Structural and Biochemical Studies of Antibiotic Resistance and Ribosomal Frameshifting

YANG CHEN
Abstract


Protein synthesis, translation, performed by the ribosome, is a fundamental process of life and one of the main targets of antibacterial drugs. This thesis provides structural and biochemical understanding of three aspects of bacterial translation.

Elongation factor G (EF-G) is the target for the antibiotic fusidic acid (FA). FA binds to EF-G only on the ribosome after GTP hydrolysis and prevents EF-G dissociation from the ribosome. Point mutations in EF-G can lead to FA resistance but are often accompanied by a fitness cost in terms of slower growth of the bacteria. Secondary mutations can compensate for this fitness cost while resistance is maintained. Here we present the crystal structure of the clinical FA drug target, Staphylococcus aureus EF-G, together with the mapping and analysis of all known FA-resistance mutations in EF-G. We also present crystal structures of the FA-resistant mutant F88L, the FA-hypersensitive mutant M16I and the FA-resistant but fitness-compensated double mutant F88L/M16I. Analysis of mutant structures together with biochemical data allowed us to propose that fitness loss and compensation are caused by effects on the conformational dynamics of EF-G on the ribosome.

Aminoglycosides are another group of antibiotics that target the decoding region of the 30S ribosomal subunit. Resistance to aminoglycosides can be acquired by inactivation of the drugs via enzymatic modification. Here, we present the first crystal structure an aminoglycoside 3’ adenyltransferase, AadA from Salmonella enterica. AadA displays two domains and unlike related structures most likely functions as a monomer.

Frameshifts are deviations the standard three-base reading frame of translation. -1 frameshifting can be caused by normal tRNASer at GCA alanine codons and tRNAThr at CCA/CCG proline codons. This process has been proposed to involve doublet decoding using non-standard codon-anticodon interactions. In our study, we showed by equilibrium binding that these tRNAs bind with low micromolar Kd to the frameshift codons. Our results support the doublet-decoding model and show that non-standard anticodon loop structures need to be adopted for the frameshifts to happen.

These findings provide new insights in antibiotic resistance and reading-frame maintenance and will contribute to a better understanding of the translation elongation process.

Keywords: protein synthesis, elongation, elongation factor G, fusidic acid, antibiotic resistance, aminoglycoside adenyltransferase, ribosomal frameshifting

Yang Chen, Uppsala University, Department of Cell and Molecular Biology, Box 596, SE-751 24 Uppsala, Sweden.

© Yang Chen 2013

ISSN 1651-6214
urn:nbn:se:uu:diva-205131 (http://urn.kb.se/resolve?urn=nbn:se:uu:diva-205131)
献给我最亲爱的父亲母亲
List of Papers

This thesis is based on the following papers, which are referred to in the text by their Roman numerals.


III Yang Chen, Joakim Näs­vall, Dan I. Andersson and Maria Selmer. (2013) Crystal structure of AadA at 2.5 Å resolution- an aminoglycoside 3” adenylyltransferase. (manuscript)

IV Yang Chen, Suparna Sanyal and Maria Selmer. (2013) tRNA\textsuperscript{Ser} and tRNA\textsuperscript{Thr} induce -1 frameshifting using alternative anticodon-loop structures. (manuscript)

Reprints were made with permission from the respective publishers.
Contribution to papers

Paper I. I collected data at the synchrotron, solved the wild-type structure, refined and analyzed the structure. I wrote a first draft of the paper and actively participated in later writing.

Paper II. I collected data for all the mutants, solved the mutant structures and analyzed the structures. I wrote a first draft of the structural part of the paper and actively participated in later writing.

Paper III. I purified and crystallized the AadA protein, collected both anomalous and native data, solved the structure and analyzed the structure. I took major responsibility in writing the manuscript.

Paper IV. I prepared all the components for the filter-binding assay, performed the assay, and analyzed the data. I took major responsibility in writing the manuscript.

Additional publications

CONTENTS

Introduction ................................................................................................... 11

Translation in bacteria .................................................................................. 12
  Initiation ................................................................................................... 14
  The elongation cycle ............................................................................... 14
    Decoding .......................................................................................... 15
    Peptide bond formation ............................................................... 17
    Translocation .................................................................................. 17
  Termination .......................................................................................... 21
  Ribosome recycling............................................................................. 22

Translation inhibition by antibiotics ............................................................. 23
  Fusidic acid and its action ................................................................... 23
  Aminoglycosides and their mode of action ........................................... 24

Antibiotic resistance ...................................................................................... 27
  Resistance mechanisms to fusidic acid .............................................. 27
    Fitness cost and compensation ....................................................... 28
  Resistance mechanisms to aminoglycosides ...................................... 28
    Aminoglycoside-modifying enzymes (AMEs) ............................. 29

Ribosomal frameshifting ............................................................................. 31

Methodology ................................................................................................ . 34
  Component preparation ..................................................................... 34
  X-ray crystallography ....................................................................... 35
  Filter binding assay .......................................................................... 36

Aims of thesis ............................................................................................... 37

Summary of current research ........................................................................ 38
  Paper I: Mapping of FA resistance mutations on the drug target ....... 38
    The structure .................................................................................. 38
    Conformational space of EF-G ....................................................... 39
  FA-resistance mutations ................................................................. 40
  Mysterious density – capture of free phosphate? ............................ 41
Paper II: A clever network to gain fitness: analysis of EF-G mutants.... 41
   Overall structures of the mutants...................................................... 42
   Phe88 and Switch II region .............................................................. 42
   Met16 and the hydrophobic core.................................................... 43
   Crucial salt bridges...................................................................... 44

Paper III: The first structure of ANT (3”) - AadA ............................ 44
   The structure.................................................................................. 44
   Conserved residues in AadA form a pocket for substrate and cofactor binding.................................................. 45
   AadA most likely functions as a monomer .................................... 46
   Candidate residues for catalysis and substrate binding............... 46

Paper IV: Alternative anticodon-loop structure to induce -1 frameshifting .............................................................. 49

Concluding remarks........................................................................ 52
Future perspectives ......................................................................... 53
Sammanfattning på svenska .............................................................. 54
Acknowledgements.......................................................................... 56
References...................................................................................... 58
Abbreviations

A-site aminoacyl-site
aa-tRNA aminoacyl-tRNA
AAC aminoglycoside acetyltransferase
AME aminoglycoside-modifying enzyme
ANT aminoglycoside nucleotidyltransferase
APH aminoglycoside phosphotransferase
ASL anticodon stem loop
AU asymmetric unit
Cryo-EM cryoelectron microscopy
Da dalton, the standard unit indicating mass. One Dalton is equivalent to 1g/mol
DNA deoxyribonucleic acid
E-site exit-site
EF- ribosomal elongation factor-
FA fusidic acid
fMet N-Formylmethionine
GMPPCP β,γ-methyleneguanosine 5’-triphosphate
GTPase guanosine triphosphate hydrolase
IC initiation complex
IF- ribosomal initiation factor-
KNTase kanamycin nucleotidyltransferase
MD molecular dynamics
MIC minimal inhibitory concentration
mRNA messenger RNA
NCS noncrystallographic symmetry
NMR nuclear magnetic resonance
P-site peptidyl-site
POST post-translocation
ppGpp guanosine tetraphosphate
PRE pre-translocation
PTC peptidyl transferase center
RF- ribosomal release factor-
RNA ribonucleic acid
RRF ribosomal recycling factor
rRNA ribosomal RNA
S Svedberg unit for sedimentation rate
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>SAD</td>
<td>single-wavelength anomalous diffraction</td>
</tr>
<tr>
<td>SCV</td>
<td>small-colony variant</td>
</tr>
<tr>
<td>SD</td>
<td>Shine-Dalgarno</td>
</tr>
<tr>
<td>smFRET</td>
<td>single molecule fluorescence resonance energy transfer</td>
</tr>
<tr>
<td>tRNA</td>
<td>transfer RNA</td>
</tr>
<tr>
<td>WT</td>
<td>wild type</td>
</tr>
</tbody>
</table>
Introduction

Genomes are maintained and expressed with remarkably high fidelity throughout biology. This is partly achieved by the fundamental process of life - protein synthesis, which is implemented in the amazing cellular machines - the ribosomes. They translate the encoded information from the genome provided by messenger ribonucleic acid (mRNA) and link together the amino acids carried by transfer ribonucleic acid (tRNA) to produce polypeptides, which will fold into functional proteins.

The prokaryotic and eukaryotic ribosomes both consist of two subunits, but their size and structure are different, and the translation process also slightly differs between prokaryotes and eukaryotes. These differences open a window for drugs to target the bacterial ribosomes. Indeed, antibiotics, such as aminoglycosides, tetracyclines, cyclic peptides, macrolides and so on, specifically target the bacterial ribosomes and inhibit translation. These ribosome-targeting antibiotics are not only used clinically to cure severe infections but also for probing various states of protein translation and completing the picture of the bacterial translation cycle. Yet, bacteria are clever enough to acquire resistance to antibiotics, and understanding of the mechanisms are of great interest and important for potential therapeutics in the future. Although the ribosomes are sophisticated to do its job with high fidelity, errors can occur, such as missense mutations, frameshifting and so on. Protein translation is a fundamental but complex process. Although each step of the process has been extensively studied over the past few decades, much remains unresolved.

The work in this thesis is mainly focusing on three aspects of the elongation step in the translation cycle using structural and biochemical methods. First, the ribosomal elongation factor G is a drug target for fusidic acid (FA), point mutations on EF-G can lead to FA resistance and often a secondary mutation compensates the fitness cost – typically a reduced growth rate. Second, resistance to aminoglycosides can be acquired by inactivation of the drugs by enzymatic modification, and the aminoglycoside adenyltransferase is one of these aminoglycoside-modifying enzymes. Third, -1 ribosomal frameshifting can be caused by normal tRNAs using an unknown mechanism. These three aspects are independent, but all have something to do with errors occurring at the elongation cycle. Each project will be introduced and insights from our results will be presented in this thesis.
Translation in bacteria

The sequence of three-base codons on mRNA carrying the genetic information directs the synthesis of a protein polypeptide chain. This process is named translation and it takes place on a large and complex molecular machine – the ribosome. The ribosome, found in all living cells, is made from both RNAs and proteins and is therefore a ribonucleoprotein particle. It consists of two subunits in all species. In bacteria, the two subunits are designated 50S and 30S respectively, and together form the 2.5-mega-dalton (MDa) 70S ribosome; in eukaryotes, their counterparts are the 60S and 40S subunits and the 80S ribosome. The 50S subunit is composed of 23S ribosomal RNA (rRNA), 5S rRNA and about 30 ribosomal proteins; the 30S subunit is composed of 16S rRNA and about 20 ribosomal proteins. During translation, the ribosome works as a dynamic machine together with various ribosomal translation factors, many of which are guanosine triphosphate hydrolases (GTPases).

Much progress towards understanding of the detailed mechanism of translation has been made during the past two decades. This has been achieved by the application of various tools such as pre-steady-state and fast kinetics, cryoelectron microscopy (cryoEM), molecular dynamics (MD) simulations, X-ray crystallography, single molecule fluorescence resonance energy transfer (smFRET) and so on.

On the ribosome, there are three binding sites for the tRNAs: 1) aminoacyl-site (A-site), which accommodates the incoming aminoacylated tRNA; 2) peptidyl-site (P-site), which holds the translocated tRNA attached to a nascent polypeptide chain, and 3) exit-site (E-site), to which the deacylated tRNA proceeds after the formation of a peptide bond before leaving the ribosome. The tRNAs occupy the inter-subunit space with their anticodon stem loops (ASLs) protruding to the mRNA codons on the 30S subunit and their amino-acid-carrying ends reaching the peptidyl transferase center (PTC) at the 50S subunit where peptide formation occurs. The tRNA and mRNA binding sites and some basic features of the ribosome are shown in Figure 1. The mRNA is located at the cleft between the “head” and the “body” of the 30S subunit. The tRNAs have codon-anticodon interactions with the mRNA, and the 3’-CCA end of the A- and P-site tRNA are in the PTC, whereas that of the E-site is around 50 Å away. Since the first high resolution crystal structures of the ribosomal subunits and the 70S ribosome were solved (Ban 2000; Wimberly et al. 2000; Yusupov et al. 2001), detailed interactions be-
between the ribosomal subunits and mRNA, tRNAs, as well as the translation factors in various states of protein synthesis have been elucidated. An excellent review of ribosome structures (Schmeing & Ramakrishnan 2009) summarizes what is known about the structural mechanism of translation.

