

Dissertation zur Erlangung des Doktorgrades
der Fakultät für Chemie und Pharmazie
der Ludwig-Maximilians-Universität München

**Function of the *E. coli* chaperone Trigger factor - Role in nascent
chain binding and folding delay of multi-domain proteins**

Rashmi Gupta

aus

New Delhi

India

2010

Erklärung

Diese Dissertation wurde im Sinne von § 13 Abs. 3 bzw. 4 der Promotionsordnung vom 29. Januar 1998 von Herrn Professor Dr. F. Ulrich Hartl betreut.

Ehrenwörtliche Versicherung

Diese Dissertation wurde selbständig, ohne unerlaubte Hilfen erarbeitet.

München, am 19.10.2010

.....
Rashmi Gupta

Dissertation eingereicht am	19.10.2010
1. Gutachter:	Professor Dr. F. Ulrich Hartl
2. Gutachter:	PD Dr. Konstanze Winklhofer
Mündliche Prüfung am	02.12.2010

Acknowledgements

I am extremely grateful to Prof. F. Ulrich Hartl for his support and guidance throughout the course of this study. His constant encouragement and discussions have been extremely crucial for the development of this project and thesis. I would also like to thank Dr. Manajit Hayer-Hartl for her ideas and advice.

I would like to thank PD. Dr. Konstanze Winklhofer for being a co-referee of my thesis committee.

I sincerely thank Dr. Hung-Chun Chang for his enthusiasm in mentoring me during the initial phases of this work and Dr. Stephanie A. Etchells for helping me take this work to completion. Their constant involvement and motivation have been significantly important for this thesis.

I take this opportunity to thank Dr. Kausik Chakraborty and Dr. Martin Vabulas for always being there to answer my queries and providing suggestions related to this work. I would like to mention the help and suggestions provided by Dr. Raluca Antonoae, Dr. Florian Brandt and Florian Ruessmann during the development of this work. To Shruti, Jyoti, Rajat and Niti for always being there with encouragement during the happy as well as the tough times.

My heartfelt gratitude to the Department office especially Andrea for her kind help throughout my stay in Munich. Thanks to all the technical staff and members of the department for making this journey smooth and memorable.

This work could not have been possible without the constant support and encouragement of my colleague, project-partner and husband Sathish. No amount of words could express my gratefulness to him for always being there for me during this adventurous journey.

Most importantly, to my parents and siblings in Delhi whose unwavering belief in me has made it possible for me to come so this far.

I SUMMARY	1
II INTRODUCTION	3
II.1 Proteins and the process of translation	3
II.2 Organization of protein structure	3
II.3 Principles and mechanisms of protein folding	6
II.4 Protein folding and aggregation inside the cell	8
II.4.1 Molecular chaperones are conserved in every kingdom of life	9
II.4.2 Cylindrical chaperones: The Chaperonins	10
II.4.2.1 Group I Chaperonins	11
II.4.2.2 Group II chaperonins	12
II.4.3 The Hsp70 Chaperone system	13
II.4.3.1 Regulation of DnaK cycle	14
II.4.3.2 Regulation of Hsp70 cycle in eukaryotes	15
II.4.4 Ribosome-associated chaperones	16
II.4.4.1 Nascent chain Associated Complex (NAC)	17
II.4.4.2 Ribosome Associated Complex (RAC)	18
II.4.5 Prokaryotic ribosome associated factors	19
II.5 Trigger factor	20
II.5.1 Model of TF function based on various crystal structures	21
II.5.2 Model of TF function based on various biochemical studies	23
II.5.3 Functional significance of the PPIase domain of TF	26
II.5.4 TF cooperates with DnaK in <i>de novo</i> folding	26
II.6 Folding of multi-domain proteins: Co-translational and post-translational folding mechanisms	27
II.7 Folding of eukaryotic proteins in <i>E. coli</i>	28
III AIM OF THE STUDY	31
IV MATERIALS AND METHODS	32
IV.1 Materials	32
IV.1.1 Chemicals	32
IV.1.2 Enzymes	34
IV.1.3 Materials	34
IV.1.4 Instruments	35
IV.1.5 Media and buffers	36
IV.1.5.1 Media	36
IV.1.5.2 Antibiotic stock solutions	36
IV.1.5.3 Buffers	36
IV.1.6 <i>E. coli</i> strains	38
IV.1.7 Antibodies	38
IV.2 Methods	38
IV.2.1 General molecular biology methods	38
IV.2.1.1 Preparation and transformation of <i>E. coli</i> competent cells	38
IV.2.1.2 Plasmid purification	39
IV.2.1.3 PCR amplification	40
IV.2.1.4 DNA restriction digestion and ligation	41
IV.2.1.5 DNA analytical methods	41
IV.2.2 Cloning of various constructs for this study	41

IV.2.3 Site directed mutagenesis	44
IV.2.4 Protein purification methods	45
IV.2.4.1 Purification of TF and TFNC and their cysteine mutants	45
IV.2.4.2 Purification of eukaryotic chaperones	46
IV.2.5 Protein analytical methods	46
IV.2.5.1 Determination of protein concentration	46
IV.2.5.2 SDS-PAGE (sodium-dodecylsulfate polyacrylamide gel electrophoresis)	46
IV.2.5.3 Autoradiography	47
IV.2.5.4 Western blotting	48
IV.2.6 <i>In vitro</i> assays	48
IV.2.6.1 Site specific labeling of single cysteine TF proteins	48
IV.2.6.2 <i>In vitro</i> translation in the PURE system	49
IV.2.6.3 Separation of ribosome-chaperone complexes	49
IV.2.6.4 Post-translational folding of Luc	50
IV.2.6.5 Limited proteolysis of Ras-DHFR	50
IV.2.7 <i>In vivo</i> co-expression experiments	50
IV.2.8 Luciferase specific activity and solubility measurements	51
IV.2.9 Fluorescence measurements	51
V RESULTS	53
V.1 Analysis of TF-nascent chain interactions using real-time fluorescence spectroscopy	53
V.1.1 TF is recruited to nascent chains exposing hydrophobic motifs	54
V.1.2 Additional TF molecules are recruited towards elongating Luc nascent chains exposing hydrophobic segments	56
V.1.3 Dissociation of TF from Luc RNCs depends on the location of the fluorophore	58
V.1.4 TF dissociates from different hydrophobic regions of Luc with different rates	61
V.2 The PPIase domain of TF delays the folding of eukaryotic multi-domain proteins relative to their translation	63
V.2.1 Dissociation of TFNC150-NBD from Luc-RNCs is faster than that of TF150-NBD	64
V.2.2 TFNC improves the folding of Luc in PURE system	66
V.2.3 TFNC improves the folding of Luc in the Rapid Translation System (RTS)	67
V.2.4 TFNC-mediated folding is more co-translational	68
V.2.5 TFNC improves the folding of Luc in Δ <i>tig E. coli</i> cells	70
V.2.6 Folding of Ras-DHFR in presence of TF and TFNC	71
V.2.6.1 TFNC improves the solubility of Ras-DHFR in Δ <i>tig E. coli</i> MC4100	72
V.2.6.2 TFNC mediated efficient co-translational folding of Ras-DHFR	73
V.2.6.3 Effect of TF and TFNC on the folding of the Ras domain	75
V.3 Eukaryotic chaperones mediate efficient folding of Luc	76
V.3.1 Eukaryotic <i>in vitro</i> translation system is capable of efficient Luc folding	77
V.3.2 Supplementation of purified eukaryotic chaperones in PURE system enhance the folding of Luc	78
V.3.3 Eukaryotic Hsp70 system promotes efficient Luc folding in PURE system	81
V.3.4 The eukaryotic Hsp70 machinery mediates efficient co-translational folding of Luc in PURE system	83
V.4 Generation of an <i>in vivo</i> co-expression system to test folding of multi-domain proteins in presence of eukaryotic chaperones	84

