CHAPTER 1

Literature Review
The Central Dogma of Molecular biology initiates with the process of transcription that synthesizes the mRNA corresponding to the protein encoding genes. The process is successfully completed only when the protein produced by translation of the messenger RNA is folded into its biologically active form. The process of translation of polypeptide chain is performed by a ribonucleoprotein complex called the “ribosome” and the proper folding of proteins in the cell is ensured by a class of molecules called the “Molecular chaperones”.

Section I. Protein synthesis in the cell

The Ribosome

Ribosomes are large ribonucleoprotein complexes with a total molecular mass of ~ 2.4 MDa for bacteria and ~ 4 MDa for eukaryotes [Kramer G et al, 2009]. The ribosome comprises of two subunits: a large subunit and a small subunit. Bacterial ribosome is 70S type and the subunits are the 50S (large subunit) and the 30S (small subunit). The 30S subunit consists of a 16S rRNA (1542 nt) with 21 proteins and the 50S subunit contains a 23S rRNA (2904 nt), 5S rRNA(120 nt) and 31 proteins [Voisset C et al, 2008]. The eukaryotic ribosome and their subunits are 80S, 60S (large) and 40S (small) respectively [Wittmann-Liebold B, 1986]. The 40S subunit consists of 18S rRNA (1900 nt) and 33 proteins where as 60S subunit comprises of 28S rRNA (4700nt), 5.8S rRNA (160 nt), 5S rRNA (120 nt) and ~ 49 proteins.
The ribosomal subunits are held together by several inter subunit bridges [Frank J et al, 1995].

The ribosomal RNA molecules is folded into defined domains, for example the 23S rRNA contains six secondary structural domains (I–VI) formed by over 130 RNA helices. The ribosomal RNA has a structural role acting as a scaffold defining the positions of the ribosomal proteins and also performs critical functions in the ribosome that allow protein synthesis to occur. Indeed, the key catalytic activity of the ribosome – the creation of a chemical bond between two amino acids (known as a peptide bond) – comes from the RNA component of the ribosome. Hence, the ribosome is the most abundant natural ribozyme that synthesizes proteins and is the only natural RNA-based polymerase [Rodnina MV et al, 2007]. The peptidyl transferase (PT) activity of the ribosome is located on the large ribosomal subunit. The domain V of 23S rRNA contains the peptidyl transferase centre (PTC). Crystal structures of the 50S subunit have revealed that the PTC is composed of RNA only, with no protein within 15Å° of the site [Nissen P et al, 2000].
Figure 1.2: Bacterial ribosomal RNA. a) Bacterial ribosome showing six secondary structural RNA domains with different color code. b) Secondary structure of bacterial 23S rRNA showing the individual domains [Picture adapted from Ban et al., 2000].
The two non-ribosomal RNA, involved ribosome mediated protein synthesis are mRNA (messenger RNA) and tRNA (transfer RNA). With the mRNA codons specifying amino acid sequence during protein synthesis while as tRNA carries the appropriate amino acid thus functioning as an adaptor molecule. In bacteria 30S ribosomal subunit harbors the decoding center, where molecular recognition between the mRNA codon and the corresponding tRNA anticodon takes place. Three distinct binding sites for tRNA are located in the interface between large and small subunit, A-site (accepting amino acylated tRNA), P-site (holding peptidyl tRNA) and E-site (holds the deacylated tRNA).

**Bacterial Protein Synthesis**

During protein synthesis ribosome reads the genetic message encoded in the messenger RNA with the help of different protein factors. Transfer RNAs (tRNAs) play an equally important role as adaptors that can bind an amino acid which then recruited on ribosome for translation. In protein synthesis tRNA molecules make specific contact with the mRNA codon and deliver their amino acid one at a time to the growing polypeptide. The mechanism of translation can be conveniently divided into four phases: initiation, elongation, termination and ribosome recycling. The steps of translation are shown in the following figure.
Figure 1.3: Schematic diagram of bacterial protein synthesis. The mains steps comprising the translation process: initiation, elongation cycle, termination and recycling. The mRNA is depicted as a strand running horizontally along the small (30S) subunit, with alternating white and black segments, each representing one codon. The tRNAs bind at A, P and E sites. The nascent polypeptide is shown as a string of spheres. The individual structures and cartoons are not drawn to scale [Picture adapted from Agirrezabala X et al, 2010].

Ribosome targeting antibiotics

The antibiotics are natural or synthetic compounds that selectively kill or inhibit the growth of microorganisms, and many of them are sufficiently specific to be useful for the treatment of bacterial infections. A large fraction of these compounds are inhibitors of bacterial protein synthesis, and almost all such antibiotics interact with the ribosome. The peptidyl transferase center (PTC) is the target of many of these antibiotics, and some of them are targeted on the translation elongation or ribosome recycling process. In our studies the ribosome binding
antibiotics have been used to study the process of ribosome chaperoning and unfolded protein mediated dissociation of ribosomal subunits (Discussed below in Section III.2). The site of binding and the mode of action of the antibiotics used in our study is outlined below.

**Inhibitors of peptidyl transferase activity**

_Chloramphenicol_

The antibiotic chloramphenicol inhibits protein synthesis by interacting with the peptidyl transferase center on the large ribosomal subunit [Long KS et al, 2003]. Two chloramphenicol binding sites have been reported in structures of antibiotic-ribosomal subunit complexes, solved through X-ray crystallography. Chloramphenicol binds to the ribosomal A site that hinders substrate binding, which direct inhibition of peptide bond formation during protein synthesis [Schlünzen F et al, 2001]. In another way chloramphenicol binds to a hydrophobic crevice at the entrance to the peptide exit tunnel that inhibits protein synthesis by preventing the entry of nascent polypeptides into the ribosomal exit tunnel [Hansen JL et al, 2003].