During translation, a ratchet-like rotation between the two ribosomal subunits and a relative movement between the “head” and the “body” of the 30S subunit (swiveling of the head) have also been observed (Frank & Agrawal 2000; Schuwirth et al. 2005). These movements are fundamental features of the ribosome.

![Figure 1](image)

*Figure 1.* Structure of the ribosome in complex with mRNA and A- P- and E- site tRNA. The 50S ribosomal RNA is shown in cyan and 50S ribosomal proteins are shown in blue. RNA in the 30S subunit is shown in yellow and ribosomal proteins in the 30S subunit are shown in orange. The A- P- and E- site tRNAs are shown in magenta, green and yellow with surface representation, respectively. The short mRNA is shown in black. Coordinates are from a structure of the *T. thermophilus* 70S ribosome (Voorhees et al. 2009). Figure was made in PyMol.

The bacterial translation cycle consists of four steps: initiation, elongation, termination and ribosome recycling. Each step involves GTP-hydrolysis-mediated complex formation or dissociation. A summary of the translation cycle and the involved protein translation factors based on current biochemical and structural knowledge will be introduced in the following sections.
Initiation

Initiation is the rate-limiting step for translation. To start synthesizing a protein, the ribosome needs to recognize a start codon (AUG, GUG, UUG, or AUU) on the mRNA, which can encode the initiator N-formylmethionine (fMet)-tRNA$_{\text{fMet}}$, and it needs to position the initiator tRNA over the P-site. The most common initiation triplet is AUG, whereas the others are rare. Another element on an mRNA other than the start codon is the Shine-Dalgarno (SD) sequence (Shine & Dalgarno 1975), which is also necessary for ribosome recognition. The SD sequence is optimally 5-9 bases upstream of the start codon, complementary to the 3’-end of 16S rRNA (Steitz and Jakes 1975), containing the consensus AGGAGGU bases. The accurate positioning of the ribosome is directed by binding of mRNA, the initiator tRNA and initiation factors (IF) 1-3 (Gualerzi and Pon 1990). The first step of initiation is the association of IF3 and the 30S ribosomal subunit that has been split from the 50S after translational termination (See termination section below). The binding of mRNA, IF1, IF2 and initiator tRNA to the 30S-IF3 complex results in the formation of 30S initiation complex (30S-IC). The formation of 70S-IC is stimulated by IF2, which is a GTPase, and coupled with release of IF3 (Milon et al. 2008). With the assistance of IF2-dependent GTP hydrolysis, phosphate release and subsequently IF2 release, the initiator tRNA at the P-site reaches the PTC and prepares the ribosome for elongation (Grigoriadou et al. 2007).

The elongation cycle

The completion of initiation leaves the 70S ribosome with an aminoacylated initiator tRNA at the P-site and an empty A-site ready for the elongation cycle. An outline of the elongation cycle is shown in Figure 2. The elongation cycle consists of several steps that are involved in sequentially adding amino acids to a polypeptide chain according to the sequence in the mRNA. The first step is decoding, during which the next aminoacyl-tRNA (aa-tRNA) is delivered to the ribosomal A-site in complex with elongation factor Tu (EF-Tu) and GTP. This is followed by peptide bond formation at the PTC resulting in the addition of one amino acid to the nascent polypeptide chain. Following peptidyl transfer, the tRNAs move spontaneously in the 50S subunit, from the “classical” state to a “hybrid” state, until elongation factor G (EF-G) binds. EF-G then catalyzes translocation under GTP hydrolysis, which moves the tRNAs and mRNA with respect to the ribosome and leaves an empty A-site for the next round of elongation.
Figure 2. The elongation cycle. Each cycle consists of three major steps: decoding, peptidyl transfer, and translocation. This figure is reproduced from (Shoji et al. 2009) with permission from the publisher.

Decoding

Decoding ensures that the ribosome selects the correct aa-tRNA for the A-site. This is extremely accurate; the missense error rate in protein translation in vivo has been estimated to be no more than $3.4 \times 10^{-4}$ per codon (Kramer & Farabaugh 2007). Base pairing between the mRNA and the tRNA ASL is the basis for codon recognition. However, the cognate base pairing, which has a perfect match for all the three bases, and the near-cognate base pairing, which has one base mismatch, have too small energy differences for the base pairs alone to be selected accurately (Xia et al. 1998; Ramakrishnan 2002). Nonetheless, the ribosome is able to discriminate the difference between cognate and near-cognate binding. The ribosome plays a direct role in ensuring discrimination and fidelity.

Most aa-tRNAs in the cell exist as ternary complexes with EF-Tu and GTP. The crystal structure of the ternary complex of EF-Tu, Phe-tRNA\textsuperscript{Phe} and a GTP analog was solved about two decades ago (Nissen et al. 1995). The binding of a correct aa-tRNA to the ribosomal A-site takes place in a series of kinetic steps, which have been elucidated by pre-steady state kinetics measurements (Rodnina and Wintermeyer 2001) and smFRET (Blanchard et al. 2004) studies. As the ternary complex is delivered to the A-
site, the ribosome can discriminate tRNAs in two stages: initial selection, which takes place before GTP hydrolysis; and proofreading, which occurs after GTP hydrolysis and before peptide bond formation. Initial selection involves binding of the ternary complex, codon recognition, GTPase activation and GTP hydrolysis. Upon GTP hydrolysis and phosphate release, EF-Tu rearranges its conformation into GDP form and dissociates from the ribosome. Thereafter the A-site tRNA is either accommodated in the PTC, ready for peptide bond formation, or rejected by ribosome proofreading.

Structural studies of various ribosomal complexes have elucidated the direct roles of the ribosome in decoding. The universally conserved bases on the 16S rRNA - G530, A1492 and A1493 - were found to be shielded by A-site tRNA from chemical probing (Moazed and Noller 1986). Nuclear magnetic resonance (NMR) spectroscopy studies have also shown that the error-inducing aminoglycoside paromomycin (see the aminoglycoside section below) induces local conformational changes at the A-site of 16S rRNA including A1492 and A1493 (Fourmy et al. 1998; Fourmy et al. 1996). Crystal structures of the 30S subunit with mRNA, cognate tRNA at the A-site with and without paromomycin (Ogle et al. 2001) reveal that the essential bases A1492, A1493 and G530 interact with the minor groove of the first two base pairs. These bases monitor the geometry of Watson-Crick base-pairing and discriminate against near-cognate tRNA, whereas the third position of the codon is free to accommodate U-G wobble base pairs. The geometry checking enhances the specificity of codon-anticodon recognition to a much greater extent than base-pairing alone. The recognition of cognate tRNA by the ribosome has been proposed to be conducted by an induced fit model coupled with a transition from an open to a closed form of the 30S subunit (Ogle et al. 2001; Ogle et al. 2003; Ogle et al. 2002). Another structure of the 70S ribosome with three tRNAs bound in A-, P- and E-sites (Selmer et al. 2006) confirms the results by Ogle and coworkers for 70S. A recent crystallographic study shows that the 30S domain closure upon binding of a tRNA to the A-site forces the first two base pairs to assume the Watson-Crick geometry (Demeshkina et al. 2012). Furthermore, the crystal structure of EF-Tu and aa-tRNA bound to the 70S ribosome (Schmeing et al. 2009) reveals details of how tRNA distortion allows communication between the decoding center on the 30S subunit and the EF-Tu GTPase center on the 50S subunit. A more recent crystal structure of the 70S ribosome, EF-Tu, aa-tRNA and a GTP analog captures EF-Tu in its active form (Voorhees et al. 2010), showing a critical and conserved interaction between the catalytic histidine of EF-Tu and A2262 at the sarcin-ricin loop of the 23S rRNA, suggesting a universal mechanism for GTPase activation on the ribosome.
Peptide bond formation

The accommodation of the correct aa-tRNA to the PTC is followed by rapid and irreversible peptide-bond formation (Wohlgemuth et al. 2006), which is the central chemical reaction during protein synthesis. The α-amino group of the A-site aa-tRNA makes a nucleophilic attack on the ester carbon of the peptidyl-tRNA to form a new peptide bond. The peptide then sits on the A-site tRNA (Green & Lorsch 2002). To do this, the CCA ends of the aa-tRNA and peptidyl-tRNA bind to domain V of the 23S rRNA, which is the catalytic site. The PTC has been precisely localized in the *Haloarcula marismortui* 50S subunit structure (Nissen et al. 2000), in which no protein is visible near the catalytic site. The structures of 50S subunit in complex with peptidyl transferase substrate analogues (Schmeing et al. 2005) suggest an induced fit model in which the active site bases and the substrates reposition to allow peptide bond formation. Later, structures of the intact ribosome containing the active site at the PTC (Voorhees et al. 2009) have confirmed the high similarity in the PTC conformations for the large subunit and the intact ribosome, but they also reveal interactions between the ribosomal proteins L16 and L27 and the tRNA substrates. This sheds light on the importance of these proteins in peptidyl transfer together with previous biochemical (Moore et al. 1975; Maguire et al. 2005) and computational data (Trobro & Aqvist 2008). The N-terminal tail of L27 may be involved in stabilizing the tRNA substrates in the PTC and helping to position their 3’ ends for peptidyl transfer. Ribosomal protein L16 may also be involved in stabilizing the A-site tRNA.

Translocation

Peptide bond formation results in a ribosomal complex with the A-site tRNA attached to a nascent polypeptide chain and a deacylated P-site tRNA, termed pre-translocation (PRE) complex. The tRNAs and the mRNA need to move relative to the ribosome such that the mRNA moves precisely one codon, and the tRNAs translocate from the A and P sites to the P and E sites before the next round of elongation cycle. Translocation takes place in several steps. First, following peptide bond formation, the acceptor ends of the tRNAs move spontaneously within the 50S subunit resulting in a hybrid-state complex (Moazed and Noller 1989). In a hybrid state, the anticodon stem loops of the peptidyl-tRNA and deacylated tRNA reside at the A and P sites of the small subunit and their acceptor ends interact with the P and E sites of the large subunit, occupying the A/P and P/E site (the first and the second letter present the binding site on 30S and 50S subunits), respectively. After binding of GTPase EF-G, the codon-anticodon helices move in the 30S subunit, where peptidyl-tRNA and deacylated tRNA occupy the P and E sites of both subunits. At this point, the complex is a post-translocation
(POST) complex. The POST complex has a vacant A-site, ready for the next round of elongation.

**Ribosome motions during translocation**

The EF-G-dependent translocation is coupled with conformational changes of the ribosome. Before peptidyl transfer or after translocation, when there is a peptidyl-tRNA bound at the ribosomal P-site, the ribosome is “locked” in a classical state, see Figure 1 (Valle et al. 2003). After peptidyl transfer, the ribosome undergoes a ratchet-like rotation between the ribosomal subunits where it becomes “unlocked”, i.e., there is a 3-10 degree of counterclockwise rotation of the body of the 30S subunit with respect to the 50S subunit (Frank and Agrawal 2000; Valle et al. 2003). This movement is essential for translocation (Horan and Noller 2007). The hybrid state tRNAs, coupled with spontaneous ratcheting of the ribosome, were visualized using cryo-EM (Agirrezabala et al. 2008; Julián et al. 2008). The formation of tRNA hybrid states is believed to occur in two steps: the P/E state forms first and then the A/P state (Munro et al. 2007). Single-molecule FRET studies show that the ribosome oscillates between the classical and rotated states. It is unlocked in the PRE complex until the binding of EF-G stabilizes the rotated state (Blanchard et al. 2004; Cornish et al. 2008; Munro et al. 2010; Spiegel et al. 2007).