V.4.1 Effect of TFNC on the folding of Luc in BL21 (DE3) cells	85
V.4.2 Effect of Hsc70 on the folding of Luc in BL21 (DE3) cells	86
V.4.3 Imbalance in the expression levels of eukaryotic Hsp70 chaperone machinery with TFNC causes inhibition of Luc translation	88
V.4.4 Hdj2 is responsible for the inhibition of Luc translation <i>in vivo</i>	89
VI DISCUSSION	91
VI.1 TF interaction with nascent chains is based on the presence of hydrophobic motifs in the primary sequence	92
VI.2 Complex kinetics of TF dissociation from Luc nascent chains	93
VI.3 TFNC is more efficient than TF in assisting multi-domain protein folding	94
VI.4 TFNC-mediated folding pathway of Luc and Ras-DHFR is more co-translational	94
VI.5 The eukaryotic Hsp70 chaperone system mediates efficient folding of Luc	96
VI.6 Generation of a co-expression system to test the folding of different aggregation-prone multi-domain proteins in presence of various chaperones in <i>E. coli</i>	97
VI.7 Perspectives	98
VII REFERENCES	99
VIII APPENDICES	111
VIII.1 Abbreviations	111
VIII.2 Plasmid maps	114
VIII.3 Publications	118
VIII.4 <i>Curriculum vitae</i>	119

I SUMMARY

Heterologous protein production in *E. coli* often results in protein aggregation arising due to the high local concentrations of fast translating polypeptides in the dense, crowded environment of the bacterial cytosol (Ellis and Minton, 2006). Chaperone co-expression is one of the commonly used methods to alleviate protein insolubility and aggregation as chaperones shield aggregation prone regions of translating nascent chains and maintain them in a folding competent state until productive folding has occurred (Baneyx and Mujacic, 2004). The *E. coli* proteome contains mostly small proteins with an average length of 317 amino acids (Netzer and Hartl, 1997) and hence the *E. coli* chaperone machinery may not be optimally adapted to the folding of large multi-domain eukaryotic proteins (Agashe et al., 2004). Trigger factor (TF) is the major ribosome-associated chaperone in *E. coli* and has homologs present only in eubacteria. Thus, TF is of particular interest for understanding the difference between the bacterial and eukaryotic folding environment. It is a modular protein composed of three domains: the N-terminal ribosome binding domain, the middle PPIase domain and the C-terminal chaperone domain which has the main chaperoning activity. The PPIase domain is dispensable for TF function *in vivo* but provides a secondary nascent chain binding site (Kaiser et al., 2006; Lakshmipathy et al., 2007).

In this study, we studied the direct interaction of TF with nascent chains and the contribution made by its domains in this interaction, using real-time fluorescence spectroscopy. TF was site-specifically labeled with an environment sensitive probe, NBD, and the increase in its fluorescence, which signified the binding of TF to hydrophobic motifs of the nascent chains, was monitored. We found that TF specifically interacts with nascent chains exposing hydrophobic regions like Luc and not with nascent chains that lacked such regions like α -Syn. We also found that multiple TF molecules bind to elongating Luc nascent chains exposing an increasing number of hydrophobic regions. The dissociation of TF from such nascent chains occurred with more than one phase. We also demonstrate that the PPIase domain of TF provides a secondary binding site for nascent chains and can interact with only a fraction of nascent chains exposing hydrophobic regions. The PPIase domain deletion mutant of TF, called TFNC,

dissociated faster from full length Luc indicating that TFNC resides on these nascent chains for a shorter time.

TF delays the folding of certain multi-domain proteins in *E. coli* relative to their translation (Agashe et al., 2004). Hence we analyzed the chaperoning effect of TFNC on the folding of multi-domain proteins like Luc and Ras-DHFR and found that TFNC enhances the folding of these proteins compared to TF via a more co-translational folding mechanism. We imply that by binding to the hydrophobic regions of the nascent chains, the PPIase domain delays the folding of multi-domain proteins in *E. coli* relative to their translation. This deletion is beneficial for the folding of these proteins that rely on domain-wise co-translational folding.

In the second part of the thesis, we examined the effect of eukaryotic chaperones on Luc folding in the bacterial PURE system and found that purified human Hsc70, Hdj2 and Bag1 in a molar ratio of 10:1:6 mediate the efficient folding of Luc. Supplementation of TFNC to this combination enhanced Luc folding further by allowing the nascent chains to fold co-translationally during translation.

II INTRODUCTION

II.1 Proteins and the process of translation

Proteins are the important building blocks of life. They are linear polymers built from a series of 20 different L-amino acids. The amino acids in a polypeptide chain are linked by peptide bonds between the carboxyl and amino groups. The process of polypeptide synthesis is called translation in which the genetic code on the mRNA is converted into the polypeptide. Cells have specialized macromolecular complexes called ribosomes made up of protein and RNA where the process of translation takes place.

Nascent polypeptide chains emerging from the ribosomes have to be properly folded to reach their native state. Non-native interactions between nascent chains during on-going translation might lead to aggregation of these nascent chains. Cells have developed mechanisms to prevent such unwanted interactions between nascent chains and shift the equilibrium towards the correctly folded structure with the help of folding helpers called chaperones. Molecular chaperones are proteins that help the cellular proteins to attain their native structure and inhibit the off-pathway aggregates without being part of their final structures.

In the forthcoming sections of this introduction, the processes of protein folding and aggregation will be discussed in detail. The mechanisms by which different classes of molecular chaperones promote protein folding in the cell and their importance in various human neurodegenerative diseases as well as in the recombinant protein production in biotechnology, will also be discussed.

II.2 Organization of protein structure

There are different levels of organization of protein structure. The primary structure of the protein comprises the linear amino acid sequence in its polypeptide chain. This structure is determined by the genetic information coded in the DNA sequence.