_Blasticidin_

Blasticidin S is an aminohexose pyrimidine nucleoside antibiotic. It inhibits the ribosomal peptidyltransferase activity [Hansen JL et al, 2003]. Blasticidin S interacts with two specific guanine nucleotides on the 23S rRNA: G2251 and G2252. Interaction of blasticidin with G2251 is much stronger than with G2252. These are the two critical bases of the P loop of 23 S rRNA that base-pair with C75 and C74, respectively, of tRNA, and thereby position the CCA sequence of P-site bound tRNAs in the peptidyl transferase center [Hansen JL et al, 2003]. These data lead to the expectation that blasticidin S should inhibit protein synthesis by competing with P-site substrates.
**Puromycin**

Puromycin is an aminonucleoside antibiotic. It causes premature peptide chain termination during translation on the ribosome. It resembles the aminoacyl end (3' of the tRNA) of aminoacylated-tRNA. Puromycin enters into the ribosomal A site and forms puromycylated – nascent peptide complex which subsequently released from the ribosome [Maden BE, 2003]. An attractive possibility of the mechanism of action is that, nucleophilic substitution of the amino group of the p-methoxyphenylalanyl moiety of puromycin on the C-terminal acyl group of the growing polypeptide chain with displacement of the soluble RNA to which the peptide chain is considered to be esterified.

**Figure 1.4:** Superposition of antibiotic binding locations and hydrophobic crevices. Surface representation of the ribosome (gray contour) shows that many antibiotics (stick figures) interact in part with either the active-site hydrophobic crevice (green contour, upper middle) or the exit tunnel hydrophobic crevice (green contour, lower right) [Picture adapted from Hansen JL, et al, 2003].
**Inhibitors of Ribosome recycling**

Fusidic acid and aminoglycosides are well known antibiotics having inhibitory effects on ribosome dissociation/recycling. Fusidic acid (FA) is a inhibitor of EFG function. FA inhibits EFG turnover, in either GTP hydrolysis or translocation [Johanson U et al, 1996] and it also blocks ribosome disassembly by EFG/RRF.

Aminoglycoside antibiotics in the 2-deoxystreptamine (2-DOS) family are broad-spectrum bactericidal agents that are used to treat Gram-negative bacterial infections. Their main target is ribosome. Members of this class exhibit various effects on translation that decrease tRNA selection accuracy, inhibit spontaneous and EFG–catalyzed translocation, and affect ribosome recycling [Feldman MB et al, 2010]. Two potent antibiotics of this class are neomycin and paromomycin. We used these two aminoglycosides to study the effect of the same on ribosome dissociation into subunits.

**Neomycin**

Crystallographic data revealed that neomycin binds to both the helix44 decoding site of 16S rRNA and helix69 of the large subunit 23S rRNA. High-affinity and low affinity neomycin binding are associated with helix44 and helix69 respectively. Recent functional studies also suggest that neomycin binding to helix69 begins to block the essential translation reactions even at low concentrations (>10 nM) [Wang L et al, 2012]. Neomycin inhibits translation process by preventing both the ribosome recycling and substrate (mRNA and tRNA) translocation. Neomycin binding at helix69 allosterically regulates ribosome dynamics and prevents the specific interaction between RRF and helix69, essential for ribosome recycling. Stabilization of
classical tRNA positions on ribosome by neomycin is an important and general mechanism for translocation inhibition.

**Figure 1.5: Interaction of neomycin with ribosomal RNA.** Overview of neomycin binding sites in h44 (dark green) of the small ribosomal subunit (blue) and in H69 (lightblue) of the large ribosomal subunit (gray). tRNA is shown in green, mRNA is shown in red, and Large subunit (L) proteins are shown in purple [Picture adapted from Wang L et al., 2012].

**Paromomycin**

Structural evidence suggests that paromomycin can bind both the helix44 and helix69 [Feldman MB et al, 2010]. The best-known function of paromomycin is its miscoding effect. Like neomycin, paromomycin may also block the mRNA and tRNA translocation by stabilizing the tRNA on ribosome. Alternatively, paromomycin inhibits the antiassociation activity of initiation factor 3 (IF3), which is an important component of the disassembly reaction of posttermination ribosomal complexes. This inhibitory effect comes from the fact that paromomycin strengthens the interaction between ribosomal subunits and even induces association of the subunits at low Mg$^{2+}$ concentrations [Hirokawa G et al, 2007].
stabilization of 70S ribosomes by paromomycin may be responsible for its inhibitory effects on translocation and ribosome recycling.

Section II. Protein folding in the cell

After synthesis at the ribosome in the highly crowded cellular environment, the polypeptide side chains are exposed to manifold non-specific interactions. The proteins however usually manage to selectively form the restricted set of intramolecular contacts that funnel the folding of the protein towards the unique, thermodynamically stable and functional native structure [Jahn TR et al, 2008]. Christian B. Anfinsen in his pioneering in vitro work [Anfinsen CB, 1973] demonstrated the re-formation of native structure of ribonuclease by diluting out the denaturant and concluded that the information for correct folding of a protein lies in its amino acid sequence. Although the physical principles of protein folding have been studied intensely for almost 50 years, how the folding process is determined by the amino acid sequence remains one of the most important problems in biology [Fersht AR, 2008]. Several ideas had been proposed to explain the mechanisms of folding from experiment and computer simulation. Levinthal [Levinthal C, 1968] by simple calculations however suggested that even a small protein of 100 amino acids would take over a billion years to find its native state if folding were to occur by a random search of all possible conformations. He raised the question how does the protein “know” what conformations not to search? These observations thus led to the view that there must be a specific pathway for folding and the crucial aim was to understand how a protein finds the ‘right’ pathway and avoids the others. In addition, it was appreciated that metastable, partially folded states of proteins exist and, if they could be detected and studied, these species might provide important clues about the stepping stones to the native state [Dinner AR et al, 2000].
Molecular chaperones are proteins that interact with, stabilize, or help a non-native protein to acquire its native conformation. The chaperones are not permanent components of protein structures when they are performing their normal biological functions [Hartl FU et al., 2009]. Chaperones are also involved in a multitude of cellular functions, including de novo folding, refolding of stress-denatured proteins, oligomeric assembly, intracellular protein and transport, protein disassembly processes, such as the partial unfolding and subunit dissociation that occur when some proteins carry out their functions, and the repair and degradation of proteins partially denatured by chemical or physical agents. The first identified molecular chaperone was nucleoplasmin, which mediate the assembly of nucleosomes [Laskey RA et al, 1978]. The chaperones include about 20 protein families of different molecular weights, structures and cellular locations. Majority of these proteins are known as heat shock protein (Hsp) as their biosynthesis is regulated by cell stress, specifically heat. Chaperones are usually classified according to their molecular weight i.e Hsp40, Hsp60, Hsp70, Hsp90, Hsp100 and the small Hsps.

