Chapter 1

Introduction

Translation of the genetic code, namely protein synthesis, is one of the most complicated biological processes, involving literally hundreds of specific macromolecules. Not the least of this complexity is the structure of ribosome itself. The ribosome is a large ribonucleoprotein complex that is responsible for the protein synthesis in all living cells. The ribosome was discovered by the cytologists Palade and Siekevitz in the mid-1950s, and by 1960 it was apparent that it catalyzes protein synthesis (Tissieres et al., 1974). Although a minimal model for protein synthesis was put forward in the early 1960s (Watson, 1964), a detailed description of the mechanism of translation continues to present challenges to ribosomologists. However, the recent arrival of crystal structures of the 50S and 30S ribosomal subunits and the intact 70S ribosome opened a new phase in our understanding of protein synthesis. To identify and fill the gaps in our current knowledge of the mechanisms involved in translation is now better possible with the correlation of structural and biochemical data. Especially, the relative movements of its molecular components can only be revealed through a combination of structural and functional analysis, which leads the main concern for this thesis, namely, the movement of tRNAs on the ribosome during the elongation cycle of translation.

In the following part of the introduction, general information and recent advances about ribosome, tRNA, and translation process will be discussed in order to shed light on explaining the aim of this research.

1.1 Overview of the Ribosome

The structural core and fundamental mechanisms of the ribosome is conserved in all organisms. Depending on their origin, we can distinguish five

different classes of ribosomes (Subramanian, 1985). These are the 70S ribosomes of eubacteria, the 80S ribosomes present in the cytoplasm of all eukaryotes, the archaebacterial ribosomes which display characteristics of both the 70S and 80S type, the chloroplast ribosomes present in plants and algae and finally the mitochondrial ribosomes found in all eukaryotes.

The bacterial ribosome, which has been the most extensively studied, has a mass of approximately 2.4-2.8 MDa, a relative sedimentation rate of 70S and a diameter of 200-250 Å. As all the other classes of ribosomes, it is built up of two independent and unequal subunits that associate upon the initiation of protein biosynthesis (Figure 1.1A). Each subunit is a ribonucleoprotein particle with one-third of the mass consisting of protein and the other two-thirds of ribosomal RNA (rRNA). The small subunit (30S) contains 16S rRNA (~1,500 nucleotides) and about 20 different proteins. The large subunit (50S) contains two RNAs, namely 5S rRNA (120 nucleotides) and 23S rRNA (~2,900 nucleotides), and 36 proteins (33 different species).

The genetic information is transported from the genome, via mediator molecules called messenger RNAs (mRNAs), to the ribosome. During the protein synthesis, the ribosome deciphers the nucleotide message using transfer RNAs (tRNAs), which translate the triplet code (a codon) of the mRNA into amino acids, the building blocks of proteins. Reading of each codon by the ribosome results in the incorporation of one amino acid into a gradually lengthening the protein chain. Once the ribosome reaches a stop codon on the mRNA, translation stops, the ribosomal subunits separate and detach from the mRNA and the completed protein is released.

Although interaction between the large and small subunit of the ribosome is a fundamental property of the translation, each subunit has distinct functions. The small subunit, 30S, is the site of the codon-anticodon interaction (decoding site) between the mRNA and the tRNA substrates (Figure 1.1B). Thus, it is responsible for translational fidelity. The large subunit, 50S, catalyses peptide bond formation between the nascent polypeptide chain and the incoming aminoacylated tRNA. One of the special features of the 50S