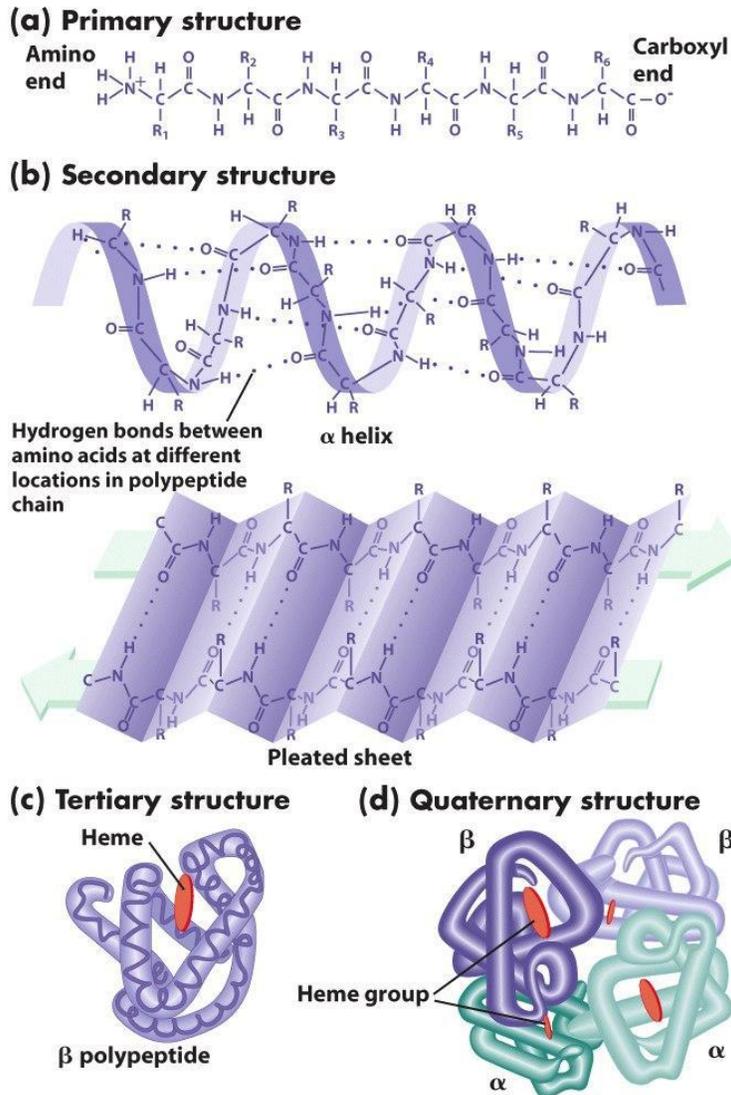


Figure 1: Levels of protein structure

Representation of primary, secondary, tertiary and quaternary structures of Hemoglobin. (a) Primary structure is represented by the amino acid sequence of the peptide chain. (b) Secondary structure comprises of highly regular sub-structures such as the α helix and β sheet. (c) Tertiary structure refers to the three-dimensional structure of a single protein molecule. (d) Quaternary structure refers to the complex of several protein molecules or polypeptide chains. Adapted from *An Introduction to Genetic Analysis* by Griffiths, Miller, Suzuki, Lewontin, and Gelbart, 2000.

The secondary structure of the protein molecules refers to the regular pattern of amino acid arrangement in the polypeptide chain due to the formation of hydrogen bonds between the amino acid main chain. The two most common types of secondary structures are called the α helix and β pleated sheet (Pauling and Corey, 1951a, b). The tertiary

structure of a protein refers to the arrangement of the helices and the sheets in space to generate a three-dimensional structure. This arrangement is stabilized by multiple weak, non-covalent interactions which include hydrogen bonds, electrostatic interactions and hydrophobic interactions. Disulphide bond formation between adjacent cysteines may also contribute to the tertiary structure. Many proteins like hemoglobin are composed of two or more polypeptide chains, adding another level of structural organization (Perutz et al., 1960). The assembly of individual polypeptide chains into a homo- or hetero oligomeric complex is called its quaternary structure.

In a polypeptide chain, the peptide bonds are planar due to their partial double bond character. The N-C_α and C_α-C bonds of the amino acids are free to rotate and are represented by the Phi (φ) and Psi (ψ) angles, respectively. These dihedral angles define the conformation of a single amino acid in a polypeptide chain. The Ramachandran plot developed by G. N. Ramachandran and V. Sasisekharan is a visual representation of all the possible φ and ψ angles for a polypeptide chain. By calculating the energy contained in various pairs of φ and ψ angles, the α and β conformations were found to be the two most stable pairs (Ramachandran and Sasisekharan, 1968). The α helix and the β sheet are also the most common conformations found in nature. Given how common these structures are, it is interesting to examine how a protein proceeds from a linear structure as it exits the ribosomal tunnel to obtain not only secondary structures but tertiary structures as well.

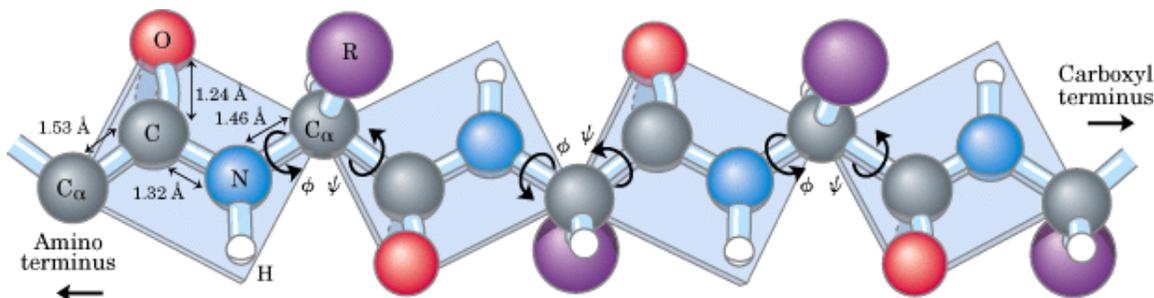


Figure 2: Torsion angles in an extended polypeptide chain

Each amino acid in a polypeptide chain contributes three bonds to its backbone. The peptide bond is planar and is represented by blue planes. The bond between nitrogen and α-carbon atoms is represented by φ and the bond between α-carbon and carbonyl

carbon atoms is represented by ψ . Adapted from Lehninger Principles of Biochemistry, Fourth Edition, 2004.

The time taken by a protein to reach its native conformation varies dramatically between different proteins. Since every bond in an amino acid can adopt numerous possible conformations, a protein could take an astronomical amount of time to sample all the possible conformations. Clearly, this time scale is beyond the time scale of any biological process and in a cell, protein folding takes place within few seconds or minutes. The discrepancy between the calculated and actual protein folding time is called the Levinthal paradox (Levinthal et al., 1962; Zwanzig et al., 1992). Levinthal proposed that the folding process occurs through a series of intermediate states and not through a random conformational search. These intermediates containing locally stabilized secondary structural elements drastically reduce the number of possible conformations available, allowing the protein to reach its native conformation in a biologically relevant time scale.

II.3 Principles and mechanisms of protein folding

Ribonuclease A, a small 14 kDa protein was used as a model protein by Christian Anfinsen to examine protein refolding *in vitro*. He postulated that all the information required by a protein to adopt its final conformation is encoded in its primary amino acid sequence. This final conformation is the polypeptide conformation with the lowest free energy. In his pioneering experiments, Anfinsen demonstrated that purified, denatured Ribonuclease A regains its native, enzymatically active state upon removal of the denaturant (Anfinsen, 1972, 1973; Taniuchi and Anfinsen, 1969). As a result of Anfinsen's work it was thought that protein folding is generally a spontaneous process.