The chaperones that actively participate in protein biogenesis are ribosome associated chaperones, Hsp70 and chaperonins (Hsp60s). These chaperone classes functions by interacting with the exposed hydrophobic portion on the unfolded or non-native proteins and assist in folding through ATP-dependent binding and release cycles. There is no known chaperone that can contribute steric information to the folding process, but rather they function by optimizing the efficiency of folding [Vabulas RM et al, 2010].
Ribosome-associated Chaperones

II.1. In prokaryotes

Trigger factor

The well known ribosome-associated chaperone is bacterial trigger factor (TF). Trigger factor is a highly abundant protein (20,000 copies per cell) that binds to the large subunit of the ribosome, close to the polypeptide exit site [Hesterkamp T et al, 1998]. TF is an ~50 kDa protein, consisting of an amino-terminal ribosome binding domain, a peptidyl-prolyl isomerase (PPIase) domain, and a carboxy-terminal domain [Ferbitz L et al, 2004]. Thus TF displays both the peptidyl prolyl cis-trans isomerase [Stoller G et al, 1995] as well as chaperoning activities. The amino-terminal domain mediates the interaction with the ribosome [Lakshmipathy SK et al, 2007; Merz F et al, 2008] and the carboxy-terminal domain interacts with the hydrophobic segments of nascent chain [Merz F et al, 2006; Lakshmipathy SK et al, 2007]. Binding of TF to free ribosomes is salt sensitive, but this interaction is stabilized during translation. Ribosome binding facilitates the TF-nascent chain interaction [Baram D et al, 2005; Kaiser CM et al, 2006]. TF then moves with the nascent chain, preventing premature or incorrect folding. Following release, TF could cycle back to the ribosome to wait for the next substrate. The eventual dissociation of TF facilitates folding or polypeptide transfer to downstream chaperones, such as the Hsp70 protein DnaK. High abundance and high degree of conservation of TF in the bacterial kingdom suggest important role of TF in assisting folding. However it has been reported that E. coli cells lacking TF show no general defects in de novo protein folding at 30°C [Kramer G et al, 2004; Genevaux P et al, 2004].
The HSP70 chaperone family

Hsp70 homologues are widespread in prokaryotes as well as in eukaryotes where they occur in the cytosol, mitochondria, chloroplast and the endoplasmic reticulum. The main Hsp70 protein in bacteria is DnaK [Chang HC et al, 2007]. Hsp70s generally function in cooperation with Hsp40 or J-proteins and with various nucleotide exchange factors (NEFs) in an ATP-dependent manner. The striking characteristic of these chaperones is their binding to short, linear segments of unfolded proteins or peptides containing hydrophobic residues [Rüdiger S et al, 1997; Fourie AM et al, 1994; Hesterkamp T et al, 1998]. DnaK proteins have three domains, an N-terminal ATPase domain, a central substrate binding domain and a C-terminal domain with potential regulatory function. The ATP-bound state of DnaK allows peptide binding. Hydrolysis of ATP to ADP is strongly accelerated by Hsp40, leading to stable peptide binding. Interaction of substrate with Hsp70 is mediated by the so-called J-domain that is present in all Hsp40s [Mayer MP et al, 2000]. Following ATP-hydrolysis, various NEFs (GrpE in bacteria) bind to the Hsp70 ATPase domain and catalyze ADP-ATP exchange, leading to substrate release and completing the reaction cycle. DnaK also show disaggregation activity towards protein aggregates. It has been reported that DnaK can associate with E. coli ribosome [Ghosh J et al, 2003] possibly at 5S rRNA [Okada S et al, 2000]. In vivo, however Δ dnaK cells are viable between 30°C and 37°C.

Heat Shock Protein 15 (Hsp 15)

Hsp 15 is a highly abundant protein, about 12,000 molecules per cell at 29°C [Korber P et al, 2000]. This protein binds with the ribosomal free 50S subunit with high affinity. The binding site of Hsp15 on the 50S subunit is accessible only in the absence of 30S subunit and blocked upon formation of 70S ribosome. At translation termination the nascent peptide chain is released by the action of termination factors before ribosome gets dissociate into subunit. During heat
stress premature translation termination leads to tight binding between 50S and nascent peptide chain. This prevents the peptide bound 50S to recycle for new round of translation initiation. Hsp15 binds with the part of rRNA of 50S subunit that changes its conformation and repairs it to re-enter the translation cycle [Korber P et al, 2000].

The Chaperonins

Chaperonins are large double-ring complexes of approximately 800 kDa. There are two groups of chaperonin exist [Horwich AL et al, 2007; Tang YC et al, 2007]. Group I chaperonins (called Hsp60s) occur in bacteria (GroEL), mitochondria and chloroplasts. They have seven-membered rings and functionally cooperate with co-chaperone Hsp10 (GroES in bacteria), which form the lid of the folding cage. The group II chaperonins in archaea (thermosome) and in the eukaryotic cytosol (TRiC/CCT) have eight or nine membered rings, interact with the co-chaperone prefoldin [Siebert R et al, 2000]. The GroEL/GroES chaperonin system of E. coli has been widely studied [Hartl FU et al, 2002]. GroEL interacts with approximately 250 different cytosolic proteins, most of these are between 20 and 50 kDa in size. GroEL exhibits a complex overall structure composed of two heptameric rings of the large subunit GroEL (57 kDa) stacking back to back and forming a 14-subunit hollow cylinder with two identical binding sites for non-native proteins. The smaller co-chaperone GroES (10 kDa) forms a seven membered dome shaped single ring. The ring system of GroEL contains the ATPase domain. The non-native proteins are encapsulated in the chaperonin ring by GroES (cis-ring) [Brinker A et al, 2001; Tang YC et al, 2006]. GroES binding to GroEL is ATP-regulated and form a cage with a highly hydrophilic, net negatively charged inner wall. Encapsulated protein up to approximately 60 kDa is free to fold in this environment for 10–15 seconds. Protein substrate leaves the cage on GroES dissociation, which is also mediated by ATP binding in the opposite ring (trans-ring). Thus the GroEL/GroES
chaperonin system avoids aggregation during folding by enclosing the unfolded protein, one molecule at a time. GroEL and GroES are the only chaperones, which are essential for the growth of *E. coli* under all conditions [Fayet O et al, 1989].