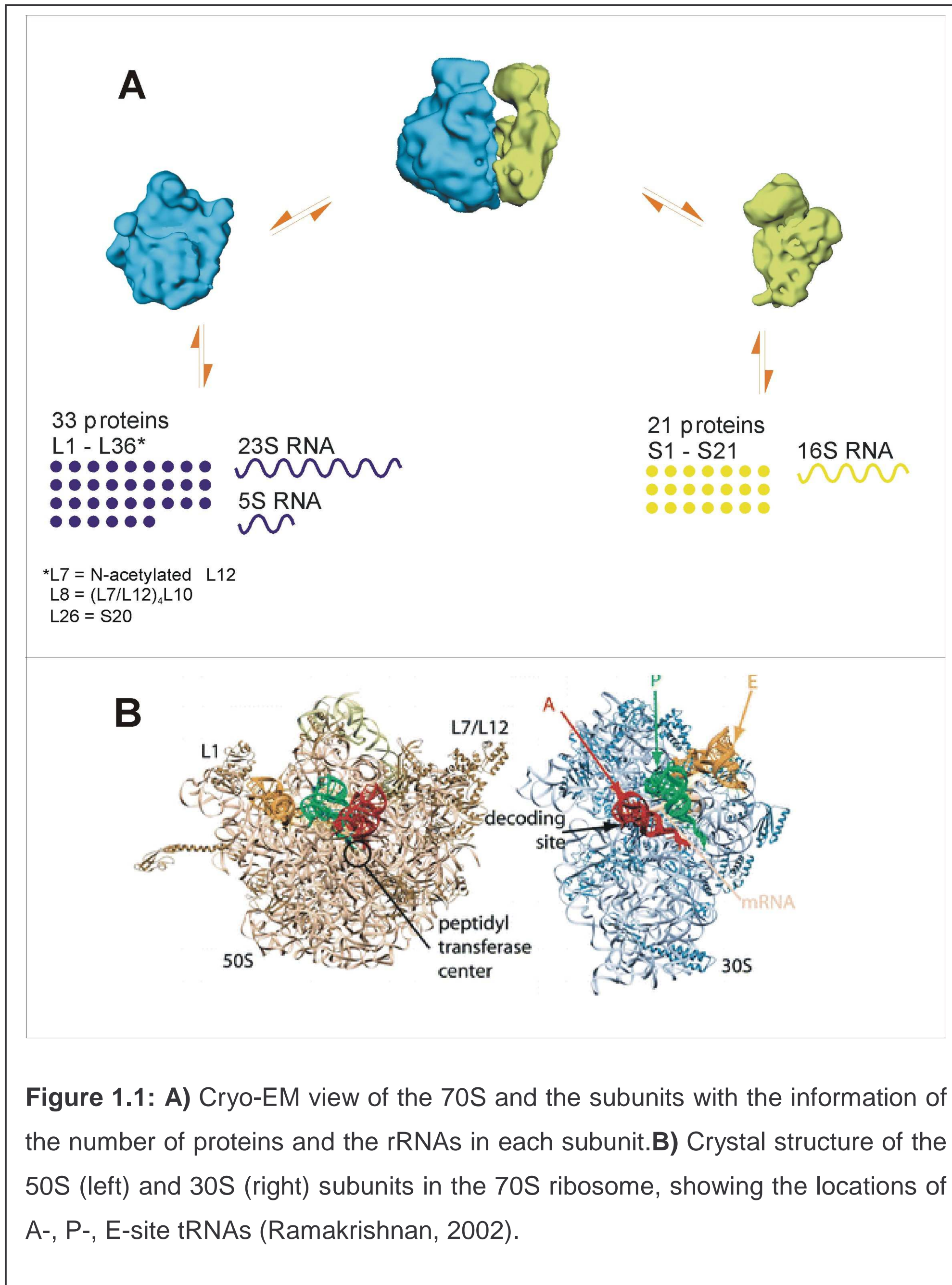

Figure 1.1: A) Cryo-EM view of the 70S and the subunits with the information of the number of proteins and the rRNAs in each subunit.**B)** Crystal structure of the 50S (left) and 30S (right) subunits in the 70S ribosome, showing the locations of A-, P-, E-site tRNAs (Ramakrishnan, 2002).

subunit is the tunnel. It runs from the peptidyl-transferase (PTF) center at the foot of the central protuberance up to the base at the cytoplasmic side of the subunit with a length of about 100 Å and a width of 10 to 20 Å (Ban et al., 2000; Stark et al., 1995). Additionally, 50S subunit has a factor-binding center and all of the G-protein factors involved in protein synthesis interact with it during at least part of their duty cycles. Both subunits are involved in translocating the mRNA by one codon in each cycle.

Structurally, the 50S subunit is spherical with three almost cylindrical extensions (Figure 1.1B). These extensions are called L1 protuberance, the central protuberance and the L7/L12 stalk. A striking difference between the two subunits has to do with the relationship between the secondary structures of their RNAs and their overall morphology. The six secondary structure domains of 23S rRNA are intricately interwoven in the 50S subunit to form a monolithic structure. On the other hand, the 30S subunit is divided into three domains (head, body, and platform). Each of these domains contains one of the principal secondary structure domains of 16S rRNA: The 5' major domain represents the body, the central domain the platform and the 3' major domain the head of the small subunit. The 3' minor domain of the 16S rRNA forms an extended helix (h44 in *E. coli* helix numbering) and runs down the long axis of the 30S subunit surface that interacts with the 50S subunit. All four domains of the 30S particle join at a narrow neck region. The two active sites (the decoding center and the peptidyl transferase center) face each other across the subunit interface and are functionally linked by both the two ends of A-site tRNA and a prominent intersubunit bridge (bridge 2B). The architectural difference between the subunits may reflect a greater functional need for flexibility on the part of the small subunit and, not necessarily alternative, a different evolutionary age of the subunits, where the large one is probably the older one (Sardesai et al., 1999).