Several protein folding mechanisms have been proposed to overcome the "Levinthal Paradox" by allowing step-wise progression of the folding process through smaller sub processes (Fersht, 2008). The initial stage of the protein folding pathway has been described by the "Hydrophobic collapse model" driven mainly by hydrophobic forces (Baldwin, 1989). The collapsed state is known as a molten globule arising due to the repulsion between the hydrophobic side chains of the protein and the hydrophilic

water molecules in the environment. Usually the molten globule state contains most of its native secondary structure but not the tertiary structure. Some of the first direct evidence for this model was found using the model protein barstar (Agashe et al., 1995).

In the “Framework model”, elements of secondary structure are formed first independently of the tertiary structure. These elements then coalesce to form the framework for the native tertiary structure (Kim and Baldwin, 1982, 1990; Ptitsyn and Rashin, 1975). The “Nucleation-condensation model” unified the features of both the hydrophobic collapse model and the framework model. It was shown for small proteins like Chymotrypsin inhibitor 2 that their folding can occur via two-state kinetics without the accumulation of intermediates where the nucleus is composed of a relatively unstable α -helix stabilized by long-range interactions (Jackson and Fersht, 1991).

The energy landscape of protein folding is represented by the “folding funnel” (Clark, 2004). It describes the progression of the polypeptide from its unfolded conformation to attain its local energy minimum. During this progression the polypeptide can form different folding intermediates and can get entrapped in local minimas called kinetic traps. These traps are formed due to the intramolecular contacts between the hydrophobic patches exposed to the hydrophilic environment in the partially folded intermediates. The number of such traps and the energy required to overcome them denotes the degree of frustration of the polypeptide chain (Onuchic, 1997). The folding pathway overlaps with the aggregation pathway which is governed by the intermolecular contacts between the polypeptide chains. Such non-native contacts can lead to the formation of aggregated species with the highly stable amyloid fibrils being the global minimum of this pathway. Such amyloid species have been shown cause toxicity in a number of human neurodegenerative diseases (Barral et al., 2004).

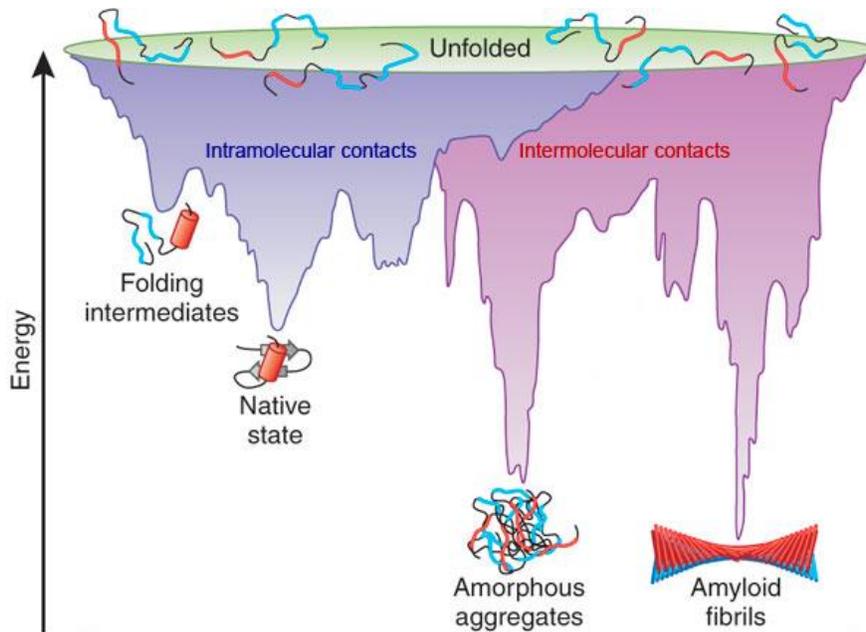


Figure 3: Energy landscape of protein folding and misfolding

Different conformations of unfolded polypeptide enter the top of the funnel and travel towards the bottom of the tunnel to attain a global energy minimum. The intermediates formed during this transition are the local energy minima. A native state is acquired following productive intramolecular interactions whereas non-productive intermolecular interactions lead to formation of aggregates. Adapted from (Hartl and Hayer-Hartl, 2009).

II.4 Protein folding and aggregation inside the cell

Protein folding in the crowded environment of the cytosol which can have a macromolecular concentration up to 400 g/liter is an enormously challenging task compared to protein folding in the diluted environment *in vitro* (Ellis and Minton, 2006). When a polypeptide chain is synthesized in a vectorial fashion on the ribosomes it does not have access to the entire folding information before the completion of translation. These unfolded nascent chains expose hydrophobic regions which have the tendency to make non-native contacts leading to inter-molecular aggregation. Moreover, the faster translation speed in *E. coli* of about ~20 amino acids per second compared to ~5 amino acids per second in eukaryotes makes the folding process in *E. coli* even more challenging. The severity of the protein aggregation problem is evident from a number of

human diseases caused by aberrant protein misfolding and aggregation (Barral et al., 2004).

It is therefore extremely important for the cell to prevent aggregation by having mechanisms that aid in proper cellular protein folding. Folding helpers called molecular chaperones are a versatile class of proteins that bind to the aggregation prone exposed hydrophobic regions and folding intermediates to prevent protein aggregation (Hartl and Hayer-Hartl, 2009; Young et al., 2004). Chaperones have diverse functions and mechanisms and they form a network in the cytosol functioning at different stages of polypeptide translation and folding.

II.4.1 Molecular chaperones are conserved in every kingdom of life

A molecular chaperone can be defined as any helper protein that binds to a partially folded aggregation prone protein and mediates attainment of the native conformation without being part of the final structure. Chaperones are involved in many cellular functions including *de novo* protein folding, refolding of proteins under stress, protein transport and degradation (Hartl and Hayer-Hartl, 2009). Chaperones called heat shock proteins (Hsps) are expressed under the cellular stress conditions whereas others are constitutively expressed. In the next sections major classes of chaperones will be discussed in detail.