**II. 2. Eukaryotic chaperone network**

The eukaryotic chaperone machinery can be classified as two groups with distinct functions [Albanèse V et al, 2006]. A set of Chaperones Linked to Protein Synthesis (CLIPS) associates specifically with ribosomes to facilitate *de novo* folding. CLIPS include the ATP-independent nascent polypeptide-associated complex (NAC), and GimC/Prefoldin; as well as ATP-dependent chaperones, including Hsp70 family members such as SSB in yeast, and the eukaryotic chaperonin TRiC/CCT [Albanèse V et al, 2006]. While the other chaperone group, called Heat Shock proteins are induced under stress conditions to protect the proteome and mediate either refolding or quality control. The ribosome–associated nascent polypeptide-associated complex (NAC) is a highly conserved dimeric complex that can be crosslinked to very short nascent chains [Raue U et al, 2007]. Yeast cells contain alpha and beta isoforms that can form hetero- and homo oligomers [Beatrix B et al, 2000]. The N-terminus of beta NAC interacts with the ribosomal protein Rpl31 close to the exit tunnel [Zhang Y et al, 2012]. NAC acts as a cotranslational binding factor for most of the nascent chains [del Alamo M et al, 2011].

Another broad classes Hsp70s were observed to bind cotranslationally to nascent chains and is best studied for the *S. cerevisiae* Hsp70 SSB. The yeast SSB has two isoforms, Ssb1 and Ssb2. SSB associates with ribosomes and directly binds a large fraction of newly translated polypeptides [Willmund F et al, 2013]. ATP hydrolysis by SSB is stimulated by the Ribosome-Associated Complex RAC, which is composed of the Hsp70 Ssz1 and the J-domain protein
Zuo1 [Huang P et al, 2005]. RAC binds directly to ribosomes but does not interact with nascent chains itself [Yam AY et al, 2005].

The ring-shaped chaperonin TRiC/CCT associates co- and posttranslationally with approximately 5–10% of newly made proteins [Yam AY et al, 2008], predominantly beta-sheet rich proteins that likely require the protected environment of its central cavity to fold [Douglas NR et al, 2011]. This eukaryotic chaperone network for in vivo folding is very resilient, as deletions of single components like GimC/PFD, NAC or SSB are not lethal [Albanèse V et al, 2010; Albanèse V et al, 2006; Koplin A et al, 2010].

Figure 1.6: Chaperone network. In prokaryotes, Trigger factor is the main ribosome-associated chaperone. It has been directly involved in folding of new proteins or passing them to the Hsp70 system of DnaK/DnaJ or the chaperonin GroEL/ES to facilitate de novo and stress-induced folding. In eukaryotes, ribosome-bound chaperones and co-chaperones like SSB, RAC, SRP, and NAC compete for overlapping ribosomal binding sites. Prefoldin, TRiC and Hsp90 are the main downstream chaperone systems that are actively involved in protein folding [Picture adapted from Pechmann S et al, 2013].
It has been reported [Bukau B et al, 2000] that about 25% -30% of prokaryotic cytosolic proteins and 15% -22% of eukaryotic cytosolic proteins are get assistance by Hsp70-Hsp40-Chaperonins and by Hsp70/Hsp40 respectively. Hence the above molecular chaperones can assist in folding a relatively small fraction of proteins in the cell. The remaining vast portion of cellular proteins might follow an unassisted default pathway that involves the cotranslational formation of native or native like domain structure, followed by the rapid completion of folding upon chain release from the ribosome [Agashe VR et al, 2004]. Based on previous in vitro and in vivo report, we suggest that the cellular translation machinery- the ribosome, can act as a molecular chaperone.

II. 3. Ribosome as chaperone

It has been reported that the ribosomes from different sources like E. coli, B. subtilis (prokaryotes), wheat germ (plant), rat liver (animal) and Methanosarcina barkery (archaebacteria) can refold a number of proteins with diverse physicochemical properties like – Bacterial Alkaline Phosphatase, Beta Galactosidase, Glucose 6-Phosphate Dehydrogenase, Glucose Oxidase, Lactate Dehydrogenase, Malate Dehydrogenase, Bovine Carbonic Anhydrase, Human Carbonic Anhydrase, Beta Lactamase, Restriction Endonuclease EcoR1, Restriction Endonuclease BamH1, DnaK, HspH from Bradirhizobium japonicum, GFP, Horse radish peroxidase, Luciferase etc.[ Das B et al, 1992; Das B et al, 1996; Chattopadhyay S et al, 1996; Ghosh J et al, 2003]. It was also shown that Ricin-A chain [Argent RH et al, 2000] and Rhodanase [Kudlicki W et al, 1995], could be folded by ribosomes [Argent RH et al, 2000].