Cryo-EM and X-ray maps of the 70S ribosome show a number of highly conserved bridges connecting the subunits (Cate et al., 1999; Frank et al., 1995; Gabashvili et al., 2000). The intersubunit interface of both subunits, especially the part that binds mRNA and tRNAs, is largely free of protein.

1.2 Overview of the tRNA

The amino acids are brought to the ribosome by tRNA molecules. The ribosome contains three tRNA binding sites, namely A, P, and E sites, (Burkhardt et al., 1998; Graifer et al., 1992; Grajevskaja et al., 1982;

Rheinberger et al., 1981; Saruyama and Nierhaus, 1986; Triana-Alonso et al., 1995) (Figure 1.1B). The A site contains the ribosomal decoding center that selects the cognate aminoacyl-tRNA during protein synthesis (A for **a**minoacyl-tRNA). The P site carries the peptidyl-tRNA before peptide bond formation (P for **p**eptidyl-tRNA), and the E site is specific for deacylated tRNA (E for **e**xit). During the protein synthesis, each tRNA passes through the ribosomal binding sites in the sequence of A → P → E.

Compared to the ribosome, tRNA molecules are relatively small. Every cell contains a population of tRNAs that differ in sequence (75-93 nucleotides), but have similar relative molecular mass (25-30 KDa), similar cloverleaf secondary structures, and the same L-shaped ternary structure (Figure 1.2). Since the anticodon of a tRNA determines its amino acid specificity and thus its role during protein synthesis, a "tRNA species" is defined solely by its anticodon. Concerning this criteria the *E. coli* cell contains 41 different tRNA species (Ikemura, 1981). The first three-dimensional structure of tRNA was solved for yeast $tRNA^{Phe}$ roughly 30 years ago (Kim et al., 1974; Robertus et al., 1974). The cloverleaf structure with self-complementary bases makes four distinct helical segments and three loops. The 3' ends of all tRNAs end in the single-stranded sequence $N^{73}CCA_{OH}$, where the free 2' or 3'-hydroxyl groups on the terminal adenosine contain the amino acid attachment site. The parts of the cloverleaf structure are the acceptor stem, the dihydrouridine (D) stem-loop, the anticodon stem-loop, and the TψC stem-loop. This structure is folded in three dimensions into an L-shaped that consist of two domains. One domain is formed by coaxial stacking of the TψC stem-loop onto the acceptor stem and that terminates at the 3'-end, while the other results from stacking the anticodon stem onto the D stem (Steer and Schimmel, 1999). At the distal end of the longer arm of the L there is a three-base, anticodon sequence in every tRNA that is complementary to one of the mRNA base triplets that encodes a specific amino acid. At the distal end of the short arm of the L of every tRNA is a 3' terminal CCA sequence to which the amino acid specified by the anticodon is attached. Thus, the two domains segregate distinct functions of the tRNA and thus, aminoacylated-tRNA molecules carry two essential functions. First, they bring the amino acids in an activated form to

the ribosome, because the amino acids are not direct substrates for the ribosome and are bound via ester bonds to the 3'-CCA ends of the tRNAs. Second, the tRNAs link the information coded in the nucleotide sequence of the mRNA into the amino acid sequence of the growing peptide. The distance between the tips of the L-shape is approximately 75 Å. Several lines of evidence suggests that the two domains of tRNAs arose independently (Chihade and Schimmel, 1999). Significantly, the two domains of tRNA

