IF3 in start codon selection has long been established, the mechanism by which IF3 acts to enhance fidelity has been unclear. It has been proposed that the difference in stability between a matching and nonmatching
codon-anticodon helix is not sufficient to explain the observed level of start codon discriminations, and IF3 destabilizes noncanonical complexes to a greater degree than canonical complexes and thereby enhances start codon selection (58,62,84). IF3 shifts equilibrium of 30SIC formation in both complexes (matching versus nonmatching) in a way that favors matching one. According to this view, IF3 probably ‘‘senses’’ or ‘‘inspects’’ a matched codon-anticodon helix instead of unique feature of the start codon (99). An alternative hypothesis has been proposed by Ehrenberg and coworkers (7). They observed that IF3 uniformly destabilizes tRNA in the P site regardless of the nature of the bound tRNA (initiator and elongator). They suggested that by increasing $k_{off}$ of fMet-tRNA and slowing the rate of 50S docking, IF3 would allow the energetic cost of a mismatch to be fully exploited and a high level of selectivity to be achieved. Our results confirm the latter hypothesis. First, we show that the programmed P site exhibits a high degree of selectivity ($K_{rel}$ (AUG) / $K_{rel}$ (AUC = ~100) in the absence of initiation factors. This is sufficient to account for the observed fidelity of initiation in vivo. Second, we show that the presence of all initiation factors did not obviously shift the balance. The intrinsic selectivity for AUG may reflect the property of the primordial 30S subunit. In the modern world, translation initiation machinery has to balance accuracy and speed. To enhance speed, IF2 greatly stimulates the forward reaction of 70SIC formation (e.g., increases in the association rate of fMet-tRNA and the docking of 50S). When 50S docking is rapid, the 30SIC intermediate exists with a short half-life, providing little time for dissociation of a mispaired fMet-tRNA. IF3 slows down 50S docking on the one hand, and increases dissociation rate of P-site
tRNA on the other, giving mispaired fMet-tRNA opportunity to dissociate and hence allowing the intrinsic selectivity of the P site to be efficiently used. Notably, in our model, IF3 does not have to increase the dissociation rate differentially since the intrinsic differences are large. This model explains how IF3 enhances start codon selection in the presence of IF2, allowing initiation to occur both rapidly and accurately.
Chapter 5

Conclusions and future directions

The translation machinery has evolved to ensure that genetic information is faithful and rapidly decoded. Precise selection of the translational start site is critical, for this establishes the correct reading frame. The fidelity of initiation depends on the contribution of initiation factors, initiator tRNA, and the ribosome to specifically recognize key features of the TIR of mRNA. In this study, we focused on the role of the 16S rRNA in start codon selection and gained insight into the molecular mechanisms that govern initiation.

During translation initiation, the fMet-tRNA enters the 30S P site, where 16S rRNA makes extensive contact with P-site tRNA and mRNA (12,25,80,141). In Chapter 2 and 3, we proposed that these interactions are tuned to balance efficiency and accuracy. In these studies, we screened a set of P-site mutations for their effects on start codon selection. G1338A and A790G decrease the fidelity of start codon selection, while other P-site mutations increase it. We provided evidence that G1338A enhances interactions of fMet-tRNA and the 30S subunit, which can compensate for mismatches in the codon-anticodon helix, while other P-site mutations (e.g., G1338U) weaken these interactions, making initiation more dependent on the correct codon-anticodon
pair, thus increasing the stringency of start codon selection. Increased fidelity in start codon selection gained by destabilization of fMet-tRNA could be mechanistically similar to the role of IF3, as described in the introduction. We also observed that further enhancing stringency by G1338U is at the cost of reduced level of translation. Like G1338A, another P-site mutation A790G also decreases the fidelity of start codon selection. A790G does not stabilize fMet-tRNA; instead, it destabilizes IF3, a factor known to be crucial for start codon selection.

In Chapter 3, we randomly mutagenized the 16S rRNA gene and isolated a number of novel mutations that increase the translation from non-canonical start codon AUC. These mutations cluster to distinct regions of 16S rRNA (h44, 790 loop, and neck). The mutations in h44 map near the IF1 binding site, implicating IF1 in the mechanisms of start codon selection. While certain h44 mutations (A1408 and A1410G) modestly affected IF1 binding, others further ‘down’ the helix did not. These latter mutations converted non-canonical or wobble base pairs to Watson-Crick base pairs, suggesting that the conformation of h44 plays a role in initiation fidelity. These data provided further evidence that decreasing IF3 binding by 16S rRNA mutations stimulates spurious initiation. We suspect the mutations in the neck may influence the dynamics of the head movements, a hypothesis that remains to be investigated.

In Chapter 4, we investigated the structural changes of 30SIC upon recognition of canonical start codon by chemical probing. From ~110 target nucleotides, we identified 6 positions showing differences between 30SIC(AUG) and 30SIC (AUC). One is A1408, which is located in h44, near IF1 binding site. Early chemical studies
showed that IF1 binding causes hyper reactivity of A1408 (34,106). We found canonical codon-anticodon pairing in the P site reduced the hyper reactivity of A1408, while non-canonical codon-anticodon pairing did not. In addition, h44 mutations identified previously or streptomycin similarly reduces the hyper reactivity of A1408 regardless of the start codon. Furthermore, reduction of A1408 hyper reactivity is associated with increase in 50S docking rate to 30SIC. Based on these observations, we proposed the docking is controlled in part by a conformational switch in the 1408 region--IF1 holds the A1408 region in a docking unfavorable conformation, which is reversed upon start codon recognition. Canonical start codon pairing in the P site also deprotects U701, an IF3 binding site, suggesting that in 30SIC(AUG), IF3 is destabilized.