The rotation of the body of 30S relative to the 50S subunit is followed by another motion that is involved in the ratcheting of the ribosome. This motion is the rotation of the head domain of the 30S subunit, which is believed to play an important role in controlling the position of the tRNAs and facilitating tRNA-mRNA movement (Frank et al. 2007; Schuwirth et al. 2005). The rotation of the head is about 12 degrees, equivalent to a 20 Å movement at the subunit interface (Schuwirth et al. 2005).

The “opening” and “closing” of the ribosomal L1 stalk, which is composed of helices 76-78 from the 23S rRNA and the ribosomal protein L1, is also required in translocation to facilitate the movement of the tRNAs (Shoji et al. 2009). The L1 stalk interacts with the elbow region of the E-site tRNA in the closed conformation (Selmer et al. 2006; Korostelev et al. 2006). Single-molecule FRET studies show that the L1 stalk spontaneously establishes interaction with the newly deacylated tRNA after peptide bond formation (Fei et al. 2008). In the unrotated classical ribosome in which the P-site tRNA is charged, the L1 stalk is in an open form without interacting with the tRNA. By contrast, in the rotated ribosome in which the P-site tRNA is deacylated, the L1 stalk is in a closed form and interacts with tRNA (Valle et al. 2003; Fei et al. 2008).

**The role of EF-G and GTP hydrolysis in translocation**

Early kinetics studies show that EF-G can promote a single round of translocation in the presence of nonhydrolyzable GTP analogs, but GTP is needed
for multiple-turnover translocation (Kaziro 1978). EF-G with GDP or without any guanine nucleotide does not have any translocase activity (Katunin et al. 2002; Rodnina et al. 1997). The presence of both A-site and P-site tRNAs is required by EF-G to catalyze translocation, but it seems that only the ASL of the A-site tRNA and full P-site tRNA are essential for translocation (Joseph and Noller 1998). EF-G joins the PRE complex in the GTP-bound form, and GTP hydrolysis is followed by tRNA-mRNA movement and phosphate release, inducing further structural rearrangement of EF-G and the ribosome (Wilden et al. 2006; Savelsbergh et al. 2003). In solution, EF-G does not undergo a global conformational change when switching between the GTP- and GDP-bound forms. Rather, ribosome binding seems to trigger EF-G conformational change and the active overall conformation of EF-G is only acquired when it is bound to the rotated state of the ribosome (Hauryliuk et al. 2008). A recent single-molecule FRET study tracks the binding of EF-G and the ribosome conformation in real time, demonstrating the locking and unlocking model of the ribosome and correlation of EF-G binding with ribosome conformation (Chen et al. 2013). This study indicates that the EF-G-GTP complex continuously samples both rotated and classical states of the ribosome and binds with higher affinity to the rotated state, consistent with previous kinetic and single-molecule FRET data.

**Structure of EF-G and its interaction with the ribosome**

EF-G is a five-domain protein, with a molecular mass of ~78 kDa. The crystal structures of *Thermus thermophilus* EF-G in its apo and GDP-bound form were solved almost two decades ago independently by two groups (AEvarsson et al. 1994; Czworkowski et al. 1994).

![Figure 3. Structure of EF-G bound to the ribosome. Domains are shown in Roman numerals. Numbers 1-8 indicate ribosome contact areas, see description to the right. The coordinates are from ribosome-bound *T. thermophilus* EF-G (PDB code 2wri) (Gao et al. 2009).](image)
There are a number of GTPases involved in the bacterial translation cycle, namely IF2, EF-Tu, EF-G, RF3, SelB and LepA. All of these GTP-hydrolyzing enzymes share a highly conserved GTPase (G) domain - the catalytic domain. They bind to an overlapping region on the ribosomal 50S subunit (Helgstrand et al. 2007). GTPases are molecular switches with intrinsic GTPase activity. Proteins in the GTPase family undergo conformational changes upon GTP hydrolysis and switch between the functional GTP state and the inactive GDP state (Bourne et al. 1991). In EF-G, the G domain (domain I) is the largest domain, which is homologous to that in the GTPases of the Ras superfamily. There are three functionally important regions in the G domain: the P-loop (phosphate-binding loop), switch I (effector loop) and switch II. The P-loop has a consensus GXXXXGKS/T (X is any amino acid) motif, interacting with the β and γ phosphates of the nucleotide. The switch regions usually undergo conformational changes upon GTP hydrolysis, where the release of γ-phosphate relaxes the switch regions into the inactive GDP form. Domains I and II of EF-G share high similarity to the corresponding domains of other translational GTPases, such as EF-Tu and IF2 (Caldon et al. 2001; Caldon et al. 2003), although there is an extra ~90-amino acid insertion of a G’ subdomain in EF-G. Domains III, IV, and V are connected with the other two domains via a loop. The overall conformation of EF-G was found to have striking similarity to the EF-Tu ternary complex (Nissen et al. 1995), with domains III-V mimicking a tRNA. The structure of EF-G with a nonhydrolyzable GTP analog is known (Hansson et al. 2005a), but no major difference was found between the GDP- and the GTP-bound forms. This is unlike crystal structures of other GTPases, e.g. EF-Tu, where GTP hydrolysis involves a large global conformational change and helix unwinding at the nucleotide binding site (Kjeldgaard et al. 1993; Polekhina et al. 1996). All previously solved EF-G structures are from *T. thermophilus* (Hansson et al. 2005a; Hansson et al. 2005b; Laurberg et al. 2000; Czworkowski & Moore 1997; Czworkowski et al. 1994; Al-Karadaghi et al. 1996; AEvarsson et al. 1994), and the switch I region is disordered in all of them. Switch II, on the other hand, displays various conformations (Figure 7B). In a crystal structure of EF-G-2 bound with GTP (a homolog of EF-G with sequence identity of 34%), the switch I region is ordered as two helices (Connell et al. 2007).

Large-scale movement of EF-G is seen with cryo-EM when EF-G is locked on the ribosome during translocation either with a nonhydrolyzable GTP analogue (Connell et al. 2007) or with the antibiotic FA and GDP (Stark et al. 2000; Agrawal et al. 1998; Agrawal et al. 2001). Domain I interacts with the L7/L12 stalk on the 50S subunit, and domain IV contacts the shoulder of the 30S subunit in the PRE complex and reaches the decoding center after translocation. The 3.6 Å crystal structure of the 70S ribosome and EF-G-GDP complex trapped by FA in the POST state reveals more details of EF-G-ribosome interactions (Gao et al. 2009), including its interac-
tion with the L11 region and L10-L12 stalk and the interaction between domain IV and the decoding center. The overall structure of EF-G when it is bound to the ribosome and its interactions with it are illustrated in Figure 3. Recent crystal structures of the *E. coli* 70S ribosome in complex with EF-G and the nonhydrolyzable GTP analogue β,γ-methyleneguanosine 5’m-triphosphate (GMPPCP) in different states of ribosomal subunit rotation (Pulk & Cate 2013) indicate that EF-G binding to the ribosome stabilizes its switch regions. Thus, EF-G·GMPPCP is in a compact and rigid conformation favoring the rotated ribosome. The same study suggests that EF-G controls translocation by going from a rigid to a relaxed conformation after GTP hydrolysis. Another recent crystal structure of the *T. thermophilus* 70S ribosome in complex with EF-G and GMPPCP also reveals an EF-G-bound ribosome in the rotated state before GTP hydrolysis (Tourigny et al. 2013). In this structure, the switch regions are ordered and an inward movement of the L1 stalk has been observed, stabilizing the P/E tRNA. The key conserved residues Asp22, Lys25, and His87 in EF-G interact with the 50S subunit in a different conformation compared to the isolated and fully translocated structures. This study suggests a similar mechanism of activation of GTP hydrolysis between EF-G, EF-Tu, and possibly other translational GTPases.

**Termination**

The termination of a coding sequence requires a UAA, UAG or UGA stop codon. The elongation cycle ends when a stop codon moves into the A-site. A type I release factor (RF) recognizes the stop codon and promotes hydrolysis of the peptidyl-tRNA linkage in the PTC, resulting in release of the nascent peptide chain from the ribosome. In bacteria, there are two type I release factors, RF1 an RF2. Both of them can recognize the UAA stop codon, whereas UAG is only recognized by RF1 and UGA is only recognized by RF2 (Scolnick et al. 1969). The high resolution structures of the 70S ribosome in complex with RF1 (Laurberg et al. 2008) and RF2 (Weixlbaumer et al. 2008; Korostelev et al. 2008) provide structural details of translational termination. The conserved GGQ motif of the release factors plays a major role in peptide hydrolysis. They show that it is positioned in the PTC in a special conformation for which the two glycines are critical. The type II release factor RF3 facilitates the release of type I release factors from the ribosome after peptide release. RF3 is a GTPase and is stably bound to GDP in solution *in vivo*. The crystal structure of RF3-GDP is strikingly similar to that of EF-Tu-GTP (Gao et al. 2007). The ribosomes in complex with RF1/2 act as guanine nucleotide-exchange factors, allowing binding of RF3-GTP to the ribosome (Zavialov et al. 2001). The binding of RF3-GTP induces conformational changes of the ribosome and release of the type I release factors.
GTP hydrolysis is required for dissociation of RF3 from the ribosome (Zavialov et al. 2001).

Ribosome recycling

Dissociation of RF3 from the ribosome under GTP hydrolysis leaves the ribosome with an mRNA and a deacylated tRNA in the P-site. To initiate a new round of protein synthesis, the ribosome must be recycled as subunits. This process is carried out by a ribosomal recycling factor (RRF) together with EF-G (Hirashima & Kaji 1973). The crystal structure of RRF from *Thermotoga maritima* contains two domains (Selmer et al. 1999). The detailed interactions of the RRF with the ribosome are revealed by the complex crystal structure of the 70S ribosome and RRF from *T. thermophilus* (Weixlbaumer et al. 2007). However, a crystal structure of the 70S ribosome with both RRF and EF-G is still lacking to date. EF-G is the only translation factor that has two distinct functions in different stages of translation. EF-G hydrolyzes GTP in both translocation and ribosome recycling, but when GTP is replaced with nonhydrolyzable analogs, ribosome recycling is completely blocked (Zavialov et al. 2005). Detailed interactions between EF-G and RRF can be seen with cryo-EM (N. Gao et al. 2007), where domains III-V of EF-G interact with the hinge region, the head domain and the C-terminus of RRF, respectively. EF-G undergoes domain rearrangement and switch I conformational change upon GTP-hydrolysis. These conformational changes are proposed to induce head domain rotation of RRF and eventually subunit dissociation (N. Gao et al. 2007).

In summary, the translation process is a fundamental but complex pathway. Although the accuracy and efficiency of translation is high, any step of the translation cycle can go wrong. Errors in protein synthesis include substitution, insertion, deletion mutations, premature termination, or even a completely different protein if the reading frame is changed. The following sections will address such aspects as error-inducing antibiotics, antibiotic resistance, and ribosomal frameshifting.
Translation inhibition by antibiotics

The ribosome is one of the major in vivo targets for antibiotics. Antibiotics were originally defined as small metabolic products produced by microorganisms that inhibit the growth of or even destroy other micro-organisms (Waksman 1947). The term antibiotic is now used in a broader sense to include chemicals that inhibit micro-organisms, viruses or eukaryotic cells (Spahn & Prescott 1996).

Since their discovery, ribosome-targeting antibiotics have been of major clinical importance; however, bacterial strains with resistance to the drugs often emerge rapidly in clinical use. The ribosome-targeting antibiotics have also been used as tools for capturing functional states of the ribosome and understanding the mechanism of the various steps of translation described above. How do antibiotics interact with the ribosome and how do the drugs take their action? These questions have been studied extensively in recent years by biochemical and structural methods. Most inhibitors target functionally important sites of the rRNA on the surface of the two subunits (Poehlsgaard and Douthwaite 2005). In the 30S subunit, the best-investigated site is the decoding center, which is targeted by aminoglycosides (Tenson and Mankin 2006), and crystal structures of the 30S in complex with various antibiotics reveal the structural basis for the action of the aminoglycosides (Brodersen et al. 2000; Carter et al. 2000). In the 50S subunit, the antibiotics target mainly three regions where they interfere with GTP hydrolysis, peptide bond formation and intersubunit interactions (Poehlsgaard and Douthwaite 2005). Translation factors are also common targets for antibiotics, e.g. kirromycin locks EF-Tu-GDP on the ribosome (Parmeggiani and Nissen 2006; Wolf et al. 1974) and fusidic acid (FA) locks EF-G-GDP on the ribosome (Bodley et al. 1969) etc.