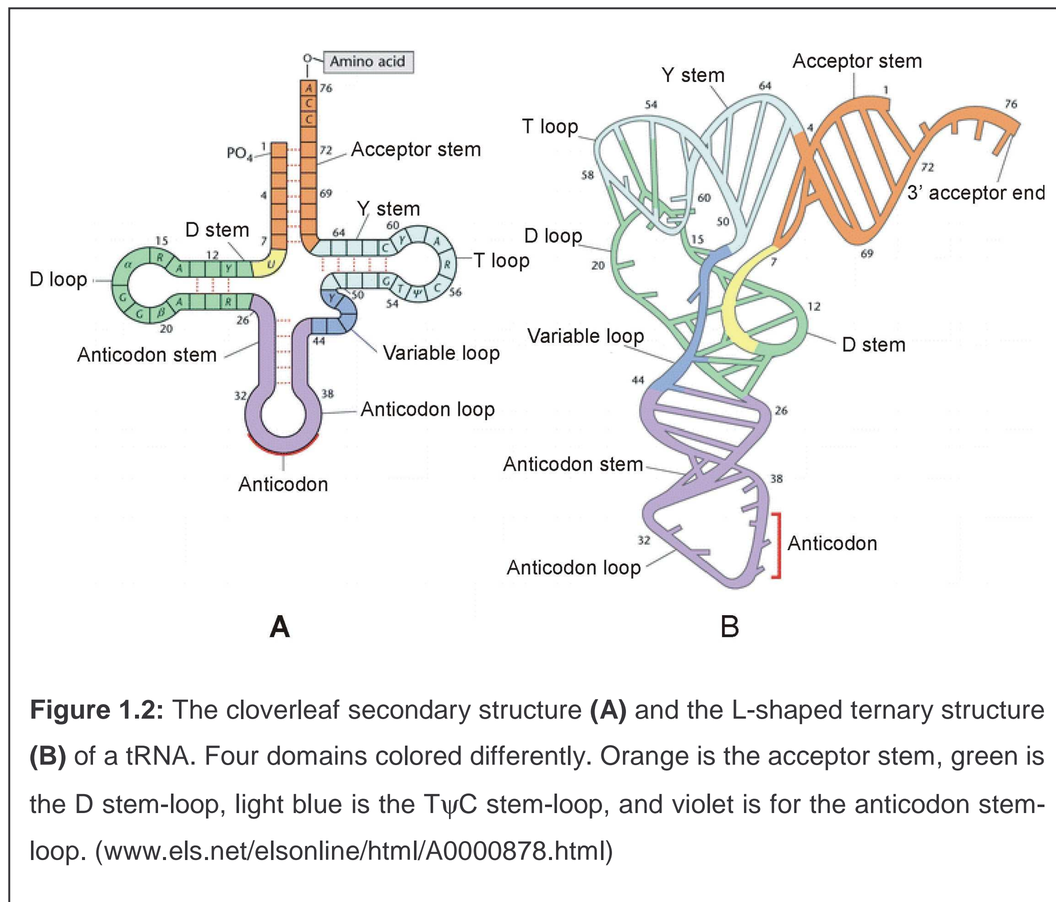

Figure 1.2: The cloverleaf secondary structure **(A)** and the L-shaped ternary structure **(B)** of a tRNA. Four domains colored differently. Orange is the acceptor stem, green is the D stem-loop, light blue is the TψC stem-loop, and violet is for the anticodon stem-loop. (www.els.net/elsonline/html/A0000878.html)

interact with separate domains of the ribosome. These observations raise the possibility that the short arm of the L shaped tRNAs and the 50S ribosomal subunit are progenitors of the full tRNA and 70S ribosome, respectively (Sardesai et al., 1999).

Since tRNAs move through functional sites on the ribosome, localization of tRNA-related functional centers such as peptidyl transferase center or the decoding center (Figure 1.1B) on the ribosome have always been an important issue in the translational field. Additionally, one of the most striking

questions in protein synthesis is how to move large tRNA molecules with a size of approximately 75 x 50 x 30 $Å^3$ and a molecular mass of 25 kDa in a very precise fashion on the ribosome. Before explaining the available models for this question, the main steps of whole protein synthesis mechanism will briefly be discussed in the next section.

1.3 Functional Aspects of the Ribosome

The protein synthesis can be disassembled into three consecutive phases: initiation, elongation and termination. All three phases are guided and controlled by additional factors, that is, the initiation factors (IFs), the elongation factors (EFs) and the termination/release factors (RFs). The outline of the protein synthesis is shown in Figure 1.3.