The only criterion for the choice of test protein was that the protein (enzyme) should be compatible with quick and quantitative assay. In the in vitro experiments for each protein, the level of regain in enzymatic activity with ribosome was around 80-90%. Upon separating subunits, the protein folding activity was detected in the large ribosomal subunit from various
sources [Das B et al, 1996]. Then the ribosomal proteins were increasingly depleted from the large ribosomal subunit with increasing concentration of lithium chloride, and the activity was traced in its large rRNA (e.g.; 23S rRNA in case of bacteria) [Das B et al, 1996; Chattopadhyay S et al, 1996]. Finally, the region in the domain V of *E. coli* 23S rRNA having the peptidyl transferase activity (the peptidyl transferase center, PTC) was found to have the protein folding activity [Chattopadhyay S et al, 1996; Pal D et al, 1997; Pal S et al, 1999; Chowdhury S et al, 2002; Sanyal SC et al, 2002; Samanta D et al, 2008]. Even when this segment of ribosomal RNA was cloned and transcribed *in vitro*, the transcribed PTC-RNA sequence (domain V RNA) could fold denatured protein equally well. The results of *in vitro* studies were further supported by *in vivo* experiments. Since the protein folding activity was detected in the PTC [Chattopadhyay S et al, 1996; Pal D et al, 1997; Pal S et al, 1999; Chowdhury S et al, 2002; Sanyal SC et al, 2002; Samanta D et al, 2008] of the large ribosomal subunit of bacteria, it could be stopped by antibiotics like chloramphenicol or lincomycin which bind to the PTC [Long KS et al, 2003; Douthwaite S et al, 1992] So, the synthesis of protein was terminated with streptomycin or kasugamycin which bind to *E.coli* ribosomal 30S subunit [Carter AP et al, 2000; Schuwirth BS et al, 2006] leaving the 50S subunit unaffected. The nascent polypeptides, already made, took some time to fold which was recorded. But if the PTC binding chloramphenicol or lincomycin was added, the nascent proteins could not fold, perhaps because the PTC region responsible for synthesis as well as folding of nascent protein remained inaccessible due to antibiotic binding. The difference between protein activities in presence of streptomycin and in presence of chloramphenicol was significant and was a measure of protein folding in *E.coli*. Ribosome mediated protein folding was independent of DnaK, DnaJ and Trigger factor. When the chemically unfolded protein was allowed to refold in presence of the ribosome/23S
rRNA/domain V RNA along with PTC specific antibiotics, its percent recovery of native activity decreased dramatically with the antibiotics. This indicated towards the presence of a common feature of \textit{in vivo} and \textit{in vitro} protein folding by the PTC RNA. Studies on the mechanisms of domain V chaperoning activity showed that it is a two step process involving its two sub-domains RNA1 and RNA2 [Pal S et al, 1999]. The initial binding of the unfolded proteins take place with the RNA1 subdomain that is the central region of PTC and although the substrate proteins possess no apparent RNA binding domain, they interact with the RNA1 region of domain V RNA via specific interactions [Samanta D et al, 2008]. The RNA2 region of this domain is responsible for the releasing the bound protein which subsequently folds into its native structure [Pal S et al, 1999]. The domain V of large subunit rRNA of bovine mitochondrial ribosome (mDV RNA) has a truncated RNA2 region (Figure 1.7) and therefore shows a delay in releasing the bound protein [Das A et al, 2011]. Unlike other cellular foldases neither the binding nor the release steps are associated with ATP hydrolysis.
Nucleotide mutations at various positions in domain V of the 23S rRNA can hamper peptidyl transferase or protein folding or both the activities of ribosome [Pang Y et al, 2013; Das A et al, 2011]. Mutations have been introduced in RNA1 and RNA2 by site-directed mutagenesis. Some of the mutants had a number of nucleotides changed and turned out to be deficient in protein folding. Many of the mutations have been known to cause a number of different changes in the functions related to protein synthesis. Most of the mutations coincide with antibiotic resistance mutations.
Figure 1.8: Point mutations on RNA1 and RNA2 subdomain of domain V RNA. a) Introduced point mutations on RNA1 and color boxes in specific nucleotide position represent degree of defect in refolding efficiency by these mutations, harsh effect; red box, mild effect; blue box. b) Mutation in the specific nucleotide (blue box) in RNA2 of domain V RNA has mild effect on refolding efficiency by this RNA domain.

Section III. Objectives of the present study

III.1. Objective 1: Effect of ribosome on folding of partially unfolded proteins

The process of aggregation of proteins often competes with the process of protein folding. Protein aggregates are non-functional and are toxic to the cell when accumulated in large quantities. One of the primary functions of chaperones is to bind to aggregation prone intermediate states of proteins and suppress their aggregation. Our objective is to study the ability of the ribosome and domain V of 23S rRNA influence the protein aggregation process. Some key aspects of aggregation and protein aggregate toxicity are discussed below.
Folding Intermediates: The Molten Globule State

Large proteins with 100 or more residues (almost 90% of all proteins in a cell) fold via formation of intermediates and have a significant tendency to rapidly collapse in aqueous solution into compact non-native conformations [Brockwell DJ et al, 2007; Bartlett AI et al, 2009]. Such intermediates might either represent on-pathway “stepping stones” toward the native state or kinetically stable, misfolded conformations prone to aggregation. In vitro studies show that proteins can also assume a partially folded, collapsed state under both equilibrium and non-equilibrium conditions [Shortle D, 1996]. Since such partially folded proteins resemble the intermediate states along the protein folding pathway, studies using these intermediates might play an important role in understanding the mechanisms of protein folding.

The molten globule state of a protein represents an ensemble of compact, partially folded protein that has native-like secondary structure and backbone folding topology, but lacks the extensive, specific side-chain packing interactions of the native structure [Ptitsyn OB, 1995]. Structural studies have shown that the side-chains in a molten globule can adopt a greater variety of conformations than in a native protein. The ability of the side chains to form the tight and specific interactions typical of a native protein, is the essential final step in the protein folding pathway. This packing process is considered to be energetically more difficult than forming the collapsed, disordered folding intermediates. The side chains in molten globules are significantly more mobile than those in the native protein. These partially folded states often tend to aggregate because of exposed hydrophobic amino acid residues, regions or patches of unstructured polypeptide backbone, that remain largely buried in the native state. The intermolecular interaction between these intermediates is driven by hydrophobic forces and predominantly results in the formation of amorphous structures. Alternatively, aggregation can lead to the
formation of ordered, fibrillar assemblies called amyloid. Although apparently restricted to a subset of proteins under physiological conditions, these thermodynamically stable structures are accessible to many proteins under denaturing conditions and are largely independent of the protein sequence. This suggests that the formation such intermediates is an inherent property of the polypeptide chain [Dobson CM, 2003]. The commonality of intermediates during folding also specify the need for molecular chaperones in all forms of life, so as to avoid problems arising from potential aggregation of these states, suggesting the co-evolution of protein sequences and their folding assistants.

**Folding Energy Landscape of Protein Folding and Aggregation**

In the recent past, the most important observation in protein folding is that there is not a single, specific folding pathway, as was suggested in some early models [Kim PS et al, 1982; Dill KA et al, 1997]. Instead, a multidimensional energy landscape or folding funnel better describes the folding process (Figure 1.9). A folding funnel is a 2D or 3D representation of the conformational space that is accessible to the polypeptide backbone during folding [Schultz CP et al, 2000]. Energetically, the breadth of the funnel represents all possible conformations of the chain (chain entropy); the broad top of the funnel represents enormous number of conformations present in the soluble denatured state; the needle like points towards the bottom of the funnel represents the more stable (free energy minima) structures including the native structure of the protein as determined by X-ray crystallography or NMR. As the chain folds to lower energy conformations, it might populate intermediates. These kinetic traps might hinder and/or promote formation of the native structure depending on their type of interactions, the barriers between the trap and native conformation and the rest of the funnel surface [Clark PL et al, 2004]. The intramolecular interactions within the same chain lead to folded native populations, whereas intermolecular
interactions between different chains generate multimolecular network called aggregates. Aggregation is a concentration dependant process, because it is multimolecular in nature. So at a low protein concentration, folding is the predominant pathway, aggregation prevails at higher protein concentrations. The aggregation process can lead to either formation of ordered, fibrillar structure called amyloid or less ordered, oligomeric intermediates.