Although a wide range of drugs target the bacterial ribosome, the following paragraphs will only introduce the modes of action for FA, which inhibits translocation, and a few aminoglycosides, which interfere with decoding.

Fusidic acid and its action

Fusidic acid, derived from Fusidium coccineum, is a bacteriostatic antibiotic discovered in the 1960s (Godtfredsen et al. 1962). The chemical structure of FA is shown in Figure 4. It has been in clinical use for the treatment of
staphylococcal infections for over four decades. *Staphylococcus aureus* is one major target for clinical treatment. FA inhibits translocation by preventing dissociation of EF-G. It acts on EF-G only when it is bound to the ribosome, locking and stabilizing EF-G there after GTP hydrolysis and translocation (Willie et al. 1975; Valle et al. 2003). FA binds to the ribosome bound EF-G-GDP with a $K_d$ of around 0.4 $\mu$M (Okura et al. 1970), whereas it does not bind to free EF-G. This indicates that the FA-binding pocket on EF-G is only formed when it is bound to the ribosome. The binding pocket can be seen in the 70S-EF-G-GDP-FA complex structure (Gao et al. 2009), where FA is surrounded by the switch II region of the G domain, domains II and III of EF-G.

![Chemical structures of antibiotics](image)

*Figure 4.* Chemical structures of antibiotics mentioned in this thesis.

### Aminoglycosides and their mode of action

Aminoglycosides are a family of antibiotics that are composed of amino-modified sugars and with considerable structural diversity. They are often broad-spectrum antibiotics produced by bacterial or fungal metabolism. The first aminoglycoside streptomycin was isolated from *Streptomyces griseus* in the 1940s (Schatz et al. 1944). Clinically aminoglycosides are for example used to treat bacteremia, endocarditis, infections of the abdomen and urinary tract. The aminoglycosides mentioned in this thesis are shown in Figure 4.
Aminoglycosides bind to 16S rRNA at the decoding region of the 30S ribosomal subunit (Moazed and Noller 1987), causing misreading and inhibiting translation (Cabanas et al. 1978; Davies et al. 1965). Most of them are inactive against eukaryotic ribosomes. It is believed that the identity of the nucleotide at position 1408 on 16S rRNA, which is an adenine in prokaryotes and a guanosine in eukaryotes, is the determinant of specificity of many aminoglycosides (Recht et al. 1999).

The crystal structure of 30S subunit in complex with the aminoglycosides paromomycin and streptomycin and an aminocyclitol antibiotic, spectinomycin, that is closely related to the aminoglycosides reveals the structural basis for their mode of action (Carter et al. 2000). Streptomycin, which can cause extensive misreading of mRNA (Davies et al. 1964), binds tightly to the phosphate backbone of 16S rRNA and makes contact with protein S12 that is encoded by \textit{rpsL} gene. The binding of paromomycin flips out A1492 and A1493 from helix 44 into a position where they could interact with the minor groove of the codon-anticodon helix, increasing the error rate of the ribosome. The binding of the aminoglycosides also induces a domain closure of the 30S subunit that usually takes place when a correct tRNA is bound. The binding sites of streptomycin, spectinomycin and paromomycin are shown in Figure 5.

\begin{figure}
    \centering
    \includegraphics[width=\textwidth]{figure5.png}
    \caption{Binding sites of the antibiotics on the bacterial ribosomal 30S subunit. The 30S subunit model is based on (Voorhees et al. 2009) and the position of the binding sites are based on Figure 1 in (Poehlsgaard & Douthwaite 2005). Figure was made in PyMol and Microsoft Powerpoint.}
\end{figure}

Spectinomycin interferes with EF-G-catalyzed translocation (Bilgin et al. 1999). The crystal structure shows that the antibiotic binds to the end of helix 34, making interactions with G1064 and C1192; it is also in close proximity to a loop of S5 and part of helix 28 (Carter et al. 2000). Spectinomycin
has a rigid structure and binds near the pivot point for the head rotation of 30S subunit which is required in the EF-G-catalyzed translocation. It is likely that binding of the drug sterically blocks the head rotation of the 30S subunit and thereby inhibits translocation.

Other aminoglycosides will not be discussed in this thesis in detail except to mention that the common effect of their binding is a conformational change of the ribosome mimicking the closed state induced by cognate tRNA binding.
Antibiotic resistance

Antibiotics are essential in the fight against infectious diseases, and bacterial resistance against antibiotics is widely recognized as a health threat. Therefore, understanding the mechanisms of antibiotic resistance is critical, both for maintaining the effectiveness of existing drugs and for developing new ones. Extensive studies on antibiotic resistance have provided considerable knowledge of its mechanisms, although much remains to be investigated. There are several mechanisms of resistance: inactivation of the drug, modification of the drug target, decreased uptake or increased efflux, and altered metabolism (bypassing the inhibited pathway) (Giedraitienė et al. 2011). The concept “antibiotic resistome” has been proposed to represent the collection of genes that contribute to resistance directly or indirectly (Wright 2010). In the following paragraphs, relevant parts of the current understanding of resistance mechanisms to FA and aminoglycosides will be introduced.

Resistance mechanisms to fusidic acid

_Staphylococcus aureus_ is a bacterium that is a common cause of respiratory diseases and skin infections, against which FA has been used for clinical treatment. The first mechanism of resistance to FA to be identified was alteration of the _fusA_ gene encoding EF-G, i.e., the drug target (Chopra 1976) (_fusA class_). Later, numerous _fusA_ class mutations were identified by DNA sequencing from either mutants evolved and selected in the lab (Besier et al. 2003; Johanson et al. 1994) or clinically isolated FA-resistant mutants (Besier et al. 2003; Lannergård et al. 2009; Nagaev et al. 2001; Norström et al. 2007). A subset of the _fusA_ class mutations cause a small-colony variant (SCV) phenotype (Norström et al. 2007). These are subpopulation of cells having slow growth rate and resistance towards antibiotics. A second class of mutations with the SCV phenotype have been identified in the _rplF_ gene, which encodes ribosomal protein L6; these are classified as _fusE_ mutants (Norström et al. 2007). The C-terminus of L6 is in close proximity to domain V of EF-G when it is bound to the ribosome after translocation (Gao et al. 2009).

Resistance to FA can also result from expression of the EF-G protective FusB-type proteins. The _fusB_ gene is carried on plasmid pUB101 (O’Brien 2002). FusB binds to _S. aureus_ EF-G, but not to _E. coli_ EF-G (O’Neill and
FusB was suggested to promote the dissociation of the EF-G-GDP complex from the ribosome (Cox et al. 2012). The crystal structure of FusB was solved and mapping of the FusB binding site using hybrid constructs between \textit{S. aureus} and \textit{E. coli} EF-G was performed (Guo et al. 2012). The hybrid construct containing \textit{S. aureus} domain IV and domains I, II, III and V from \textit{E. coli} show the same binding to FusB as the wild type \textit{S. aureus} EF-G, indicating that FusB binds to domain IV of EF-G. The detailed mechanism of FusB-mediated FA resistance, however, still remains unclear.

We solved the \textit{S. aureus} EF-G structure at high resolution (1.9 Å), and analyzed and mapped all known clinically isolated FA-resistance mutations to our structure (Paper I).

**Fitness cost and compensation**

Any characteristic that enables an organism to survive may be defined as fitness. The most common measure of fitness is the growth rate. Antibiotic resistance is often associated with a fitness loss, i.e., slow growth for the resistant bacteria (Andersson and Levin 1999). The FA-resistant mutations in EF-G often are evolved to be associated with a secondary mutation that partly or fully compensates the fitness loss (Nagaev et al. 2001). One of the primary mutations that cause strong FA resistance is F88L. This mutation is located at the tip of the switch II loop and involved in direct contact with FA when bound to the ribosome (Gao et al. 2009). The F88L mutant exhibits a significant growth defect, yet an additional mutation M16I can compensate the fitness loss and retain FA resistance (Nagaev et al. 2001). Interestingly, the M16I mutation itself confers an FA hypersensitive phenotype. We tried to clarify the mechanism of how the resistance is retained and how the fitness loss is compensated. In our study, crystal structures of EF-G mutants F88L, M16I, and the fitness compensated double mutant F88L/M16I were solved and analyzed together with kinetic data (Paper II).

**Resistance mechanisms to aminoglycosides**

Bacterial resistance towards aminoglycosides is acquired by several mechanisms: inactivation of the drug by aminoglycoside-modifying enzymes; structural alteration of the drug binding site on the ribosome, e.g. 16S rRNA methylation and ribosomal point mutations; extrusion of drugs from the cell by efflux pumps; and decreased cell membrane permeability. Among these mechanisms, drug modification by enzymes gives the highest level of aminoglycoside resistance (Azucena & Mobashery 2001), and only this mechanism will be discussed here.
Aminoglycoside-modifying enzymes (AMEs)

Aminoglycoside-modifying enzymes transfer a functional group to the aminoglycoside structure resulting in inactivation of the antibiotic via diminished binding to the drug target. There are three types of such enzymes: aminoglycoside nucleotidyltransferases (ANTs), which transfer a nucleotidyl group from a nucleotide triphosphates; aminoglycoside acetyltransferases (AACs), which transfer the acetyl group from acetyl-CoA; and aminoglycoside phosphotransferases (APHs), which transfer the phosphoryl group from ATP (Shaw et al. 1993). Each type also consists of different enzymes that differ in the position of the substrate to which the modification is added. A detailed classification of each class of AME based on the modification sites on the drugs is summarized in a relatively recent review (Ramirez & Tolmasky 2010). Representative crystal structures for each class have shed light on enzyme function, and reveal some unexpected connections to other enzyme families (Wright 1999). Most of the members of the APH family share more than 25% sequence identity, and they show similarity to protein kinases even though the sequence identity with them is below 5%. A few known AAC structures show similar folds to each other. In contrast, the ANT enzymes share little overall sequence identity with each other.

ANT is the smallest class of AMEs and is the least studied. The only known ANT crystal structure is a kanamycin nucleotidyltransferase (KNTase) from *S. aureus*. Structures for it exist both in its apo form and in ternary complex with kanamycin and a nonhydrolyzable ATP analog AMPCPP (Pedersen et al. 1995; Sakon et al. 1993). The KNTase, an ANT (4’) modifying the 4’ position of the drug, functions as a homodimer. The active site is at the interface of the two monomers and residues from both monomers form and stabilize the binding pocket for kanamycin and ATP (Pedersen et al. 1995). Each monomer consists of two domains of approximately the same size. The N-terminal domain is a nucleotidyltransferase domain, and the C-terminal domain is an up-and-down α helical bundle. This enzyme shares structural similarity with DNA polymerase B. In the same KNTase structure, there is no specific interaction between the enzyme and the adenine ring, although there is extensive hydrogen bonding network between AMPCPP and the enzyme, which explains why this enzyme can also utilize other nucleotides as substrate.

Aminoglycosides have positive charges, therefore they tend to bind to negatively charged pockets in structured RNAs or proteins. Results from molecular simulations suggest that the most favorable binding sites for aminoglycosides are similar in the modifying enzymes and the rRNA (Romanowska et al. 2013).