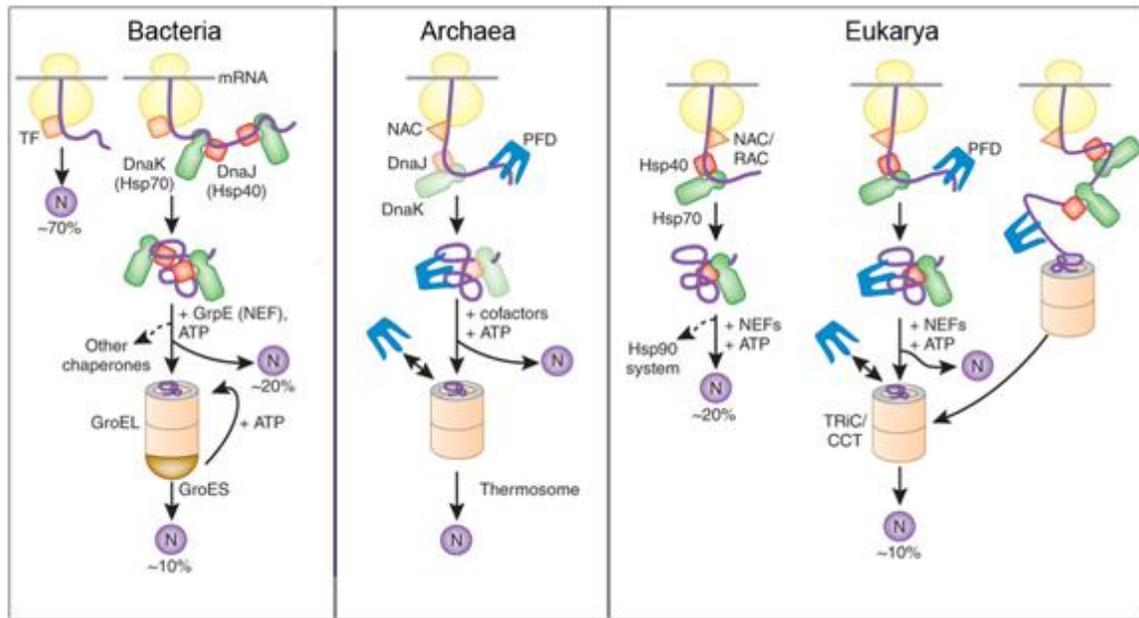


Figure 4: Models of chaperone-assisted folding in various kingdoms of life

Native protein is denoted as N. In *E. coli*, Trigger factor (TF) interacts with the translating nascent chains. Most of the smaller proteins (~70 %) are proposed to attain their native structure without further assistance. 20 % of the proteins are assisted by the DnaK chaperone machinery to fold and the remaining 10 % are transferred to the GroEL chaperonin complex to attain their final structure in an isolated environment. In Archaea, prefoldin (PFD) and the nascent chain-associated complex (NAC) interact with the elongating polypeptides. In eukaryotes, NAC and ribosome-associated complex (RAC) perform the function of a ribosome-associated chaperone like TF. The nascent chains then interact with the Hsp70 system and a part of the nascent chains are transferred to Hsp90. About 10% of chains are then co- or post-translationally transferred to the chaperonin TRiC with the assistance of PFD. Adapted from (Hartl and Hayer-Hartl, 2009).

II.4.2 Cylindrical chaperones: The Chaperonins

Chaperonins are multimeric complexes made up of two rings stacked back to back on one another that provide an isolated folding cavity for their substrates. There are two types of chaperonins Group I and Group II (Horwich et al., 2007). Group I chaperonins, also termed Hsp60s, are found in eubacteria and organelles of bacterial origin like mitochondria and chloroplast, whereas the Group II chaperonins are found in archaea and eukaryotes. The folding mechanism of both the groups is regulated by ATP binding and hydrolysis.

II.4.2.1 Group I Chaperonins

These 800 kDa complexes consist of two homo-heptameric rings stacked back to back and require the action of cochaperones of the Hsp10 family for their action. One of the most intensively studied members of this group is the *E. coli* chaperonin GroEL (Hartl, 1996; Hartl and Hayer-Hartl, 2002; Horwich et al., 2007). It functions together with its cochaperone GroES and provides an isolated cavity for a substrate protein of up to 60 kDa in size to fold in isolation.

The GroEL folding mechanism involves multiple rounds of binding and release in order to form a native structure (Brinker et al., 2001; Mayhew et al., 1996; Weissman et al., 1996). Each of the 14 subunits has two distinct domains: the equatorial ATP binding domain and the apical substrate binding domain that contains hydrophobic amino acids protruding inside the cavity. These two domains are connected to each other via a hinge like domain (Xu et al., 1997). The two rings of the GroEL chaperonin termed as the *cis* and *trans* rings are not in the same-nucleotide bound state. Binding of ATP molecules to the *cis* ring causes binding of GroES to the same ring (Langer et al., 1992b; Ranson et al., 2001). This binding leads to a conformational change in the apical domain that releases substrate into the cavity lined with hydrophilic side chains where it remains for almost 10 s as its folding process commences (Xu et al., 1997). ATP hydrolysis weakens the interaction of GroES with GroEL and signals ATP to bind to the *trans* ring. This conformational rearrangement leads to dissociation of ADP, GroES and the partially/completely folded substrate protein from the *cis* ring (Rye et al., 1997). This cycle continues as the *trans* ring becomes the new *cis* ring and another round of ATP hydrolysis helps the partially folded substrate to bury its hydrophobic regions and acquire its native state following which the correctly folded protein is released into the cytosol.

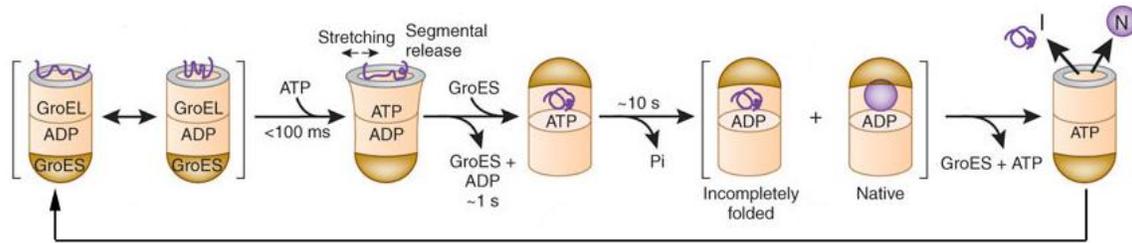


Figure 5: Protein folding cycle inside the GroEL-GroES cage

The folding intermediate is denoted by I and native protein folded inside the cage by N. Folding occurs with the help of ATP hydrolysis. Multiple rounds of chaperonin action are needed for some substrates to achieve the N state. Both I and the N state accumulate after each round of binding and release. The substrate leaves the cage after forming the N state and the I state is rebound by GroEL and the chaperonin cycle might continue until the N state is attained. Adapted from (Hartl and Hayer-Hartl, 2009).

II.4.2.2 Group II chaperonins

The members of this group are much more divergent in sequence and structure from the Group I chaperonins and are thought to function without any cochaperones. The Group II chaperonin present in the eukaryotic cytosol is called TRiC (Tailless Complex Polypeptide-1 [TCP-1] Ring Complex) or CCT (Chaperonin Containing TCP-1). The Group II chaperonin present in archaea is called the Thermosome.

TRiC is composed of eight different subunits per ring and has two such rings stacked back to back on each other (Frydman et al., 1992; Gao et al., 1992). The folding process mediated by TRiC is not clearly understood but is an absolute requirement for the folding of essential proteins like actin and tubulin. This chaperonin complex binds to actin and tubulin both *in vitro* and *in vivo* facilitating their folding (Frydman et al., 1992; Gao et al., 1992; Stemp et al., 2005; Sternlicht et al., 1993; Thulasiraman et al., 1999; Yaffe et al., 1992). Proteomic studies have revealed that the substrate repertoire of TRiC consists mainly of tryptophan-aspartic acid repeats (WD-40) containing proteins (Ho et al., 2002). Surprisingly, most of these substrates exceed the TRiC cavity size. It is speculated that TRiC mediates the folding of individual domains in a sequential manner. The folding mechanism of TRiC is governed by binding and hydrolysis of ATP, similar to the GroEL chaperonin system.