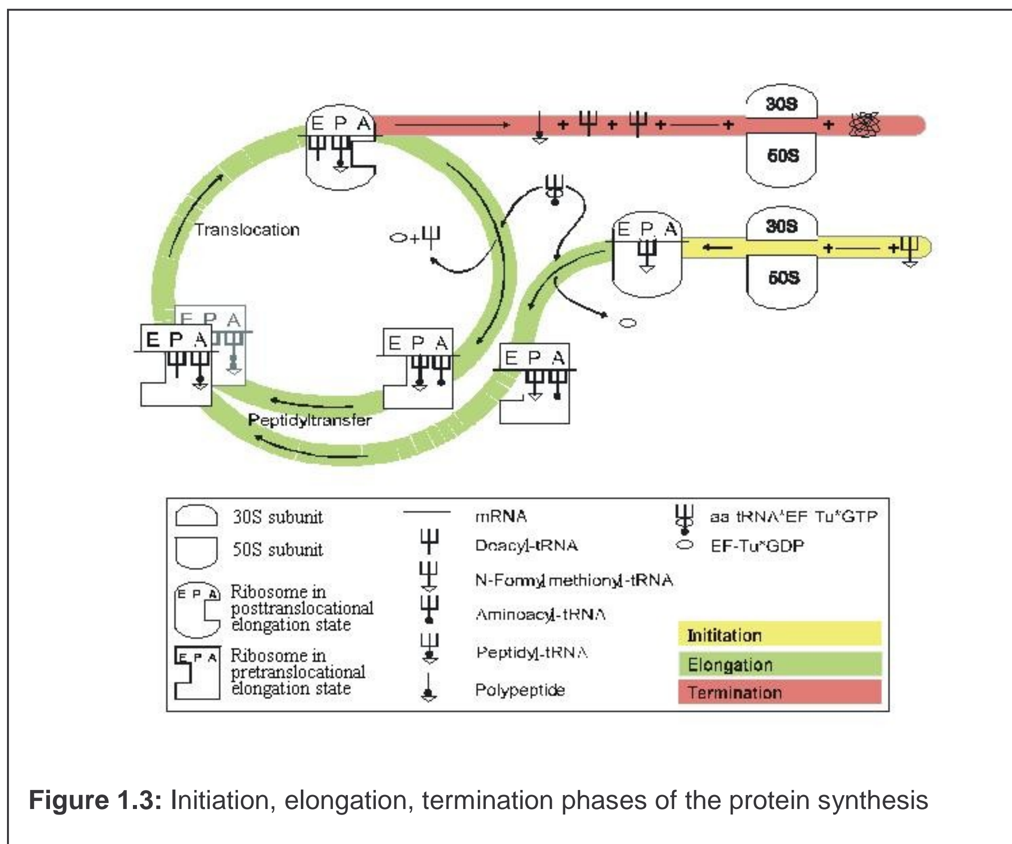

Figure 1.3: Initiation, elongation, termination phases of the protein synthesis

Initiation Phase of translation in bacteria involves the interaction of the 30S subunit with the Shine-Dalgarno sequence on mRNA that is complementary to the 3' end of 16S RNA. The process also involves three initiation factors (IF1, IF2 and IF3), which form an intermediate initiation complex consisting of an mRNA, initiator tRNA (fMet-$tRNA^{fMet}$) and the 30S subunit (reviewed in (Gualerzi and Pon, 1990)). Although the exact role of the factors and the chronology of events in initiation are still controversial, IF3 is known to bind strongly to the 30S subunit and prevent its association with the 50S subunit. IF3 also binds to the stem of the initiator tRNA and destabilize the binding of an elongator tRNA to the P site of the 30S initiation complex (Hartz et al., 1990; Risuleo et al., 1976). IF2, as a binary complex with GTP, binds the 30S subunit and directs the initiator tRNA. IF2 specifically recognizes the blocked α-amino group of the initiator tRNA, thus excluding aminoacylated elongator tRNAs in the initiation step, meanwhile IF1 binds to the A site, preventing binding of the elongator tRNA to the ribosomal A site (Carter et al., 2001). In doing so, it also induces conformational changes that may prevent the transition state in the equilibrium between the subunit association and dissociation. Although structures of bacterial IF1 (Sette et al., 1997), IF3 (Biou et al., 1995) and archeabacterial IF2 homolog (Roll-Mecak et al., 2000) have been solved, the locations of IF3 and IF2 in the 30S complex are not certain.

When the large subunit joins, the ribosomal initiation complex is achieved: the initiator codon AUG is in codon-anticodon contact with the initiator tRNA, and the A site exposes the next down-stream codon and is ready to accept the corresponding aminoacyl-tRNA. The aminoacyl-tRNA is not the direct substrate for the A site, but rather a ternary complex aminoacyl-tRNA•EF-Tu•GTP containing the elongation factor EF-Tu. With the binding of such a ternary complex the ribosome enters the elongation cycle.

Elongation cycle: Elongation cycle is the central part of the gene expression, where the genetic information of the mRNA is translated into the sequence of amino acids (Figure 1.3). Each cycle can be divided into three basic reactions:

1. A-site occupation: Aminoacylated tRNA is brought into the A site as a ternary complex with EF-Tu and GTP. This step can be further divided into two steps. (a) The decoding reaction, which restricts the aminoacyl-tRNA and ribosome interactions to codon-anticodon interactions (low affinity interaction). Correct codon-anticodon interactions result in conformational changes in the ribosome, which stabilize tRNA binding and trigger GTP hydrolysis by EF-Tu. This leads to the release of the aminoacyl end of A-site tRNA by EF-Tu. (b) The accommodation reaction where high affinity binding of the whole tRNA to the A-site results in the docking of the aminoacyl residue into the peptidyl transferase (PTF) center of the large subunit.
2. Peptide bond formation: The nascent peptide chain is transferred from the P-site tRNA (deacylation of P site tRNA) to the A-site tRNA.
3. Translocation: the movement of mRNA:$tRNA_2$ on the ribosome by one codon length so as to place the deacylated tRNA into the E site and peptidyl-tRNA into the P site that frees the A site for the next coming aminoacyl-tRNA. Translocation of mRNA and tRNAs is facilitated by EF-G, which is also GTPase. Different models have been proposed for the translocation process that is going to be explained in more detail in the next section, since this thesis is concerned with one of these models.