A streptomycin/spectinomycin adenyltransferase gene (*aadA*) was first identified and sequenced in *E. coli*. (Hollingshead & Vapnek 1985). Its encoded protein AadA is an ANT (3”) modifying the 3” position of the sub-
strates; however, no crystal structure of ANT(3”) has been reported to date. The expression of a cryptic, chromosomally located \( aadA \) gene in \( Salmonella enterica \) has shown activation in SCVs or by growth in minimal media, and the gene expression is positively regulated by the stringent response regulator tetraphosphate (ppGpp) (Koskinemi et al. 2011). The fact that AadA can act on both streptomycin and spectinomycin despite their structural differences makes it interesting to solve its structure. We have solved the crystal structure of AadA from \( Salmonella enterica \) and analyzed the structure in relation to the current knowledge of this enzyme family (Paper III).
Ribosomal frameshifting

The accuracy of translation requires not only the correct tRNA corresponding to the A-site codon to be delivered to the ribosome. It also requires the ribosome to maintain a reading frame of three nucleotides. The consequence of a shift of the reading frame would be a complete alteration of the protein sequence starting from the frameshift site, and the ribosome would continue reading in the new frame until it encounters a stop codon. Frameshifts are extremely rare errors during translation. The error rate is estimated to be $10^{-5}$ or less (Parker 1989), which is at least 10-fold lower than missense errors. Some genes have evolved sequences, termed programmed frameshifting sites, which can manipulate the ribosome to promote non-canonical decoding, thereby inducing frameshifting efficiently (Farabaugh 1996). Programmed frameshifts have been mainly described in viruses, retrotransposons, and also some cellular genes. Ribosomes can also shift frame at simple frameshift-prone sequences. In some genes, the ribosome shifts the reading frame by one base in the upstream direction, causing a -1 frameshift; in others, the reading frame is shifted by one base in the downstream direction, namely a +1 frameshift. The ribosome is even capable of reading through a stop codon, or bypassing a short piece of nucleotide sequence to continue translation either in-frame or in a new frame. All of these frameshifting events require the ribosome to pause during translation elongation and let the kinetically unfavorable alternatives to occur (Farabaugh 1996). The frameshift sites have also been used as tools to probe the mechanisms of reading frame maintenance, but the detailed mechanism of how the reading frame is maintained is still unclear.

A major step forward towards understanding of the mechanism of ribosomal frameshifting is the structural knowledge of ribosome in complex with tRNAs and mRNAs. The P-site tRNA clearly has more extensive interactions with both ribosomal subunits compared to the A-site tRNA (Selmer et al. 2006), such that the P-site tRNA is held tightly by the ribosome and maintains the reading frame. An excellent review on ribosomal frameshifting has been written by Atkins and Björk (Atkins & Björk 2009), where they emphasized the pivotal role of the ribosomal P-site on frameshifting events by summarizing how various alterations of tRNA, rRNA, or the ribosomal proteins induce frameshifting. The structure of a tRNA is given in Figure 6. When a frameshift has occurred, the action of a suppressor tRNA which reads a non-triplet codon may restore the proper reading frame. Some +1
frameshift mutation suppressor tRNAs contain an extra base in their anticodon loops (Roth 1981), and the interactions between a few four-base anticodon ASLs and the mRNA at the ribosomal decoding center have been visualized in crystal structures (Dunham et al. 2007). Other frameshift mutation suppressor tRNAs have either alterations in their primary sequence or modification deficiency. Studies on a number of +1 frameshift mutation suppressors support a model where after a three-base translocation, the grip of the ribosome to the P-site tRNA is altered. The mRNA encounters a +1 slip, thereby inducing a frameshift. This model has gained further support from an analysis of independently isolated frameshift-suppressor mutants (Jäger et al. 2013).

Figure 6. Structure of a standard tRNA with the conventional numbering for the locations of the different nucleotides (left). Three-dimensional structure of yeast tRNA\textsuperscript{Phe} with various regions indicated (right). This figure is reproduced from (Atkins & Björk 2009) with permission from the publisher.

Although it is generally accepted that most types of frameshifting occur at the ribosomal P-site, there are cases that most likely do not, e.g. some -1 framshifts were proposed to depend on doublet base pairing in the A-site. This is the simplest explanation but it remains to be confirmed experimentally. Two normal tRNAs promote ribosomal frameshifting when the MS2 virus genome is translated \textit{in vitro} using \textit{E. coli} cell extract (Atkins et al. 1979). Wild type \textit{E. coli} tRNA\textsuperscript{Ser} (anticodon GCU) is able to decode a GCA alanine codon to cause -1 frameshifting when extra tRNA\textsuperscript{Ser} is added to the \textit{E. coli} cell extract. Even without a perturbed balance of tRNAs, wild type tRNA\textsuperscript{Thr} (anticodon GGU) decodes CCG or CCA proline codons, inducing -1 frameshifting.

An early study with anticodon replacement experiments showed that tRNA\textsuperscript{Phe} with a GCU anticodon also promotes frameshifting, indicating that
the anticodon loop, but not the rest of the tRNA structure, makes a major contribution to the frameshifting event (Bruce et al. 1986). The same study also investigated the importance of bases 33-36 at the anticodon loop to frameshifting. In conclusion, it seems that the tRNA$^{\text{Ser}}$ anticodon must be maintained to obtain efficient frameshifting, but base 33 can better tolerate substitution than bases 34-36. Characterization of the frameshift event by protein sequencing shows that tRNA$^{\text{Ser}}$ can promote frameshifting at GCA, GCU and GCC alanine codons, but GCA is the most active of the three; tRNA$^{\text{Thr}}$ induces frameshifting at CCG and CCA, but not at CCU codons (Dayhuff et al. 1986).

A simple explanation would be that the doublet base-paired tRNA-mRNA is translocated in a step size of two bases, inducing a -1 frameshift. However, there are other possibilities. In a standard codon-anticodon interaction, the 7-base anticodon loop employs a 2:5 stacking (two on the 5’ side and five on the 3’ side). A possible A-site base pairing has been proposed to be a 1:6 stacking (one on the 5’ side and six on the 3’ side). This would allow the anticodon base 33 to base pair with the third codon base and revert to a 2:5 stack at the P-site after 3-base translocation. This in turn allows the third base to be the first base of the next codon (Weiss 1984; Atkins et al. 2000).

In our study, we measured the binding affinity of the frameshift tRNAs to the frameshifting sites and examined the roles of different anticodon bases. (Paper IV).
Methodology

The following section is a brief description of the methods that are used in the studies included in this thesis. Detailed procedures can be found in the corresponding papers.

Component preparation

Protein expression and purification
Wild type *S. aureus* EF-G and three EF-G mutants were expressed and purified using a standard procedure as described by R. Koripella (Paper I and II). A pEXP5 construct of C-terminally His-tagged AadA was used for expression. Native and selenomethionine-substituted AadA proteins were overexpressed in *E. coli* BL21 star cells. In both cases the proteins were first purified using nickel-immobilized metal affinity chromatography. The purified fractions were then subjected to size exclusion chromatography and the purified proteins were concentrated for crystallization trials and storage (Paper III). An extensive high-salt wash was performed before elution in the nickel-column step to exclude nucleic acid contaminants. Purity of proteins was checked using SDS-polyacrylamide gel electrophoresis.

tRNA and mRNA production
In our study, tRNAs were cloned into plasmid pBSTNA V-2 and overexpressed in *E. coli* strain HMS174 using the constitutive lpp promoter. Total tRNA was extracted using phenol and purified using hydrophobic interaction chromatography. Fractions were collected, concentrated and stored for further use. Purity of tRNAs was checked using denaturing polyacrylamide electrophoresis. The ASLs used in this study were chemically synthesized by Dharmacon, USA.

The mRNAs used in our study were *in vitro* transcribed using T7 RNA polymerase. (Paper IV)

tRNA and ASL labeling with Phosphorus 32 ($^{32}$P)
The 3’ end of the tRNA or ASL molecules was labeled with $^{32}$P using alkaline phosphatase (Amersham Biosciences). Labeled material was gel purified, phenol extracted and ethanol precipitated. The pellet was recovered in a minimum volume of water and specific activity was determined using a scintillation counter. The labeled material was stored at -20 °C for further use.
Ribosome purification
70S ribosomes were purified from *E. coli* MRE600 cells by sucrose gradient ultra-centrifugation. The 70S peak was detected by UV$_{260}$ absorption at 40-50% sucrose concentration. The purified ribosomes were shock frozen and stored at -80 °C for further use. (Paper IV)

X-ray crystallography

Crystallization
Initial crystals of *S. aureus* EF-G grew in the Index screen (Hampton Research, USA) with condition 100mM Tris-HCl, pH 8.5, 200mM NaCl and 25% (w/v) PEG 3350 at 20 °C. The initial crystals were used for streak seeding into sitting-drop vapour diffusion-experiments. Optimized crystals were obtained a size of 100-150 µM within four weeks after setting up the drops. The same condition was used for growing and optimizing crystals of the mutant proteins. (Paper I and II)

Crystals of *S. enterica* AadA were obtained in the Morpheus screen (Molecular Dimensions, UK) with a reservoir solution containing 0.12M various alcohols, the 0.1M Morpheus buffer system 1, pH 6.5 and 30% ethylene glycol/PEG8000 at 8 °C. The crystals grew to a size of about 50-100 µM in 24 hours. (Paper III)

Soaking and co-crystallization with ligands were attempted. All the crystals were either transferred to the corresponding reservoir solution containing cryo-protectant or directly fished out (if the reservoir solution was adequate to serve as a cryo-protectant) and vitrified in liquid nitrogen prior to data collection.

Data collection and processing
A 2.1Å dataset for wild type EF-G was collected at the PXII beamline, Swiss Light Source (Villigen, Switzerland). Later a 1.9 Å dataset was collected at beamline ID14-1, ESRF (Grenoble, France). All the EF-G mutant datasets were collected at beamline ID23-1, ESRF, with resolution slightly better than 3Å. The dataset for selenomethionine-substituted AadA protein was collected at beamline ID14-4, ESRF, with resolution of 2.5 Å. Data were processed and scaled using the XDS package (Kabsch 1993).

Structure determination and refinement
All the EF-G structures were solved by molecular replacement using Phaser (McCoy et al. 2007). (Paper I and II) The AadA structure was solved by SAD phasing using PHENIX (Adams et al. 2010). (Paper III) Where appropriate, the structure models were improved by the used of automatic model building in ARP/WARP (Perrakis et al. 1999). Manual model rebuilding was
done in COOT (Emsley & Cowtan 2004). Structures were refined with CNS (Brünger et al. 1998), REFMAC (Murshudov et al. 1997) and PHENIX. Structure superposition was performed with O (Jones et al. 1991).

**Filter binding assay**

To measure the binding affinity of tRNAs to mRNAs, a nitrocellulose filter binding assay was performed. A-site binding affinity was determined using a constant amount of labeled cognate ASL, by either homologous competition with unlabeled cognate ASL or heterologous competition with unlabeled near-cognate ASL. Triplicate measurements were performed at each concentration. The buffer system and complex formation condition were consistent with what has been described previously (Ogle et al. 2002). The prepared complex solution was applied onto a nitrocellulose filter and washed quickly with cold buffer to get rid of small molecules unbound to the ribosomes. The filter was dried and dissolved in scintillation cocktail and radioactivity measured using a scintillation counter. The measurements were normalized and plotted against the concentrations of the competitors, and binding affinity was calculated. (Paper IV)
Aims of thesis

In this thesis, the research focuses on three aspects of elongation during protein translation. The first aspect is understanding *fusA*-type FA resistance and the accompanied fitness cost (paper I and paper II); second, understanding aminoglycoside resistance by a drug-modifying enzyme (paper III); third, understanding the mechanism of tRNA-induced -1 ribosomal frameshifting (paper IV).

We aimed to solve the crystal structure of EF-G from *S. aureus*, which is the drug target of FA. The goal was to solve the structures of wild-type EF-G in both apo and nucleotide-bound form. Also, we aimed to solve the structures of FA-resistant mutant F88L, FA-hypersensitive mutant M16I, and FA-resistant but fitness compensated double mutant F88L/M16I. We wanted to analyze the structures in relation to clinically isolated FA-resistance mutations and biochemical data provided by our collaborators (Koripella and Sanyal).

We also aimed to solve the crystal structure of an aminoglycoside modifying enzyme, adenyltransferase AadA, in both its apo form and with its substrates streptomycin/spetinomycin and cofactor ATP. We wanted to analyze the structure to gain insight into the mechanism of the enzymatic reaction and substrate specificity.