II.4.3 The Hsp70 Chaperone system

Members of the Hsp70 family exist in all kingdoms of life as well as in cellular organelles like mitochondria, chloroplasts and the endoplasmic reticulum in both constitutively expressed and stress inducible forms (Mayer and Bukau, 2005). The most extensively studied member of this family is the DnaK chaperone present in *E. coli*. This 70 kDa chaperone is regulated by ATP hydrolysis and cochaperones (DnaJ and GrpE). DnaK consists of two domains, an N-terminal ATPase domain of ~45 kDa (Flaherty et al., 1990) and a C-terminal substrate binding domain of ~25 kDa (Zhu et al., 1996). The two domains are separated by a small linker that couples the ATP hydrolysis by ATPase domain to the opening and closing of the substrate binding domain. The Hsp70 mediated folding mechanism involves multiple cycles of ATP hydrolysis and substrate release.

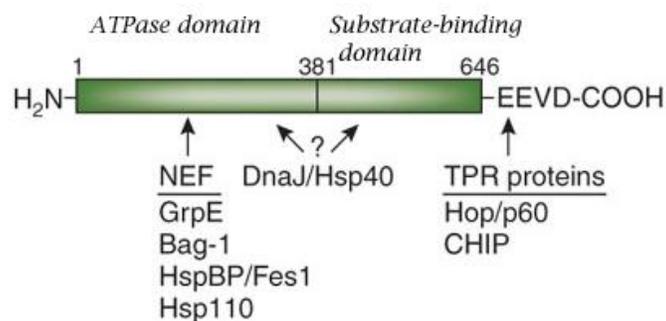


Figure 6: Linear representation of Hsp70 and its interaction with various cofactors

NEFs (Nucleotide Exchange Factors) bind to the ATPase domain of Hsp70 whereas Hsp40 probably interacts with both the ATPase and substrate-binding domain. Hsp70s of the eukaryotic cytosol contain the C-terminal EEVD sequence which interacts with tetratricopeptide (TPR) containing proteins like Hop and CHIP. Adapted from (Hartl and Hayer-Hartl, 2009).

The J domain of DnaJ is a conserved characteristic of all J domain proteins. There are 3 subfamilies of DnaJ proteins (Vos et al., 2008). The type I subfamily also called DNAJA contains the N-terminal J domain, the glycine/phenylalanine-rich region, a cysteine-rich zinc binding domain and a variable C-terminal substrate homodimerization domain. The *E. coli* DnaJ protein is a typical representative of this subfamily. The type II, DNAJB, subfamily members contain all the above mentioned domains with the exception

of the cysteine-rich zinc binding domain. The type III DNAJC subfamily members are the most divergent of all the DnaJ proteins containing only the conserved J domain.

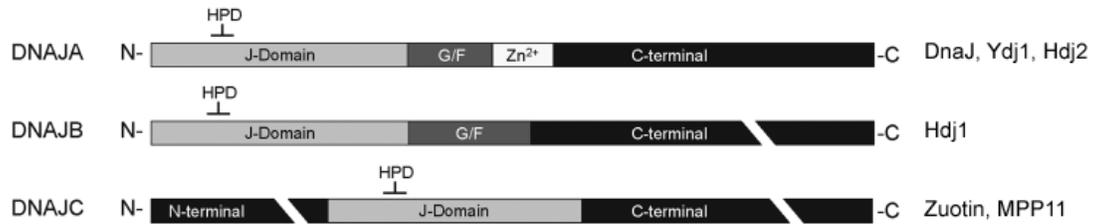


Figure 7: Linear representation of DnaJ proteins

DNAJA family contain the conserved J-domain, the Glycine-Phenylalanine (G/F) rich domain, the Zn-binding domain and the C terminal domain. Bacterial DnaJ, yeast Ydj1 and mammalian Hdj2 are examples of this family. Hdj1 is a representative of DNAJB family which contain all the above mentioned domains except Zn-binding domain. DNAJC is the most diverse DnaJ family containing only the J-domain. Yeast Zuotin and mammalian MPP11 are representatives of this family. Adapted from (Vos et al., 2008).

GrpE is a homodimeric protein that acts in concert with DnaJ to regulate the DnaK cycle. It is a NEF (Nucleotide Exchange Factor) which promotes dissociation of ADP from the nucleotide binding cleft (Harrison et al., 1997) and subsequent substrate release by stabilizing the open conformation of the ATPase domain of DnaK. The *E. coli* GrpE also possesses a putative “thermosensing” function at higher temperatures in *E. coli* but this function has not been studied in detail yet (Gelinas et al., 2003). In eukaryotes, nucleotide exchange is performed by proteins structurally unrelated to GrpE. There are three types of NEFs found in eukaryotes: Bag domain proteins, HspBP1 (Hsp Binding Protein 1) and Hsp110. The NEF activity of Bag proteins lies in their C-terminal Bag domains (C-Bag). This domain is the closest homolog of prokaryotic GrpE in terms of its nucleotide exchange action on Hsc70 although they are structurally different (Hohfeld and Jentsch, 1997). However, the mechanism of action of HspBP1 and the Hsp70 homolog Hsp110 is distinct from the nucleotide exchange action of GrpE (Dragovic et al., 2006; Shomura et al., 2005).

II.4.3.1 Regulation of DnaK cycle

DnaK is the most extensively studied of the Hsp70 family of chaperones. The ATP bound state of DnaK has a low affinity for substrates whereas the ADP bound state

has high affinity for substrates (Pierpaoli et al., 1997; Theysen et al., 1996). The cochaperone DnaJ recognizes hydrophobic segments on the nascent chains and targets them to the substrate-binding domain of DnaK (Langer et al., 1992a). Binding of DnaJ to DnaK leads to ATP hydrolysis resulting in closing of the α -helical lid (Suh et al., 1999). ADP release is then facilitated by GrpE (Harrison et al., 1997), another DnaK cochaperone. This cycle is completed by the rebinding of ATP, opening of the lid and release of the folded substrate.

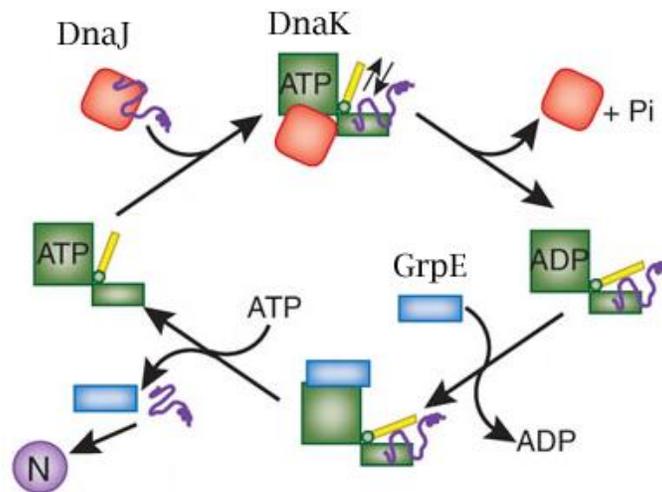


Figure 8: Reaction cycle of the DnaK chaperone machinery

Native protein is denoted as N. Rapid peptide binding and dissociation occurs in the ATP-bound state of DnaK. The substrate is delivered by DnaJ to the ATP-bound state of DnaK. The N-terminal J domain of DnaJ accelerates ATP hydrolysis by DnaK leading to stable substrate binding by DnaK. GrpE induces ADP release from DnaK and upon ATP rebinding by DnaK, substrate dissociates from DnaK, completing the reaction cycle. Adapted from (Hartl and Hayer-Hartl, 2009).