![Energy landscape scheme of protein folding and aggregation](Figure_1.9.png)

**Figure 1.9: Energy landscape scheme of protein folding and aggregation** [Picture adapted from Vabulas RM et al, 2010].

**Toxicity of protein Aggregates**

The protein quality control (PQC) system is a multifunctional unit that governs folding, counteracts aggregation, and eliminates misfolded and damaged polypeptide chains before they can exert toxic effects. Failure of the PQC system to degrade misfolded proteins may lead to formation of protein aggregates. Aggregates in the cytosol may accumulate at a single large body
called aggresome, or as soluble monomers and oligomers, which later may precipitate into long amyloid fibrils [Gregersen N et al, 2006]. These aggregates have various toxic effects on the host cell. Protein aggregates toxicity can result in diseases that range from cancer and diabetes to neurodegeneration. Neurodegenerative diseases, such as Parkinson’s disease (PD), Alzheimer’s disease (AD), amyotrophic lateral sclerosis (ALS) and the polyglutamine diseases that include Huntington’s disease (HD), arise from abnormal protein interactions in the central nervous system. In all of these diseases, there are characteristic deposits of protein aggregates in the brain, which can be cytoplasmic, nuclear or extracellular [Ross CA et al, 2005]. In aggregation diseases, large intracellular or extracellular accumulations of aggregated protein, known as inclusion bodies, are often formed. The inclusion bodies frequently contain the disease protein in a fibrillar aggregated form called amyloid [Sunde M et al, 1998]. In several other studies it has been indicated that the inclusion body formation not always linked to cytotoxicity, but it might be a cellular protective response [Ross CA et al, 2005]. The several possible intermediate species including oligomeric forms and protofibrillar forms associated with the complex multi-step process of protein aggregation are responsible for toxicity. There might be a common structural motif that are shared by various aggregating proteins termed as “…common toxic fold” [Dobson CM, 2003; Bucciantini M et al, 2004; Kayed R et al, 2003]. Alternatively, toxicity might involve the exposure of moieties that are usually hidden in globular protein structures, such as hydrophobic side chains, or main chain NH and CO groups in an abnormal β-conformation that could lead to the formation of non-native hydrogen bonds.

There are also many proposals that deal with the molecular mechanism of cellular toxicity [Ross CA et al, 2005]. One interesting hypothesis is that abnormal proteins might inhibit the proteasome - involved in degradation of non native proteins [Bence NF et al, 2001]. Long
polyglutamine stretches are difficult for proteasomes to digest and thus inhibition direct of proteosome functioning. Abnormal proteins could also alter autophagy. Mutant α-synuclein for example, can impair chaperone mediated autophagy. A more general proposal for the relationship between aggregation and toxicity is that toxicity depends on interactions with, and possibly recruitment of, other cellular constituents during the aggregation process. The active process of aggregation and interactions with other proteins in the cell might also cause toxicity.

**Cellular protective mechanisms**

The molecular chaperones that assist in refolding of proteins and prevent their aggregation form the first line of defense against toxic protein aggregate formation. A second important cellular defence against misfolded proteins involves degradation by the proteasome [Ross CA et al, 2004]. A third defence mechanism involves autophagy, which has several variants, including macroautophagy, microautophagy and chaperone-mediated autophagy.
Figure 1.10: Quality control of protein folding in the cytosol. A protein folds through different intermediates to its native, three-dimensional structure. Proteotoxic stresses, mutations in the synthesized protein or translational errors can cause protein misfolding. Folding intermediates can be refolded to the native state or be degraded by different cellular proteolysis systems. In harsh stress conditions, increased amounts of aberrant proteins aggregates can form. Their formation can be guided by molecular chaperones. These aggregates may be disordered or highly ordered β-sheet-rich amyloid fibrils. Disordered aggregates and intermediates during amyloid formation may be degraded. Arrows indicate a process that can include several single steps; dashed arrows indicate a process of minor significance. [Picture adapted from Tyedmers J et al, 2010]

In our studies on the effect of this chaperone on the protein folding and aggregation, the proteins bovine carbonic anhydrase II and chicken egg white lysozyme were chosen. The human homologues of the proteins lysozyme and carbonic anhydrase are disease associated and have
been shown to possess the ability of amyloid formation. Human lysozyme is involved in the formation of amyloid deposits in autosomal hereditary systemic amyloidosis [Pepys MB et al, 1993]. X-ray crystallography and biophysical studies have shown that the lysozyme variants are structurally similar to wild-type (WT) lysozyme [Booth DR et al, 1997] and that their propensities to form amyloid fibrils can be attributed to their reduced thermodynamic stabilities and enhanced ability to form partially unfolded amyloidogenic states [Canet D et al, 1999; Canet D et al, 2002]. Further, the disease marble brain syndrome (MBS), is a carbonic anhydrase II deficiency syndrome (CADS) that can manifest in carriers of point mutations in the human carbonic anhydrase II (HCA II) gene. One mutation associated with MBS entails the His107Tyr substitution. It was demonstrated that this mutation is a remarkably destabilizing folding mutation and the mutant protein populates the aggregation prone molten globule state of the protein [Almstedt K et al, 2004].

Our studies on the effect of ribosomal chaperone on folding and aggregation of the molten globule state of bovine carbonic anhydrase II and reduced and denatured lysozyme are outlined in chapter 2.

**III.2 Objective 2: Studies on mechanism of unfolded protein induced ribosome subunit dissociation**

Earlier studies have demonstrated that a significant outcome of the unfolded protein-ribosome interaction is the rapid dissociation of the ribosome into its subunits [Basu A et al, 2008a; Basu A et al, 2008b]. Our studies were aimed at understanding the mechanism of unfolded protein mediated subunit dissociation. The relationship between subunit dissociation and the chaperoning ability of the ribosome was also studied.
The dissociation of ribosomal subunits is an important step in protein translation. Ribosome recycling in the post termination stage of translation is ensured by cooperative action of three translation factors: ribosome recycling factor (RRF), elongation factor G (EFG), and initiation factor 3 (IF3). The mechanism of this process that is relevant to our studies outlined below.