In the frameshifting project, we aimed to reproduce the -1 frameshifting induced by normal tRNAs (tRNA\textsubscript{Ser\textsuperscript{3}} and tRNA\textsubscript{Thr\textsuperscript{3}}) in a small, defined system and measure the binding affinity of tRNA to mRNA in regular or frameshift context. We wanted to understand the mechanism of these particular frameshifts and define the roles of various bases on the anticodon stem loop of the frameshift tRNAs. We also aimed to solve the crystal structures of the 70S ribosome in complex of frameshifting mRNAs and tRNAs.
Summary of current research

Paper I: Mapping of FA resistance mutations on the drug target

The structure
Prior to our study, all available structures of EF-G were from *T. thermophilus*. Very few studies had been done using EF-G from the major drug target *S. aureus*. In this work, we solved the crystal structure of *S. aureus* EF-G at 1.9Å resolution using molecular replacement. The crystals belong to space group P2$_1$. There are two molecules in the asymmetric unit (AU), forming a two-fold noncrystallographic symmetry (NCS). The two molecules in the AU have identical conformations despite the fact that they make different crystal contacts with the neighboring molecules. A *Thermus thermophilus* EF-G structure was chosen as search model (PDB code 1fnm). The multidomain characteristic of EF-G and the propensity of inter-domain movement resulted in difficulty in molecular replacement using the whole molecule as a search model. A solution was found by using only domains I and II, which together form a globular, less flexible structure. Domain V, Domain III and Domain IV were added sequentially afterwards. Co-crystallization or soaking EF-G with GDP, GTP and nonhydrolyzable GTP analogs was attempted. Soaking with GDP and the GTP analog GMPPNP gave only partial occupancy of GDP in our structure, so only the *apo* structure is presented here. Unsurprisingly, the overall structure of *S. aureus* EF-G displays five domains similar to *T. thermophilus* EF-G (Figure 7A).

In molecule B, part of the Switch I region was built as a short helix (Figure 8). The Switch I region is disordered in all other known EF-G structures. The P-loop structure is in agreement with that in the *apo T. thermophilus* EF-G structure (AEvarsson et al. 1994). Upon soaking with GDP, a peptide-flip conformation with partial occupancy was observed, similar to that in a previous EF-G-GDP structure from *T. thermophilus* (Al-Karadaghi et al. 1996).
The switch II region displays a distinct conformation compared to other available structures (Figure 7B), and none of them adopts the same conformation as when EF-G is bound to the 70S ribosome (Gao et al. 2009). The conserved bulky residue Phe88 (Phe90 in T. thermophilus) is exposed at the surface of EF-G when it is bound to the ribosome, where it participates in forming the binding pocket of FA. Only when EF-G is bound to the ribosome does the switch II region adopt its conformation for FA-binding.

Conformational space of EF-G

As mentioned in the introduction, there is a hinge-like movement between domain G/II and domain III, IV and V when EF-G is bound to the ribosome. Many T. thermophilus EF-G structures, when not bound to the ribosome, have only small overall conformational changes possibly because of constraints imposed by crystal contacts. In our structure, crystallization of EF-G from another species avoided this problem. Our structure has a significantly different domain arrangement. This new conformation demonstrates the conformational space of EF-G. The hinge-like movement in T. thermophilus EF-G structures results in displacement of the tip of domain IV by up to 8Å.
Upon binding to the ribosome, the EF-G molecule opens up resulting in a shift of the tip of domain IV, which protrudes to the A-site during translocation, of up to 27 Å in the same direction. Our structure shows a 25 Å displacement of the tip of domain IV, in the perpendicular direction. A comparison of the position of the tip of domain IV in all available EF-G structures based on superposition of domain G and II is shown in Figure 9. The radius of gyration for the S. aureus EF-G structure was calculated to be 30.6 Å, which is in agreement with previous data from small-angle X-ray scattering (Czworkowski & Moore 1997). In conclusion, EF-G may display higher flexibility in solution than previously thought, and our structure demonstrates the size of the conformational space of EF-G when not bound to the ribosome.

Figure 9. Conformational space of EF-G. A and B are superpositions of S. aureus EF-G (magenta) with T. thermophilus EF-G on (blue) and off (yellow) the ribosome based on domain G and II. The arrows indicate the direction of projection to the circle in C. C is the comparison of the positions of the tip of domain IV in the following known EF-G crystal and cryo-EM structures. 1, E.coli EF-G+ GMPPNP+ 70S (PDB 2om7 cryo-EM); 2, T. thermophilus EF-G+ 70S+ FA+ GDP (PDB2wri); 3, E.coli EF-G+ GDP+ 70S+ FA (PDB 1jqm, cryo-EM); 4, T. thermophilus EF-G T84A+ GMPPNP (PDB 2bv3); 5, T. thermophilus EF-G+GDP (PDB 1ktv); 6, T. thermophilus EF-G apo (PDB 1elo); 7, T. thermophilus EF-G T84A+ GDP (PDB 2bm0); 8, T. thermophilus EF-G G16V + GDP (PDB 2bm1); 9, T. thermophilus EF-G H573A+ GDP (1fnm); 10, S. aureus EF-G (PDB 2xex).

FA-resistance mutations

In total there are 42 point mutations reported to cause FA-resistance (Johanson et al. 1994; Nagaev et al. 2001; Norström et al. 2007). We mapped and analyzed all the mutations on our S. aureus EF-G structure. We classified the mutation sites into four groups: group A, residues involved in FA-binding pocket formation; group B, residues altering ribosome-EF-G interactions; group C, residues related to interdomain orientations in EF-G; group D, mutations decreasing stability of domain I and III of EF-G. Several of the mutations are likely to interfere with more than one parameter: e.g.,
EF-G conformation and stability are linked to FA- and ribosome-binding. A detailed interpretation of the mutations can be found in table 2 of paper I.

**Mysterious density – capture of free phosphate?**

Pre-steady-state kinetics suggested that EF-G binding to the ribosome leads to GTPase activation, after which EF-G is in a GDP-Pi state. In this state the inorganic phosphate (Pi) is not released (Wilden et al. 2006; Rodnina et al. 1997; Savelsbergh et al. 2003). As mentioned above, soaking in GDP, GTP, and a nonhydrolyzable GTP analog were tried, but not included in paper I. Soaked with either GDP or the GTP analog, only partial occupancy of GDP was observed. An interesting observation occurred with the GTP soaked structure. In this structure, GDP was observed with full occupancy, and noticeable local changes were found around the P-loop region. A clear piece of positive-difference density was observed in between a few charged residues. The density was larger than a water molecule and had a pyramid shape (Figure 10A). Attempts to fit a phosphate group in the density agreed well with the density although the binding pocket is not a typical phosphate binding site (Figure 10B)

*Figure 10.* Electron density maps of the ligand in the GTP soaked EF-G structure. A. Un-biased Fo-Fc map at a contour level of 3 σ showing strong positive density for GDP and an extra blob of density ~9 Å away from the GDP. B. 2Fo-Fc map at a contour level of 1 σ showing the electron density map for GDP, a free phosphate, and the surrounding residues.

**Paper II: A clever network to gain fitness: analysis of EF-G mutants**

In this study, we solved crystal structures of three *S. aureus* EF-G mutants – FA-resistant mutant F88L, FA-resistant but fitness compensated double mu-
tant F88L/M16I, and the M16I mutant, which was back mutated from the double mutant and interestingly displays the FA-hypersensitive phenotype. Structures were solved using molecular replacement with the wild type S. aureus EF-G Domains I and II as a search model, and the rest of the model was manually added, domain by domain, in COOT.

Fast kinetics assays show that the M16I mutant exhibits higher activity in ribosome complex formation, tripeptide formation, ribosome recycling and GTP hydrolysis, and somewhat higher binding affinity to GTP compared to the other mutants. The F88L mutant shows significantly slower translocation and recycling, and increased tRNA drop off. However, the activity of the double mutant F88L/M16I in both reactions was increased by three to four times compared to F88L and no tRNA drop off was observed. These data provide biochemical explanations for the fitness compensatory phenotype of the double mutant. (The kinetics of paper II was done by R. Koripella.)

To further reveal the mechanism of fitness compensation, careful analysis of our crystal structures was performed. Although all the mutant structures are ribosome-free structures, the alteration of interactions that are crucial for GTP hydrolysis, domain rearrangement and FA binding pocket were in agreement with the kinetics data.

Overall structures of the mutants
The mutant crystals belong to the same space group, P2₁, as the wild type crystals. However, the size of the unit cell for the mutant crystals is dramatically different from that of the wild type (WT) despite the fact that crystals were streak seeded with WT crystals. The overall conformation of all the mutants differs significantly from the WT structure in the sense of relative domain movement, and it also differs from previous T. thermophilus structures. The tip of domain IV is displaced by 6-9 Å in the mutants with respect to the WT.

Phe88 and Switch II region
As mentioned before, switch II is located at the interface of domains II, III and V. The conformation of switch II is very similar in all three mutant structures but is different from the WT structure (Figure 11). In the WT structure Phe88 contacts domain V, whereas in all the mutant structures, the switch II loop shifts about 5 Å, resulting in the side chain of residue 88 contacting domain III instead. In the mutants, residue 88 forms hydrophobic interactions with Leu456 and Ile460 in domain III. In F88L and the double mutant, the smaller side chain Leu88 brings domains III 1.6 Å closer to domain I. Asp87 forms a salt bridge to Arg659 in domain V in the wild type structure, but this salt bridge is disrupted in the mutant structures. Clearly Phe88 and its neighboring residues can trigger global conformational change
in EF-G. Our structures provide structural support for Phe88 playing a pivotal role in transmitting the conformational change of switch II to other domains of EF-G so that EF-G acquires the active conformation for translocation (Ticu et al. 2011).

**Figure 11.** Switch II region of *S. aureus* wild type EF-G (magenta), M16I mutant (blue), F88L mutant (green), and double mutant F88L/M16I (yellow).

**Figure 12.** M16 and the hydrophobic core of domain G. Vander Waals interactions are shown as dashed lines. Superposed structures are M16I (blue), F88L (red), double mutant F88L/M16I (green) and the WT structure (grey).

**Met16 and the hydrophobic core**

Met16 is located in a hydrophobic core formed by Leu, Val and Ala residues upstream of P-loop in domain I (Figure 12). This core involves residues from two helices in domain I, which are close to the switch II loop and pack against domain II and V. The residues forming the hydrophobic core are not conserved between species but seem to compensate for each other in size and shape. When Met16 is mutated to a smaller residue (Ile) in the mutants M16I and F88L/M16I, the distances of hydrophobic interactions to sur-
rounding residues in the core notably increased. This suggests a destabilizing effect of the hydrophobic core. The F88L mutant also showed small changes in the core compared to the WT, where the distances between Cα atoms of the hydrophobic residues increase (Figure 12). A loosening of the hydrophobic core in M16I seems to favor conformational dynamics in EF-G through effects on the domain I-II interface. Thus, Leu88 can substitute for Phe88 more efficiently in the double mutant F88L/M16I. Several other fitness compensatory mutants to F88L (Nagaev et al. 2001) seem to have similar effects on the hydrophobic core and the domain I-II interface. A loosening of the hydrophobic core may make EF-G more efficient in translocation and recycling thereby improving fitness.

Crucial salt bridges

Apart from the salt bridge Asp87-Arg659 mentioned above, a few more salt bridges were observed to be involved in inter-domain interaction. They could potentially be of functional importance. These salt bridges include Glu444-Arg483, Glu117-Arg154, Glu610-Arg636, Lys23-Thr82. It would be of great interest to test the effect of these salt bridges on EF-G function.

Paper III: The first structure of ANT (3”) - AadA

The structure

In this work, we solved the crystal structure of an aminoglycoside 3” adenyltransferase AadA from *Salmonella enterica*. AadA has 262 amino acids. Both native AadA and selenomethionine-substituted protein was produced and purified. The protein was purified as a monomer from gel filtration and the structure was solved using single-wavelength anomalous diffraction (SAD) phasing. Of 262 amino acids in the sequence, 260 were ordered in the structure. Although co-crystallization and soaking with the substrates streptomycin/spectinomycin and the nonhydrolyzable analogue of the cofactor ATP was attempted, no density for the ligands was observed. The overall structure consists of two domains. An overview of the structure and the topology diagram is shown in Figure 13. The N-terminal domain has a central β-sheet surrounded by α helices, a typical fold for nucleotide recognition. This domain is classified as a nucleotidyltransferase domain. The C-terminal domain is an α-helical domain, which falls into the category of up-and-down helical bundles.
**Figure 13.** Overall structure of AadA and topology diagram for the N- and the C-terminal domains. The structure is rainbow colored from the N terminus in blue to the C terminus in red. The same color scheme is used in the topology diagrams. The numbering stands for the start and end residues of the secondary structure.