II.4.3.2 Regulation of Hsp70 cycle in eukaryotes

The Hsp70 folding cycle is significantly more complex than that of DnaK because of the presence of multiple cofactors in different cellular compartments. For example the human genome encodes 13 different Hsp70s, 41 different DnaJ proteins, various members of NEFs apart from numerous small Hsps (Vos et al., 2008).

Different cofactors bind to the two domains of Hsp70. The J domain of DnaJ proteins, Bag-1 and Hip (Hsp70 interacting protein) bind to the ATPase domain whereas

DnaJ proteins, Hop (Hsp70-Hsp90-organizing protein) and CHIP (E3 ligase C-terminal Hsp-interacting protein) bind to the C-terminal domain of Hsp70. Hip stabilizes the ADP state of Hsp70 and thus acts antagonistically to Bag-1 which promotes the ADP-ATP exchange (Hohfeld et al., 1995). Hip stabilizes the substrate-Hsp70 complex preventing pre-mature release of substrate. Hop is a dimer in solution and interacts with the C-terminal EEVD motifs of Hsp70 and Hsp90 through its TPR domains stimulating the recycling of Hsp70 (Gross and Hessefort, 1996; Smith et al., 1993). CHIP is also a dimer in solution and competes with Hop for binding to the C-termini of Hsp70 and Hsp90. It acts as an E3 ubiquitin ligase connecting the folding machinery to proteolysis (Ballinger et al., 1999; Connell et al., 2001).

II.4.4 Ribosome-associated chaperones

Ribosome-associated chaperones provide a first line of defense against aggregation as the nascent chains emerge from the ribosome. In *E. coli*, TF is the first chaperone to interact with the translating nascent chains. Eukaryotes have two chaperone complexes that have this function: NAC (Nascent chain Associated Complex) and RAC (Ribosome Associated Complex). Both of these chaperone complexes are unrelated to each other and to TF. Recent reports have demonstrated that in addition to promoting protein folding, the chaperones RAC, NAC and the J-protein Jjj1 cooperate to facilitate ribosome assembly in *Saccharomyces cerevisiae* (Albanese et al., 2010; Koplín et al., 2010).

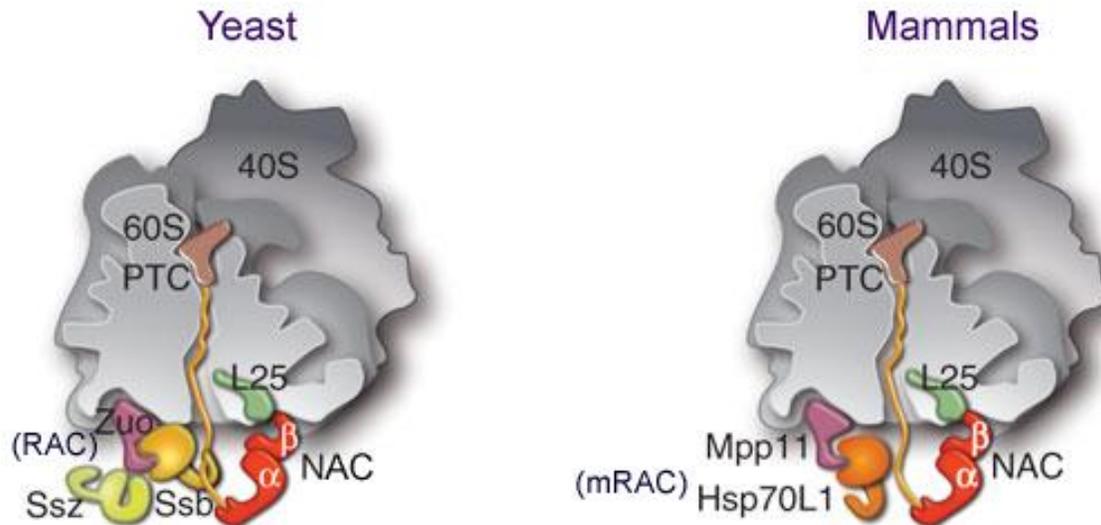


Figure 9: Model of ribosome-associated chaperones in eukaryotes

In yeast, nascent chain-associated complex (NAC) and the functional triad of Ssb/SSz/Zuo bind close to the ribosomal exit tunnel. In mammals, mRAC (ribosome-associated complex) is composed of MPP11 and Hsp70L1 that binds in close proximity to the exit tunnel together with mammalian NAC. Adapted from (Kramer et al., 2009).

II.4.4.1 Nascent chain Associated Complex (NAC)

NAC is a heterodimer of two subunits: α NAC (33 kDa) and β NAC (22 kDa). The β subunit of NAC interacts with the yeast ribosome through the ribosomal protein Rpl25, homolog of L23 in *E. coli* which is the interaction site for TF (Wegrzyn et al., 2006). Both yeast Rpl25 and *E. coli* L23 proteins are situated close to the exit tunnel in the large ribosomal subunit, an ideal location to interact with the nascent chains during translation. However, a recent report has demonstrated that β NAC binds to the yeast ribosomal protein Rpl31 which is present only in eukaryotes and archaea (Pech et al., 2010). On the yeast ribosomal exit tunnel, this adaptor site is at the opposite side of the previously described binding site Rpl25. Crosslinking studies have demonstrated that NAC interacts with ribosome-bound nascent chains (Wang et al., 1995; Wiedmann et al., 1994). NAC also has a role in the regulation of co-translational targeting of nascent chains to the endoplasmic reticulum (Lauring et al., 1995). A recent report has suggested a possible role of NAC in ribosome biogenesis (Koplin et al., 2010). In this study, the authors have reported that simultaneous deletion of NAC and Ssb led to aggregation of

ribosomal proteins and decrease in the levels of ribosomal subunits and translating ribosomes.

II.4.4.2 Ribosome Associated Complex (RAC)

RAC is a complex of two proteins Ssz1p (an Hsp70 homolog) and Zuotin (an Hsp40 homolog) (Gautschi et al., 2001). This complex acts as a J domain co-chaperone for the ribosome associated Hsp70s, Ssb1p and Ssb2p, and recruits them to the nascent polypeptides via the J domain of Zuotin (Huang et al., 2005). Together these proteins function as a chaperone triad at the yeast ribosomal exit tunnel but their exact mechanism remains to be elucidated. Deletion of any of these three components leads to a similar growth defect in yeast of cold and aminoglycoside sensitivity. The binding site of RAC at the exit tunnel has recently been identified (Peisker et al., 2008). Zuotin interacts with the ribosome through the Rpl31 ribosomal subunit. Although Rpl31 is not absolutely essential for RAC interaction with the ribosome, it might provide a structural scaffold for proper ribosome binding and function of RAC. A recent report, using amide hydrogen exchange (HX) coupled with mass spectrometry has elucidated the molecular basis of an unusual Ssz1p and Zuotin interaction (Fiaux et al., 2010). The authors reported that the C-terminus of Ssz1p and N-terminal 62 amino acids of Zuotin are involved in the assembly of RAC. This association results in the increased dynamics of the J domain of Zuotin which then activates the RAC complex to act as a co-chaperone for Ssb (Fiaux et al., 2010). Zuotin and Jjj1 (another ribosome-anchored J-protein) together with their Hsp70-partner proteins Ssb and Ssa, respectively, have recently been shown to be involved in ribosome biogenesis similar to NAC (Albanese et al., 2010).