In Chapter 4, we also found that in 30SIC, several P-site nucleotides are efficiently protected by fMet-tRNA only in the presence of canonical start codon. These results indicated that the P site is largely unoccupied in the presence of non-canonical start codon AUC. Our fMet-binding experiments using microcon or toeprinting techniques suggested that fMet-tRNA associates to the 30S subunit in a kinetically labile manner when AUC is the start codon. Our toeprinting data also showed that even in the absence of initiation factors, the stability of 30S·tRNA·AUG ternary complex is 100-fold greater than 30S·tRNA·AUC complex, indicating an intrinsic high selectivity for AUG. IF2 and IF1 enhance the stability of both complexes, with more dramatic effects on 30S·tRNA·AUC. The additional presence of IF3 destabilizes both complexes such that the binding selectivity is restored to ~100. These
data suggest that IF2 greatly promotes the forward reaction, at the cost of fidelity. IF3 shifts the equilibrium back, presumably by increasing the \( k_{off} \) thereby restoring the selectivity while maintaining the fast rate of initiation.

**Future directions**

We have identified that mutations clustering in distinct regions of 16S rRNA cause spurious initiation. We have characterized three regions (790 loop, h44 and P site). It will be of interest to investigate the remaining regions, although they have subtle effects. One of them is in the neck region. Mutations in this region are suspected to affect the dynamics of head domain, which may influence the initiator tRNA binding and/or 50S docking since the head domain adopts a different position depending on the ligand bound (67,140). Intriguingly, it is proposed that IF3 affects the conformation of 30S by altering the orientation the head domain, which may be relevant for its fidelity function. One hypothesis is that this movement opens the P site gate A790 and G1338, thus paving the way for tRNA dissociation (42). Another interesting mutation lies in the central pseudoknot, which connects the three major domains of 16S rRNA. Like the neck mutation, it may affect the domain dynamics. One potential issue in studying existing mutations in these two regions is that their effects are subtle *in vivo*. However, we can use site-directed mutagenesis to target these two regions. Hopefully, mutations with stronger phenotypes will be identified. Alternatively, we can combine the mutations (e.g., introducing three available neck mutations to the 16S rRNA). We can study the effects of the mutations *in vitro* using the techniques developed in the lab.
Although we have evidence that neck mutation G1094A does not intrinsically stabilize tRNA$^{\text{Met}}$, it would be interesting to repeat the experiments in the presence of IF3. Also, we can investigate whether the mutations stimulate 70SIC formation using stopped-flow spectroscopy. Once a strong phenotype is observed, it will be interesting to monitor the dynamics of the 30SIC. One approach is to use single molecule Fluorescence Resonance Energy Transfer (smFRET), as in (30). The fluorescence labels can be introduced into two small ribosomal proteins (e.g., S7 and S11, S3 and S4) and reconstituted into 30S subunits (67). The relative position of these two proteins should be a good indicator of domain movements.

IF3, as an anti-association factor, functions in preventing premature 50S docking to non-canonical 30SICs. However, the molecular basis is unclear. One hypothesis is that IF3 prevents subunit association by sterically clashing with H69 of the 50S subunit (35). Consistent with this idea, a biochemical study demonstrated high concentration of IF3 inhibits 50S docking, implying mutually exclusive interactions of the 50S subunit on IF3 with the 30S subunit (7). Interestingly, we observed that the 50S subunit with a deletion of H69 associates rapidly with the 30SIC even in the presence of non-canonical start codon (data not shown). One explanation is that deletion of H69 removes the clash, thus allowing subunit association regardless of the presence of IF3. It will be of interest to test this idea. For example, we can monitor binding of IF3 to the 70S ribosome with H69 deletion by sucrose gradient assay used in Chapter 2 and 3. We predict IF3 will greatly associate with this type of ribosome compared to control.
In the co-crystal structure, mRNA directly contacts G926, which explains mRNA-dependent protection of G926 (80,141). Our chemical probing showed that in the absence of fMet-tRNA, this mRNA-dependent protection of G926 occurs only in the presence of IF3. We suspect this may be due to repositioning of mRNA by IF3, as suggested in previous reports (82). The nature of this putative repositioning remains obscure. One possibility is that the region of mRNA downstream of SD enters the mRNA channel with the help of IF3 (73), thereby protecting G926. When fMet-tRNA is added in the case of AUG, G926 is further protected. Further protection could come from additional repositioning of mRNA by the tRNA. This raises the possibility that in the presence of IFs, the mRNA is in a suboptimal position for pairing with fMet-tRNA. The start codon of mRNA could stay in the “stand-by” site or moves in equilibrium between “stand-by” and “decoding” site. This suboptimal positioning could make initiation more dependent on a canonical start codon. When fMet-tRNA comes in, it pairs with AUG and helps position it in the P site. This will induce a 5’-3’ movement of mRNA, based on previous prediction (75,82). It could be interesting to monitor the movements of mRNA in the presence of absence of fMet-tRNA by smFRET. In addition, this mRNA movement is proposed to couple the SD-ASD movement (75). Based on the predicted IF3 binding site (35), the SD-ASD movement could affect the stability of IF3 on the 30S subunit. This provides one explanation why canonical start codon destabilizes IF3, as observed in our chemical probing study (U701) and others (7,102). In this case, we might see that with mRNA lacking SD (e.g., (AUG)_n and (AUC)_n), the dissociation of IF3 is insensitive to the start codon.
LIST OF REFERENCE