**Figure 14.** Surface properties of AadA (A) and KNTase (B). The upper row shows surface electrostatic potential. The color spectrum ranges from deep red (-7kT) to deep blue (+7kT). The lower row shows the surface conservation. The color spectrum ranges from magenta (highest conservation) to cyan (lowest conservation). The second subunit in the KNTase homodimer is shown as a yellow cartoon.

Conserved residues in AadA form a pocket for substrate and cofactor binding

A search for similar structures in the PDB was performed using the Dali server (Holm & Rosenström 2010). In a Dali search with the entire AadA molecule, the top hit was the kanamycin nucleotidyltransferase (KNTase) from *S. aureus* with a Z-score of 9.5 (pdb 1kny, (Pedersen et al. 1995)). Sur-
face conservation of AadA and KNTase was analyzed using the ConSurf server (Ashkenazy et al. 2010; Celniker et al. 2013), where homologue sequences were obtained by a BLAST search in the UniRef90 database. Strikingly, highly conserved surface residues of the AadA molecule are clustered in the inter-domain region, forming a pocket, whereas the outer surface of the molecule shows little conservation (Figure 14A). Surface electrostatic potential of AadA shows that the pocket formed by the conserved residues is highly negatively charged, whereas the other parts of the molecule are with a lower level of negative charge or uncharged (Figure 14A). This highly conserved and negatively charged pocket is obviously the binding pocket for the positively charged substrates and the cofactor ATP.

AadA most likely functions as a monomer

KNTase function as a homodimer, whereas AadA is monomeric both from purification and crystallization. In KNTase, the conserved residues are involved in both ATP- and substrate- binding and dimerization (Figure 14B). Among these conserved residues, the ones involved in ligand binding display a highly negatively charged surface, whereas the ones involved in dimerization are uncharged. Although AadA is potentially able to open up and expose the conserved pocket, the highly negatively charged surface is unlikely to form a homodimer. Also, the highly conserved helix η2 preceding α4 in AadA (which has no equivalent in KNTase, Figure 15) would clash with helix α6 in the C-terminal domain of KNTase when binding to the beta-sheet of the N-terminal domain. Therefore, AadA most likely functions as a monomer.

Candidate residues for catalysis and substrate binding

In the Dali search with the entire AadA molecule, a lincosamide nucleotidyltransferase LinB from Enterococcus faecium (pdb 3jz0 (Morar et al. 2009)) also shows similarity to AadA with a Z-score of 6.7. Both KNTase and LinB share low sequence identity with AadA (14% and 9%), but they display similar folds to AadA although both are homodimers. Moreover, both the structures of KNTase and LinB have ligands bound. Therefore, comparisons between AadA and these two structures were performed. Careful domain-by-domain superpositioning shows that the central β-sheet and helix α1 superpose well in all three structures when superposing the N-terminal domain; α8 and α11 also superpose well in the C-terminal domain. The nucleotide-binding pocket is conserved in all three enzymes; however, AadA and LinB show little similarity in the other regions. Therefore structure-based sequence alignment was performed only between AadA and KNTase (Figure 15). The majority of the secondary elements overlap between these two structures.
Figure 15. Structure-based sequence alignment. On the top and bottom of the sequences, the secondary structures of AadA and KNTase (PDB 1kny (Pedersen et al. 1995)) are shown. Identical residues are in white on grey background and conservative substitutions are squared. The figure was produced using the EsPript Server (Gouet et al. 2003).

The superpositioning of AadA and KNTase based on separate domains is shown in Figure 16A and B. Among the strictly conserved residues in AadA, Ser36, Ser46, Asp47, Asp49, Glu87, Thr89, Trp112, Asp182, Arg192 and Lys205 are surface exposed and point towards the inter-domain space, suggesting that they are involved in substrate binding and/or catalysis (Figure 16C). Of these, Ser36, Ser46, Asp47 and Asp49 have equivalent residues in KNTase and LinB, which coordinate with ATP phosphates (Figure 17). These residues most likely have the same role in AadA. Glu87’s equivalent residue Glu76 of KNTase is in close proximity to the substrate kanamycin without clear role, but its equivalent residue in LinB, Glu89, has been proposed to be the catalytic residue; this has been confirmed by mutagenesis (Morar et al. 2009). Glu145 in KNTase has been proposed to be the catalytic residue (Pedersen et al. 1995), but without any equivalent in AadA. However, in the overlay based on the N-terminal domain, the carboxylic oxygens of Glu145 in KNTase are within 5 Å of the ones of Glu87 in AadA. This suggests that Glu87 in AadA may have the same role as Glu89 in LinB and Glu145 in KNTase, being the catalytic residue. All the other conserved residues at the domain interface are without any equivalent residues in either KNTase or
LinB. Therefore these residues probably are responsible for substrate specificity.

*Figure 16.* Superpositioning of AadA and KNTase based on the N-terminal domain (A) with an RMSD of 1.59 Å for 81 Cα atoms and the C-terminal domain (B) with an RMSD of 2.23 Å for 57 Cα atoms. (C) Close-up view of conserved surface-exposed residues in AadA. ATP and kanamycin from KNTase are shown in light cyan and light blue when superimposed based on the N-terminal domain.

*Figure 17.* Structure of the nucleotide binding site of LinB (A), KNTase (B) and AadA (C). (A), (B) and (C) are from the same view. Superpositions are based on the N-terminal domain. Mg$^{2+}$ ions are in green spheres. 2Fo-Fc map for conserved residues in AadA is shown in C, countered at 1σ (0.23 electrons/Å$^3$). Their equivalent residues in LinB and KNTase are shown in (A) and (B).

The AadA structure is in a “closed” conformation compared to the KNTase. There are extensive inter-domain interactions in the AadA structure. Notably, Ser36 forms a hydrogen bond with Asp206, and Asp 47 and Asp49 form salt bridges with Lys205. Therefore, the ATP-binding site in our structure is blocked by inter-domain interactions. In other words, AadA probably needs to open up so as to accommodate ATP- and substrate-binding.
Paper IV: Alternative anticodon-loop structure to induce -1 frameshifting

As introduced previously, normal tRNAs (tRNA_{Ser} and tRNA_{Thr}) can cause -1 frameshifting at GCA alanine codon and CCA/CCG proline codons. A schematic view of possible codon-anticodon interactions between the frameshift tRNAs and the mRNAs is shown in Figure 18. As seen in this figure, the standard binding of tRNA_{Ser} and tRNA_{Thr} would only allow one base pair to be formed. Therefore, a doublet-decoding model has been proposed (see introduction). The main question we would like to answer in this study is if the doublet decoding at the A-site is true, and in which site the committed event of the -1 frameshifting happens. Initially, we aimed to combine biochemistry and crystallography to understand the mechanism of -1 frameshifting. Crystals of the 70S ribosome in complex with mRNA, P-site tRNA and A-site tRNA/ASL were obtained and optimized, but they were not of sufficient quality to solve the structures. Binding affinity was measured for A-site tRNAs to mRNAs while the P-site was occupied. Only the ASL part of the A-site tRNA was used for measuring the binding affinity because the full tRNA gave high background noise in the filter binding assay (possibly due to non-specific binding of the other parts of the tRNA).

Four mRNAs and six ASLs were used in our assay. The mRNAs differ only at the A-site codons. They are the frameshifting sites GCA alanine codon, CCA and CCG proline codons, and also GCG codon designed to test the effect of U33 in the ASLs (see below). In addition to cognate and frameshift-inducing ASLs, an ASL_{Ser}^{U36C} was also designed with U36 mutated to C; this was to test the effect of U36 (see below). Nine cognate or near-cognate combinations were measured. The variants and the measured dissociation constants are summarized in Tables 1 and 2.

Table 1. Sequences of RNAs used in the nitrocellulose filter-binding assay.

<table>
<thead>
<tr>
<th>RNA Oligos</th>
<th>Sequences</th>
</tr>
</thead>
<tbody>
<tr>
<td>ASL^{Ala}_{1}</td>
<td>CCUGCUUUGCACGCAGG</td>
</tr>
<tr>
<td>ASL_{Ser}^{3}</td>
<td>CUCCCCUGCUAAGGGAG</td>
</tr>
<tr>
<td>ASL_{Ser}^{U36C}</td>
<td>CUCCCCUGCGAAGGGAG</td>
</tr>
<tr>
<td>ASL_{Pro}^{1}</td>
<td>ACUGGCUUGGACCAAGU</td>
</tr>
<tr>
<td>ASL_{Pro}^{2}</td>
<td>CUUCGUUGGACCAAGG</td>
</tr>
<tr>
<td>ASL_{Thr}^{3}</td>
<td>CACCCUUUGUAAGGGUG</td>
</tr>
<tr>
<td>mRNA_{GCA}</td>
<td>GGCAAGGAGGUAAGGUGG</td>
</tr>
<tr>
<td>mRNA_{CCA}</td>
<td>GGGAAGGAGGUAAGGCGA</td>
</tr>
<tr>
<td>mRNA_{CCG}</td>
<td>GGGAAGGAGGUAAGGCGA</td>
</tr>
<tr>
<td>mRNA_{GCG}</td>
<td>GGCAAGGAGGUAAGGCGA</td>
</tr>
</tbody>
</table>
Figure 18. Schematic view of possible codon-anticodon interactions. The first column shows standard cognate codon-anticodon interactions; the second column shows that only one base pair can be formed if the frameshift tRNA interacts with the mRNA in the standard way; the third column shows doublet base pairing.

Table 2. Binding affinity of ASLs to various A-site codons.

<table>
<thead>
<tr>
<th>A-site ASL</th>
<th>Anticodon</th>
<th>mRNA</th>
<th>$K_d$ (µM)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ala1</td>
<td>UGC</td>
<td>GCA</td>
<td>0.5 ± 0.04</td>
</tr>
<tr>
<td>Ser3</td>
<td>GCU</td>
<td>GCA</td>
<td>4 ± 0.1</td>
</tr>
<tr>
<td>SerU36C</td>
<td>GCC</td>
<td>GCA</td>
<td>154</td>
</tr>
<tr>
<td>Ala1</td>
<td>UGC</td>
<td>GCG</td>
<td>3 ± 0.8</td>
</tr>
<tr>
<td>Ser3</td>
<td>GCU</td>
<td>GCG</td>
<td>2 ± 0.2</td>
</tr>
<tr>
<td>Pro1</td>
<td>UGG</td>
<td>CCA</td>
<td>0.8 ± 0.0</td>
</tr>
<tr>
<td>Thr3</td>
<td>GGU</td>
<td>CCA</td>
<td>13.8 ± 2.1</td>
</tr>
<tr>
<td>Pro2</td>
<td>CGG</td>
<td>CCG</td>
<td>3.6 ± 0.4</td>
</tr>
<tr>
<td>Thr3</td>
<td>GGU</td>
<td>CCG</td>
<td>---</td>
</tr>
</tbody>
</table>
At the GCA alanine codon, cognate binding with ASL$_{\text{Ala}}^1$ gives a dissociation constant of 0.5 µM, in agreement with cognate binding affinity measured for ASL$_{\text{Phe}}^\text{GAA}$ (GAA is the anticodon) binding to its cognate codon (Ogle et al. 2002). ASL$_{\text{Ser}}^3$ displays 10-fold lower binding affinity to the same codon, yet it is still 50-fold tighter than the binding affinity measured for ASL$_{\text{Leu}}^{\text{GAG}}$ to a non-cognate UUC codon (~200 µM) (Ogle et al. 2002). The differences in binding affinity seem to indicate that the ASL$_{\text{Ser}}^3$ forms two base pairs with the mRNA. We also wanted to test in our system if the identity of position 36 was crucial for this -1 frameshifting event (Bruce et al. 1986). ASL$_{\text{Ser}}^{\text{U36C}}$, which carries a mutation U36C in ASL$_{\text{Ser}}^3$, was determined to have a $K_d$ value in the same latitude as a non-cognate binding. This agrees with Bruce et al.’s observation. Also, this excludes the possibility of U36 being involved in base pairing. The U36C mutation would result in tighter binding if C base paired with the G before GCA alanine codon.