Termination phase: If one of the three stop codons moves into the A site during a translocation reaction, the ribosome enters the termination phase: The synthesized peptide is released and the tRNAs and mRNA dissociate from the ribosome.

1.4 Functional Insights from Structural-Biochemical Data of the Ribosome

There are three main mechanistic questions concerning translation: How are correct tRNAs discriminated from incorrect tRNAs; how is peptide bond formation catalyzed; and how are tRNAs and mRNA moved across the

ribosome's active site at the end of each cycle of amino acid addition? Although ribosomes were considered to be too large for high-resolution structural analysis for a long time, and biochemical and genetic tools were not as sophisticated as they are now to answer these questions, in the last decade rapid progress has been made due to a convergence of various approaches. The crystal structures of the ribosome have already revolutionized the field of protein synthesis by providing a three-dimensional reference for interpreting existing biochemical and genetic data.

Crystal structures for both 30S (Schluenzen et al., 2000; Wimberly et al., 2000) and 50S ribosomal subunits (Ban et al., 2000; Harms et al., 2001) have been presented at molecular resolution, which has enabled certain ligand interactions with these subunits to be identified. On the 30S subunit, interactions with initiation factors, numerous antibiotics (Brodersen et al., 2000; Carter et al., 2001; Pioletti et al., 2001) and tRNA fragments (Ogle et al., 2001) have been determined. The latter of which has led to a detailed understanding of the mechanism of ribosome decoding at the A site. In addition to antibiotics, a transition state analogue for peptide bond formation, the "Yarus inhibitor", has been soaked into crystals of the large ribosomal subunit (Nissen et al., 2000), the results of which have evoked intense discussion regarding the mechanism of peptide bond formation.

The highest resolution structure for a complete 70S ribosome is currently at 5.5 Å (Yusupov et al., 2001). At this resolution molecular interactions with bound ligands cannot be directly visualized, instead they are inferred by modeling based on the high-resolution subunit structures. This has enabled the path of the mRNA through the ribosome, encompassing 31 nucleotides from positions –15 to +16, to be determined (Yusupova et al., 2001), where the first nucleotide of the P-site codon is defined as position +1. Furthermore, the positions of A, P and E site tRNAs were determined allowing contacts with the ribosomal components to be predicted (Yusupov et al., 2001), which were in good agreement with previous studies of tRNA ribosome interactions. However, we should also note that Noller et al.'s 70S structure is a composite of two structures. The first is a 5.5-Å structure of the 70S with mRNA and

tRNA in the P and E sites. The second is a structure at 6.5-Å resolution, obtained by adding A-site tRNA to performed crystals of the 70S ribosome with P- and E-site tRNAs. It does not represent an authentic elongation ribosome with 3 deacylated tRNAs.

The general positions of tRNAs at the A, P and E sites are well known from cryo-electron microscopy studies of functionally competent complexes (Agrawal et al., 2000), but at 11.5 Å resolution relatively little information pertaining to specific tRNA ribosome interactions is available. Numerous chemical probing and crosslinking studies have been employed to map the contact sites on the rRNA from ribosome bound tRNA (or analogs thereof; refs. (Döring et al., 1994; Joseph and Noller, 1998; Joseph and Noller, 1996)), but there are few comprehensive studies examining the reverse situation, namely, protection of bound tRNAs by ribosomal components. One such study analyzed the protection against hydroxyl radical probing conferred by a 30S subunit to a P site bound tRNA (Hüttenhofer and Noller, 1992). In this study the 30S subunit shielded positions 28 to 46 (33% of the 76 positions) of the P site bound $tRNA^{Phe}$, but a number of additional protections outside this anticodon-stem-loop region were observed thus limiting the precise definition of the contact border of a tRNA bound to the small subunit. Little information is available regarding the protection pattern on a tRNA afforded by the 50S subunit.

1.5 Functional Models for the Elongation Cycle

The ribosome has a dynamic structure and the next step in understanding translation will definitely come from elucidating ribosomal dynamics. In order to get insight into possible structural changes in tRNA and rRNA in well-defined functional states of the ribosomal elongation cycle, we need now functional studies with atomic resolution. Experimental evidence for the nature of tRNA movement during the translocation accelerated with the development of footprinting techniques. Chemical modification studies are giving important

preliminary answers. Two principles have been implied. The first one is base modification with reagents like kethoxal, DMS and others. The second method is the phosphorothioate technique that modifies phosphate groups by replacing a non-bridging oxygen with a sulfur atom. The phosphorothioate method provides a tool for the analysis of function, structure, and interactions of nucleic acids with atomic resolution. Interestingly, either technique has led to a model in elongation cycle.