**Ribosome recycling**

The post termination complex (post-TC) consists of mRNA with the termination codon at the A-site, tRNAs and the ribosome. The final step of translation, the ribosome recycling is the disassembly of the post-termination ribosomal complex that is an active process catalyzed by a novel protein called the ribosome recycling factor (RRF) [Hirashima A et al, 1972]. RRF functions precisely in the presence of elongation factor G (EFG) and GTP [Janosi L et al, 1994]. The gene encoding RRF (frr) is essential for bacteria. Its deletion causes unscheduled translation termination.

There are two different models that discuss the RRF and EFG mediated disassembly of post termination complex [Karimi R et al, 1999; Peske F et al, 2005; Hirokawa G et al, 2005]. According to the model 1, RRF and EFG catalyzes both the dissociation of ribosome into subunits as well as release of mRNA and tRNA from the ribosome [Hirokawa G et al, 2005]. By contrast, in model 2, RRF and EFG are only responsible for the dissociation of the 70S ribosome into subunits and initiation factor 3 (IF3) is required for the release of tRNA whereas the mRNA is released spontaneously [Peske F et al, 2005]. The dissociation of 70S ribosomes into subunits has been studied by various independent biochemical methods: (a) fluorescence resonance energy transfer (FRET) change between fluorescence-labeled subunits [Peske F et al, 2005], (b) ribosomal subunit exchange between the radiolabeled ribosomal complex and free subunits [Zavialov AV et al, 2005], (c) decrease in the light-scattering with time during ribosome
dissociation into subunits [Hirokawa G et al, 2005], (d) sucrose-density-gradient centrifugation (SDGC) to differentiate the whole ribosomal particle from its subunits [Hirashima A et al, 1972; Hirokawa G et al, 2002].

It has been reported that application of the SDGC technique permitted the transiently dissociated subunits to re-associate to form 70S ribosomes even in the presence of RRF and EF-G and explaining why IF3 is required to stabilize the dissociated subunits. Based on various biochemical studies and structural data, model 2 is more acceptable by different individual groups throughout the world. The following is the suggested mechanism of ribosome recycling. RRF and EF-G.GTP bind to the post termination complex and after GTP hydrolysis, promote the dissociation of the 50S subunit from the 30S-tRNA-mRNA complex. The deacylated tRNA as well as mRNA remain bound to the 30S subunit in a rather stable fashion. In the second step, IF3 stimulates tRNA dissociation from the 30S subunit, followed by the release of mRNA. The resulting ribosomal subunits are free for the next round of translation [Peske F et al, 2005].
**Figure 1.11: Ribosome recycling** Model 1: RRF binds to the model post-TC containing deacylated tRNA in the P-site, followed by binding of EFG-GTP resulting in tRNA release. This is then followed by subunit dissociation and mRNA release. The transiently dissociated subunits are kept separated by IF3 binding to the 30S subunit. In the absence of IF3, the subunits re-associate to form 70S ribosomes that may again be dissociate to subunits by RRF and EFG. [Picture adapted from Seshadri A et al, 2006].

**Figure 1.12: Ribosome recycling** Model 2: RRF binds to the tRNA containing post-TC. Binding of EFG-GTP leads to subunit dissociation. IF3 prevents the re-association of the separated subunits, and also aids in the release of tRNA and mRNA from the 30S subunit [Picture adapted from Seshadri A et al, 2006].
Ribosome recycling: Interplay of Factors

Understanding of three-dimensional structure of RRF [Selmer M et al, 1999; Saikrishnan K et al, 2005; Nakano H et al, 2003]; and its complex formation with 50S subunit or whole 70S ribosome [Agrawal RK et al, 2004; Wilson DN et al, 2005]; and the studies on the role of initiation factor 3 (IF3) in ribosome recycling [Karimi R et al, 1999; Peske F et al, 2005; Hirokawa G et al, 2005; Singh NS et al, 2005] have simplified our thoughts about the mechanism of ribosome recycling. These studies also bridged the gap between the termination and the initiation steps of protein synthesis.

RRF makes extensive interactions with helices 69 and 71 of the 23S rRNA [Agrawal RK et al, 2004]. These two helices are the component of major inter-subunit bridges, B2a and B3 respectively [Cate JH et al, 1999; Gao H et al, 2003]. RRF binding to the ribosome induces conformational changes of these bridges that split the ribosome into subunits.

Figure 1.13: Intersubunit bridges. Contacts at the interface of 30S (a) and 50S (b) subunits, colour coded by type of interaction. Interaction with protein in the opposite subunit for bridge formation is shown in gold and interaction with RNA, red [Picture adapted from Schuwirth BS et al, 2005].
Crystal and solution structures of RRF from several different organisms show that it is composed of two domains that are linked to form an “L” shaped structure, similar to that of tRNA. Domain I is a three-helix bundle (α1, α3 and α4), whereas relatively mobile domain II has a three-layered β/α/β structure. Domain II is connected to the Domain I by a highly flexible linker sequence [Pai RD et al, 2008]. This near perfect mimicry of tRNA was thought to provide clues in elucidating a possible mechanism of action for RRF. Initially it had been proposed that since the overall shape of RRF mimics tRNA structure and therefore RRF binds to the tRNA binding sites on the ribosome and dissociates ribosomal subunits through EFG catalyzed translocation. Later hydroxyl radical probing data of 16S and 23S rRNA and cryo-EM studies shows that the orientation of RRF in the ribosome is different as proposed by the tRNA-mimic model.

Crystal structure of RRF with the E.coli ribosome shows that the domain I of RRF interacts with the large subunit of the ribosome and domain II is crucial in the function of RRF and EFG during recycling. These studies of RRF bound to the empty ribosome showed that RRF is positioned in the cleft of the 50S subunit that contains the peptidyl transferase center, and is close to two key elements of the ribosome—23S rRNA helix H69 of the 50S subunit and 16S rRNA helix h44 of the 30S subunit [Pai RD et al, 2008]. The position of RRF binding to the ribosome was revealed in more detail by a 3.3 Å crystal structure of an RRF variant containing only Domain I bound to the *Deinococcus radiodurans* 50S subunit. This structure revealed that RRF binding hinders the positioning of helix H69, which interacts with helix h44 to form intersubunit bridge B2a. In biochemical experiments, RRF binds with 6-fold higher affinity to the 70S ribosome compared to the 50S subunit, suggesting that stable RRF association with the isolated 50S subunit may not have any physiological role. Furthermore, RRF bound to the 50S subunit is not released by EFG [Pai RD et al, 2008].
Specific interactions between RRF and EFG are needed for ribosome recycling [Rao AR et al, 2001] and it is clear that the role of EFG in ribosome recycling is different from its classical role in ribosome translocation. Using the energy of ATP hydrolysis, EFG induces a conformational rearrangement of the ribosome which is the main driving force for tRNA-mRNA translocation. The translocation activity of EFG results in the release of tRNA, which is followed by the release of mRNA and 70S ribosome from the Post-termination complex.