At the CCA proline codon, the $K_d$ for cognate binding was 0.8 µM, and the frameshift-inducing ASL$_{\text{Thr}}^3$ bound to the CCA codon with a $K_d$ of 13.8 µM. Interestingly, at the CCG proline codon, we observed a much lower cognate binding and no binding was detected for ASL$_{\text{Thr}}^3$. The third base of the A-site codon seems to affect the frameshifting event, and base pairing between U33 on the ASLs and the third base would be a simple explanation for this. If this is valid, a “shifted anticodon” model would be appropriate, where U33 forms the third base pair with the last base of the A-site codon. To test our hypothesis, the binding affinities of ASL$_{\text{Ala}}^1$ and ASL$_{\text{Ser}}^3$ to a GCG codon altered from the GCA alanine codon were measured. If U33 participates in base pairing, altering the third base from A to G would result in a decrease in binding affinity of ASL$_{\text{Ser}}^3$ by a factor of 10 (Ogle et al. 2002). However, no decrease in binding affinity was detected. Rather, ASL$_{\text{Ser}}^3$ bound to both codons with similar affinity.

The above results strongly indicate that doublet decoding at the A-site is established in these cases of -1 frameshifting. The committed step for these frameshifting events happens at the A-site, although the real shift on the mRNA relative to the ribosome is probably established at the P-site after a two-base translocation. Our results clearly support the doublet-decoding model and shed light on the mechanism of the particular frameshift events induced by tRNA$_{\text{Ser}}^\text{G}$ and tRNA$_{\text{Thr}}^\text{G}$. 

51
Concluding remarks

The studies in this thesis provide insights for understanding of antibiotic resistance and reading-frame maintenance.

Elucidation of the crystal structure of wild type *S. aureus* EF-G allowed us to map all known FA-resistant mutations to the structure, as well as explain the roles of the mutations. Structures of the FA-resistant mutant F88L, the FA-hypersensitive mutant M16I, and the fitness-compensated double mutant F88L/M16I made it possible to provide a structural explanation of the mechanism-of-fitness compensation, which involves loosening of a hydrophobic core in the catalytic domain.

The structure of the aminoglycoside adenyltransferase – AadA is the first structure from the ANT (3”) enzyme class. Comparison of our structure with the KNTase structure from the same enzyme family with low sequence similarity showed that AadA probably functions as a monomer, and candidates for catalytic residues and mutagenesis study were proposed. The ATP binding sites seem to be conserved in these two enzymes. However, the other conserved residues in AadA that are not present in KNTase seem to be responsible for substrate binding specificity.

Measurements of binding affinities of various tRNAs to the ribosomal A-site within different mRNA contexts support the doublet-decoding model for the -1 frameshifting events caused by normal tRNAs. Unlike most framshifting events, which happen at the P-site (Atkins & Björk 2009), these cases of frameshifting seem to adopt a committed step at the A-site by doublet base pairing. Our experiments successfully demonstrated these frameshifting events in a small system with unmodified ASLs of the tRNAs. This provides evidence for the doublet-decoding model.

To sum up, it is my hope that these studies have contributed their part to the overall picture of translation elongation process.
Future perspectives

Some interesting points derived from each project will be further investigated.

First of all, regarding the EF-G structures and FA-resistance, the mysterious density speculated to be a phosphate in the GTP-soaked crystals would be of great interest to be confirmed. If it is a free phosphate, the structure would represent a transitional state after GTP hydrolysis but before Pi release. Also, effects of the proposed crucial salt bridges on EF-G function could be tested by mutagenesis.

Second, in the AadA structure, the conserved residues that we believe to be of functional importance in substrate specificity will be mutated and tested for activity with different substrates. A rescreening for co-crystallization of AadA with the substrates and the cofactor will then be attempted to produce crystals of the complex. Furthermore, docking of the substrates using our structure and MD simulations will be carried out. The binding order and affinities of substrate and cofactor to AadA would also be interesting to be determined. We would also like to find out why this enzyme can catalyze the enzymatic reaction on substrates with so different shapes.

Last but not least, simulations of the codon-anticodon interaction using our -1 frameshifting model will be implemented to predict the conformation of tRNAs in the unusual binding situations, which will help us to further understand the mechanism of the -1 frameshifting events.
Proteinsyntes är en fundamental process i alla celler. Proteinsyntes utförs av stora molekylära maskiner som kallas ribosomer. En ribosom kan översätta genetiskt material till proteiner, via molekylen mRNA. Vid avläsningen av mRNA rekryteras enskilda aminosyror med hjälp av molekylen tRNA. Dessa länkas sedan ihop och veckas till ett funktionellt protein. Den här processen kallas translation.


Dessa tre aspekter är oberoende av varandra, men är alla relaterade till fel som inträffar vid förlängningssteget under proteinsyntesen. Målet med det
arbete som presenteras i denna avhandling är att ge insikter om detaljerna bakom antibiotikaresistens och de fel som kan uppstå under translationsprocessen.

Med hjälp av kristallstrukturer av EF-G från bakterien Staphylococcus aureus kunde vi kartlägga alla kända mutationer som orsakar fusidinsyrraresistens. Strukturer av den fusidinsyra-resistenta mutanten F88L, en fusidinsyra-överkänslig mutant M16I och en F88L/M16I dubbelmutant som kompenserar för F88L-mutantens lägre tillväxt, gav strukturella förklaringar till resistens och tillväxtkompensation.

Den andra delen beskriver kristallstrukturer av aminoglykosidadenyltransferas, vilket är den första strukturen från enzymklassen ANT (3”). En jämförelse av AadA med det avlägset besläktade ANT (3’)-enzymet KNTase indikerar att AadA fungerar som en monomer. Både AadA och KNTase binder ATP och bindningsfickan är väldigt lika mellan dessa två enzymer. Ett antal aminosyror i närheten av ATP-fickan, vilka är konserverade bland AadA-enzyme men inte bland KNTaser, är sannolikt ansvariga för skillnaden i substratspecificitet mellan de två proteinerna.

Mätningar av bindningsaffiniteterna för tRNA till olika mRNA-kodon i ribosomens A-position tyder på att -1-skift i ribosomens läsram som orsakas av normala tRNA sker genom avkodning av två mRNA-baser i stället för det normala tre. Våra experiment i ett väldefinierat system med antikodonfragment av tRNA tyder på att det är speciella tRNA-structurer i ribosomens A-position som får detta att ske.

Sammanfattningvis leder dessa studier till ny kunskap om några aspekter av den bakteriella proteinyntesen och ökar förståelsen av mekanismer bakom antibiotikaresistens och förändringar i ribosomens läsram.
Acknowledgements

I would like to express my sincere gratitude to my supervisor Maria Selmer. It has been truly delightful working with you during the past six years. Thank you for accepting me as a student. You have always been there for me whenever I got stuck; you have always been nice and supportive whenever I feel frustrated. Thank you for helping me transform my poorly written drafts into publications, my ordinary poster and presentation into prize-winning ones, and for making me feel proud of myself sometimes! Having you as my supervisor is one of the luckiest things that have happened to me in my life!

I also would like to thank my co-supervisor Suparna Sanyal, for nice collaborations in the EF-G and frameshifting projects, interesting discussions, and help for my understanding of kinetics. This thesis would not have been possible without collaborations with a lot of other people. My sincere thanks to: Ravi Koripella for crystallizing EF-G and for biochemical studies on the EF-G mutants; Dan I. Andersson and Joakim Näsvall for good collaboration on the AadA project; Samuel Flores and Alex Tek for their efforts in the frameshifting project. Also, thanks to Terese Bergfors for critical reading of this thesis and Anna Jansson for the Swedish summary.

Thank you all the other former and present members in the Selmer group: Nina for help in uncountable small things in the lab, Avinash for good discussions and useful suggestions, Shiying for help in RNA production and labeling, Xiaohu for being a good office mate, Annika for good discussions at the group meetings, ChaSan for help in analyzing my EF-G mutant structures and good times, Maike for cookies and good times… also, Sandesh, Lisa, Jendrik, Zhaolin, Zhe, Jeniffer… and everybody for all the positive atmosphere.

I have learnt a lot during the past years and not only in the lab. I would like to thank Torsten for the protein-engineering course, Alwyn and Lars for the crystallography course, Tex for the crystallization course and scientific writing course, and all the language tips. Måns, for blue Monday meeting discussions and good suggestions. Thanks to everybody in X-ray for Friday seminars where I always learned something new. Thanks to Maria, Evalena and Daniel for my first ESRF trip and to Anna, Christofer, Xiaohu, Lu, Sanjee, Wimal… for a lot more trips together. I learned a lot from everybody.
I am very grateful that I have so many lovely friends and colleagues around during my PhD study. You make my life so colorful! Thanks to: Shiying, for being a big sister, you are a family to me. Anna - “sötnos”, for Swedish lessons, cooking exchange, raw food diet, yoga, tjekväll, fredagsmys and a lot more good times, you are my very own sötnos! Christofer, for all the good times- cooking, eating, drinking, partying… Annette, for always being nice and sweet! Nisse, for being a great office mate during the first few years of my study. Miao, Xiaohu, Wangshu, Li, Jie and Lu, for making me feel that I am not too far away from home. Tex, for cooking sessions. Henrik K, Lotta, for beer clubs, afterworks… Christoffer, Masoud, for good fika times. Prune, for all the lunch climblings. Karin and Nina, for pottery evenings. Yafei, for good times at your place. Akiko, for fika and making moon cakes. Erling and Christer, for fixing computer problems. Adrian, Agata, Ana Laura, Anatoly, Anna L, Anton, Chandu, Cia, Dee, Glareh, Henrik I, Jerry, Johan S, Johan Å, Kristin, Lars B, Lena, Magnus, Majid, Mats, Micke, Sanjee, Stefan, Wimal, Yasmin and all the others in Xray, ICM, RiboCore and URRC for making here such a nice place to work at.

Special thanks to Maria, Elvira, Tex, Anna, Patrik, Christofer for coming all the way from Sweden to Taiyuan for our wedding and for joining us to a mini-honeymoon 😊. It was our great pleasure to have you at our home country!

I would also like to thank some of my close friends from college: Hongqian, Yin, Yingying, Lvlv, for always being ready to talk about anything. And some dear friends I have known since almost the first day I came to Sweden: Loulou, Siding, Vivien, Yuening and Geng, for being my special ladies all these years. Also, to my friends during my living in Flogsta and ever since: Freyja, Malin, Maja, Kiyoko… thank you for all the good times. Also, thanks to my classmates and friends from the MBB master program: Kaweng, Chuan, Xiang, Feifei, Jun, Ren, Chengxi, Lin…for fighting together for the PhD… and thanks to all my Chinese friends in Sweden for all the care and company.

在此我要感谢我在国内的亲朋好友，多年来的支持和鼓励！
感谢我的公公婆婆，谢谢你们对我的支持和厚爱！
感谢我亲爱的父母，谢谢你们对女儿一直以来无微不至的关怀，有你们这样的父母是我一生的骄傲！

Shun, thank you for showing up and staying in my life! I don’t know how many times I wished we had met earlier. You make me happier than I ever thought I could be and I know you feel the same way. Life with you has been so wonderful and I’m so looking forward to the rest of our lives together!


Schmeing, T.M. et al., 2009. The crystal structure of the ribosome bound to EF-Tu and aminoaoyl-tRNA. Science, 326(5953), pp.688–94.


A doctoral dissertation from the Faculty of Science and Technology, Uppsala University, is usually a summary of a number of papers. A few copies of the complete dissertation are kept at major Swedish research libraries, while the summary alone is distributed internationally through the series Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Science and Technology.

Distribution: publications.uu.se
urn:nbn:se:uu:diva-205131