The Hybrid Site Model: This model (Moazed and Noller, 1989) is depicted in Figure 1.4A. The essence of this model is a creeping movement of the tRNAs through the ribosome. The model states that after peptide bond formation the tRNA portion on the large subunit moves to the next tRNA binding site. It means that the peptidyl-tRNA part on the large subunit moves from the A to the P site and the deacylated-tRNA moves from the P to the E site. The peptidyl-tRNA is now present in the hybrid site A/P (where the letters before and after the slash indicate the sites on the 30S and 50S subunit, respectively) with the anticodon stem-loop region still at the A site and the remaining region of the tRNA at the P site. Correspondingly, the deacylated tRNA is in the P/E hybrid site. And an "E-site pattern" - not as well defined in the poly(U) dependent system used – simply derives from the binding of deacylated tRNA to the E site, which requires an intact CCA-3' terminus (for review see ref.(Noller et al., 1990)).

The basic observation was that bases of rRNAs could be protected against chemical modifications by binding tRNAs specifically to the A, P and E sites. Each tRNA position on the ribosome is correlated with a protection pattern: A peptidyl-tRNA analogue AcPhe-tRNA bound to the P site could thus define a “P-site pattern”. Binding of a ternary complex Phe-tRNA·EF-Tu·GTP to the A site after pre filling the P site with a deacylated tRNA gives an “A-site pattern”.

However, when an AcPhe-tRNA was bound to the P site and the protection patterns were assessed before and after peptide-bond formation with puromycin, the P-site pattern shifted to an E-site pattern on the 23S rRNA after peptide-bond formation, in contrast to the unaltered P-site pattern on the

16S rRNA. The conclusion was that the tRNA was in a P/P state before and in a P/E state after the puromycin reaction. In another experiment a ternary complex was added to ribosomes carrying an AcPhe-tRNA at the P site; peptide bond formation would put an AcPhe-Phe-tRNA at the A site and a deacylated tRNA at the P site. An “A plus P” pattern was found on the 16S rRNA, whereas a “P plus E” pattern was resent on the 23S rRNA. The same pattern was observed, when AcPhe-tRNA was present at the A site and a deacylated tRNA at the P site. In this situation peptide bond formation is not possible following A-site occupation. These results were interpreted such that the analogue of the peptidyl-tRNA was in an A/P site and the deacylated tRNA in a P/E site. Following translocation the protection patterns suggested that the tRNAs were in the P/P and E sites, where the E site is located solely on the 50S subunit rather than on the 30S subunit. The model has been modified based on the crystal structure, where the E-site tRNA shows intensive contacts with the small subunit, in that the tRNA is now at the E/E site after translocation, [Yusupov, 2001 #12663].

The α–ε Model: The second method, namely phosphorothioate technique, has led to the α–ε model. This model (Dabrowski et al., 1998) is a revised form of the allosteric three-site model (Rheinberger and Nierhaus, 1983), that keeps all the features of the latter but explains the reciprocal linkage, which was stated in the allosteric model, between the A and E sites radically differently. By measuring the protection by the ribosome of phosphorothioated tRNAs against iodine-mediated cleavage, it is concluded that the pattern of protection characteristic of A- and P-site tRNAs does not change during translocation (Figure 1.4B). This conclusion is based on the observation that the contact pattern of the tRNA assessed by at least 60 phosphate positions does not change by either tRNA during translocation. This implies that there are movable domains within the ribosome that transport the tRNAs through the ribosome.

Comparison of these two models will be done in the Discussion part in more detail.