EFG converts a high affinity binding of RRF to ribosome to a low affinity binding. This change in the affinity of RRF binding has been interpreted due to EFG mediated movement of RRF on ribosome [Kiel MC et al, 2003]. This movement of RRF probably leads to release tRNA, disruption of the inter-subunit bridges (B2a, B3), mRNA release and dissociation of ribosome into two subunits, which are then kept separated by binding of IF3 to the 30S subunit [Hirokawa G et al, 2005].

**Figure 1.14: Ribosome recycling with RRF and EFG.** a) In ribosome recycling process EFG and RRF function together to dissociate the post termination complex (PoTC) into subunits. b) Schematic representation of the catalytic role of EFG loop II in ribosome recycling. At the end of translation, RRF binds to post termination complex (PoTC) inducing the conformational change of H69 and h44 to destabilize B2a. Upon EFG•GTP binding, loop II (blue) locates in the vicinity of B2a. Together with RRF, loop II disassembles the ribosome into the subunits [Picture adapted from Zhang D et al, 2015].
Recent studies also propose that there are two major pathways for ribosome recycling: the initiation factors IF1-IF3 mediated ribosome recycling and RRF-EFG induced recycling into subunit inside cell. Ribosome containing a mRNA with strong SD sequence are split/recycled rapidly by IF1 and IF3, but very slowly with RRF and EFG. Post termination complex having mRNA and a P-site bound deacylated t-RNA are split very rapidly by RRF and EFG, but slowly with IF1 and IF3. Vacant ribosomes are split by RRF/EFG much more slowly than ribosomes in post termination complex and by IF1/IF3 much more slowly than mRNA-containing ribosomes [Pavlov MY et al, 2008].

Our studies on a) the effect of partially or intrinsically unstructured proteins b) PTC and aminoglycoside antibiotics on ribosome subunit dissociation and re-association process are outlined in chapter 3.

III.3. Objective 3: Effect of ribosome on protein aggregation systems

The aggregation of lysozyme and insulin under reducing conditions is well documented and is often used a model system for protein aggregation studies [Tomoyasu T et al, 2010; Bhattacharyya J et al, 1998]. Our objective was to study the effect of the E. coli ribosome and the domain V of its 23S rRNA on a) aggregation of lysozyme and insulin under reducing conditions b) aggregation of the non-disulphide containing protein, bovine carbonic anhydrase II (that is known to form aggregation prone molten globule like intermediate during its refolding).

In this context, a recent study has demonstrating that the antiprion drugs 6-aminophenanthridine and guanabenz can significantly reduce ribosome’s chaperoning ability by selectively binding to the specific positions on the RNA1 sub-domain that are necessary for interaction with the unfolded proteins [Tribouillard-Tanvier D et al, 2008; Pang Y et al, 2013], needs further discussion.
Two structurally unrelated drugs 6-Aminophenanthridine (6AP) and Guanabenz (GA), that are used for the treatment of hypertension, were also shown to possess antiprion activity using a yeast-based assay and in the *in vivo* mouse model for prion-based diseases.

In studies that were conducted on the identification of cellular targets of these drugs using affinity chromatography matrices, it was demonstrated that 6AP and GA do not interact with the prion protein but interact with the ribosomal RNA of the ribosome. While these drugs had no effect on the peptidyl transferase activity of the ribosome or on global translation, specific inhibition of the ribosomal RNA-mediated protein folding activity of the ribosome was observed. Since these drugs inhibit ribosome’s protein folding ability, it is expected that their binding site should coincide with the chaperoning center of the ribosome residing in the domain V of the 23S rRNA (see section I). Studies were performed on the exact binding sites of 6AP to domain V rRNA using UV cross-linking followed by primer extension assay showed that the major 6AP binding sites. Earlier studies had demonstrated that three unrelated substrate proteins (human carbonic anhydrase, bovine carbonic anhydrase and dihydrofolate reductase) when refolded in presence of ribosome or the domain V RNA interact with five specific sites in the RNA1 subdomain. There is a major overlap between the protein binding sites and 6AP interaction sites on domain V rRNA (Figure 1.15).
Figure 1.15: Interaction sites of antiprion drug on ribosome. a) Binding sites of 6AP and the protein folding substrates in domain V rRNA. The green and the black boxes indicate 6AP and the protein binding sites, respectively. b) The position of the binding sites in the 50S subunit of *E. coli* ribosome. (PDB ID: 3UOS) The red highlighted portion represents residues 2435–2668 of domain V rRNA where most of the binding sites (marked with blue spheres and labeled with residue numbers) are located [Picture adapted from Banerjee D et al, 2004].
These studies explained the competitive inhibition of ribosome’s chaperoning ability in presence of these drugs that was further confirmed by mutagenesis assays. The interaction of these drugs with the rRNA domain is also believed to be significant for the implementation of their antiprion activity, although the exact mechanism still needs to be elucidated. This finding also highlights the importance of further studies that are necessary to elucidate the effect of ribosome and its 23S rRNA domain on the process of protein aggregation. Our study on BCAII and lysozyme aggregation under different conditions (stated in Chapter 4) is a continuation of our investigation on the effect of ribosomal chaperone on aggregation of partially folded BCAII and lysozyme (chapter 2). We have also included bovine insulin as a model protein in these studies. It is known that in the rare medical condition termed injection amyloidosis, full-length insulin molecules are found in fibrillar form at the site of frequent insulin injections [Dische FE et al, 1988]. Also, serum samples from patients with Parkinson’s disease have been found to display an autoimmune response to insulin oligomers and fibrils [Wilhelm KR et al, 2007], possibly indicating the presence of insulin aggregates in this disease as well. 

Our studies on the effect of the ribosomal chaperone on the process of lysozyme and insulin aggregation (under reducing conditions) and carbonic anhydrase aggregation (during refolding at high concentrations) are discussed in chapter 4 of the thesis.