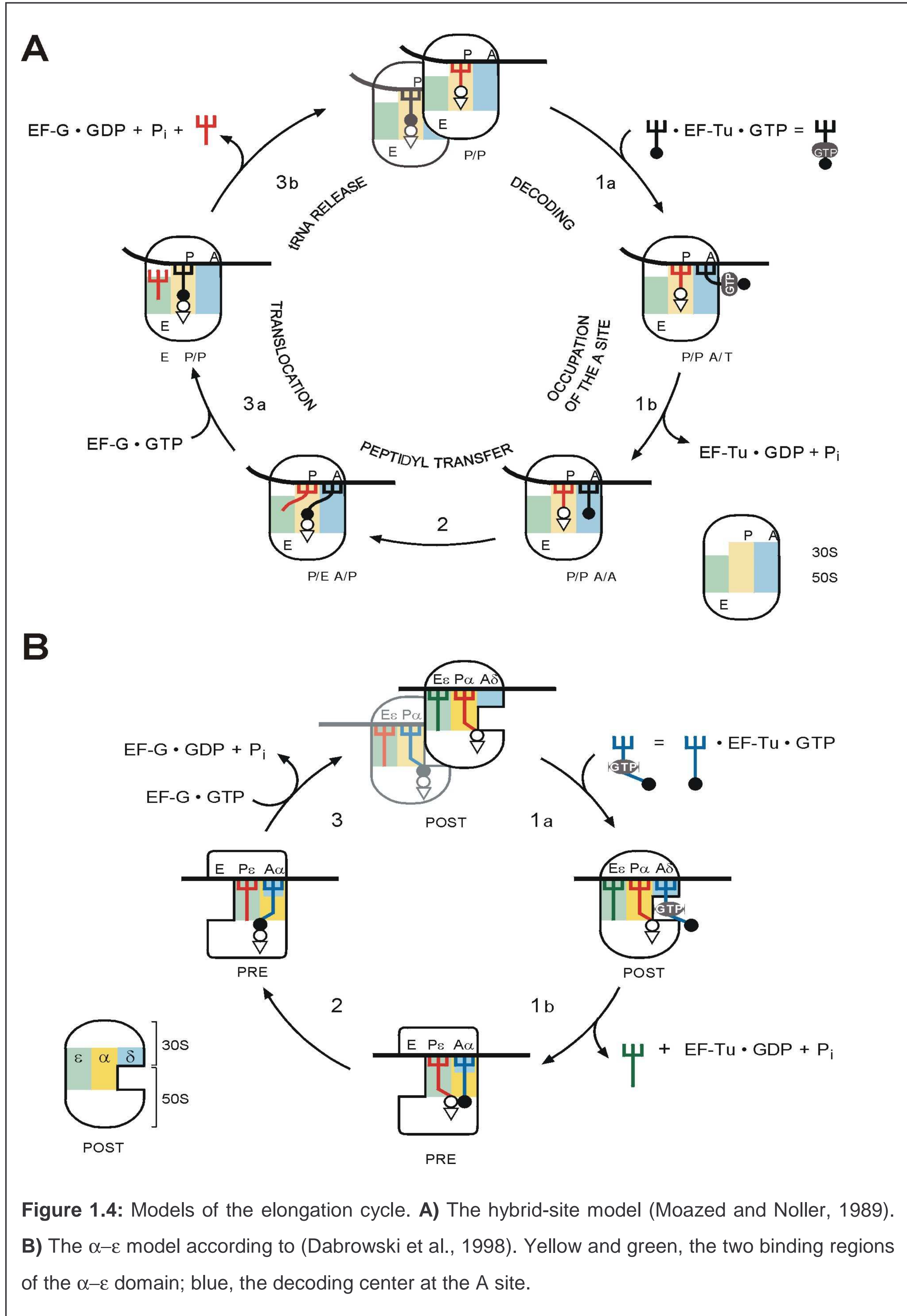

Figure 1.4: Models of the elongation cycle. **A)** The hybrid-site model (Moazed and Noller, 1989). **B)** The α–ε model according to (Dabrowski et al., 1998). Yellow and green, the two binding regions of the α–ε domain; blue, the decoding center at the A site.

1.6 Aims of this thesis

How are the enormous tRNA substrates and the mRNA ratcheted through the ribosome? Bretscher was proposing the first model for this question in 1968 in an entirely theoretical paper (Bretscher, 1968). Since then development of techniques improved our understanding, yet still it might be frightening to think how far we are to the real answers. On the other hand, solutions for mysteries of translation mechanism now seem to be in reach the more we know.

Footprinting techniques has given milestone experimental evidences to explain the nature of tRNA movement. The hybrid side model and α–ε models have sometimes been considered mutually exclusive, but they may in fact be addressing different aspects of the translocation problem. This point will be addressed in the experimental work here. However, since the systematic analysis of the elongation cycle via cryo-EM study gave no hint to the hybrid site location of the tRNAs under the in vivo conditions, the results are mainly going to be discussed in the frame of the α–ε model.

This work can be mainly divided into three parts. The first part is the setting up the experimental conditions during the whole project. The second part contains the footprinting analysis of the different ribosomal complexes in the different states of elongation cycle in the various conditions and the last part is the computational analysis of these and previously available data under the light of recent 5.5 Å X-ray crystallography map of the tRNA-70S complex.

We would like to give the data with and the discussion of the analysis and to leave the decision to your side for the degree of the laziness of tRNAs. How lazy are they? Does only a part of tRNA jump from one site to the other on the large subunit with a creeping-like movement and wait for a EF-G mediated push to complete the translocation? Or does tRNA prefer to sit on the movable domain and wait for being transported during the elongation cycle via the ribosome?