The human homolog of RAC is called mRAC composed of MPP11 (M-Phase Phosphoprotein 11) and Hsp70L1. This complex binds to the cytosolic ribosomes in various cell types tested so far, such as HeLa, liver and kidney cells and is believed to recruit the cytosolic Hsc70 to the nascent chains. mRAC was also able to complement the growth defects caused by the deletion of yeast RAC (Hundley et al., 2005; Otto et al., 2005).

II.4.5 Prokaryotic ribosome associated factors

E. coli contains various ribosome associated factors that are involved in the processing, folding and translocation of the nascent chains. Proteins like Peptide deformylase (PDF), Signal recognition particle (SRP), Methionine aminopeptidase (MAP) and TF associate with the ribosomes close to the polypeptide exit tunnel (Kramer et al., 2009).

PDF which binds to ribosomal protein L22 removes the formyl group from the N-terminal formyl methionine of the nascent chains following which MAP cleaves the N-terminal methionine residue. A concerted action of PDF and TF has been proposed by the Ban group according to which the growing nascent chain is cradled in the folding space provided by TF. Through an opening between the arms and the ribosome binding tail, TF then “passively” routes the nascent chain to the active centre of PDF where deformylation can happen (Bingel-Erlenmeyer et al., 2008). The authors also propose that the interaction of PDF with the ribosomes enhances cell viability.

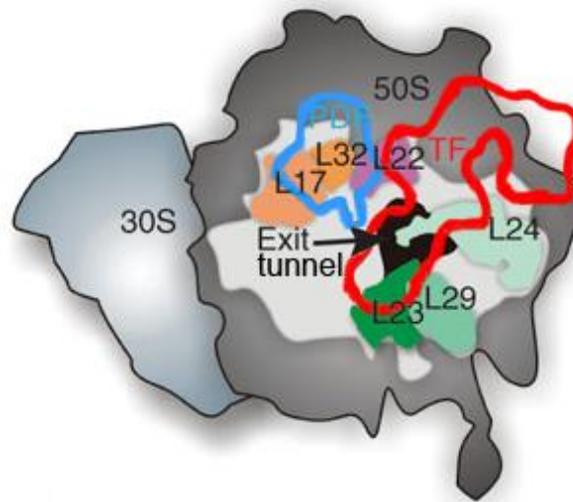


Figure 10: View of the ribosomal exit tunnel (black) of bacterial 50S subunit
The bacterial-specific ribosomal proteins L17 and L32 proteins are indicated in yellow. The ubiquitously conserved L22 (magenta), L23 (dark green), L29 (green) and L24 proteins (light green) are also indicated. The projections of PDF and TF on the ribosomal surface are shown as blue and red outlines, respectively. Adapted from (Kramer et al., 2009).

TF assists the *de novo* folding of cytosolic nascent chains. However, a part of the nascent chains translated in *E. coli* are destined for translocation. The periplasmic and outer membrane proteins are translocated post-translationally via the SecA/B pathway (Hartl et al., 1990) but the inner membrane proteins are translocated co-translationally to the SecYEG translocon by SRP and its membrane bound receptor FtsY (Macfarlane and Muller, 1995; Seluanov and Bibi, 1997). The bacterial SRP contains a protein Ffh having three domains namely N, G and M and 4.5S RNA (Poritz et al., 1990).

Interestingly both SRP and TF occupy overlapping positions on L23 making the interplay between these factors an intensively studied question (Halic et al., 2006; Kramer et al., 2002; Schaffitzel et al., 2006; Schlunzen et al., 2005). Simultaneous binding of both these factors to the *E. coli* ribosome has been demonstrated implying that they have distinct binding sites on L23 (Buskiewicz et al., 2004; Raine et al., 2004), however this co-existence has not yet been proven by a structural study. Crosslinking studies have demonstrated that these factors compete for binding to the nascent chains (Ullers et al., 2006; Ullers et al., 2003; Valent et al., 1997). It has been shown that hydrophobic signal sequences stabilize SRP but destabilize TF association with nascent chains and that SRP tends to bind to signal sequences depending on their hydrophobicity (Beck et al., 2000; Eisner et al., 2006). It has also been demonstrated that deletion of the *tig* gene accelerates protein export while overproduction of TF retards protein export (Lee and Bernstein, 2002).

II.5 Trigger factor

Like many other chaperones, TF was originally identified to be involved in a non-folding process. In the late 1980's, Wickner and co-workers found that TF maintained the proOmpA protein in an extended conformation and helped in its translocation through the outer membrane (Crooke et al., 1988a; Crooke et al., 1988b; Crooke and Wickner, 1987). This process was not affected by the depletion of TF, however the cells became filamentous both under TF depletion and overproduction (Guthrie and Wickner, 1990). Ever since then, TF has been a subject of intense analysis and has been recognized as a ribosome-associated chaperone.

TF is present only in eubacteria and there are no known homologs in eukaryotes. It is encoded by the *tig* gene and is non-essential for *E. coli* survival. The *tig* gene can be deleted in the wild type *E. coli* cells without any consequences on the normal cell growth and phenotype. But its deletion in a $\Delta dnaK$ background results in synthetic lethality at temperatures above 30 °C indicating that both these factors function in a similar cellular process (Deuerling et al., 1999; Genevaux et al., 2004; Teter et al., 1999).

TF is a modular protein of 48 kDa containing three domains namely an N-terminal ribosome binding domain, a middle peptidyl-prolyl isomerase (PPIase) domain and a C-terminal substrate binding domain. TF binds to the *E. coli* ribosome exit tunnel in a 1:1 stoichiometry via the L23, L29 subunits (Kramer et al., 2002). The cellular concentration of TF exceeds that of ribosomes (50 μ M:30 μ M) (Lill et al., 1988), hence part of its total cellular pool exists free in cytosol. The extra-ribosomal TF was proposed to form dimers in the cytosol (Kaiser et al., 2006; Patzelt et al., 2002).

II.5.1 Model of TF function based on various crystal structures

The structure of the ribosome binding domain of *E. coli* TF was solved in 2003 and it was shown to be composed of a β -sheet of four anti-parallel strands flanked by two α -helices (Kristensen and Gajhede, 2003). The main sites of interaction between TF and L23 ribosomal subunit consists of a 17 amino acid long conserved region called “TF signature” on the N-terminal domain of TF and the surface exposed region of L23 near the polypeptide exit tunnel (Kramer et al., 2002). The main ribosome-binding motif of TF is the loop region connecting the two helices comprising Phe 44, Arg 45 and Lys 46. The replacement of these amino acids by Ala (the mutant FRK/AAA) renders TF unable to bind to the ribosomes (Kramer et al., 2002). Similarly replacement of Val 16, Ser 17 and Glu 18 in L23 by Ala (the mutant VSE/AAA) prevents TF binding to the ribosomes (Kramer et al., 2002).

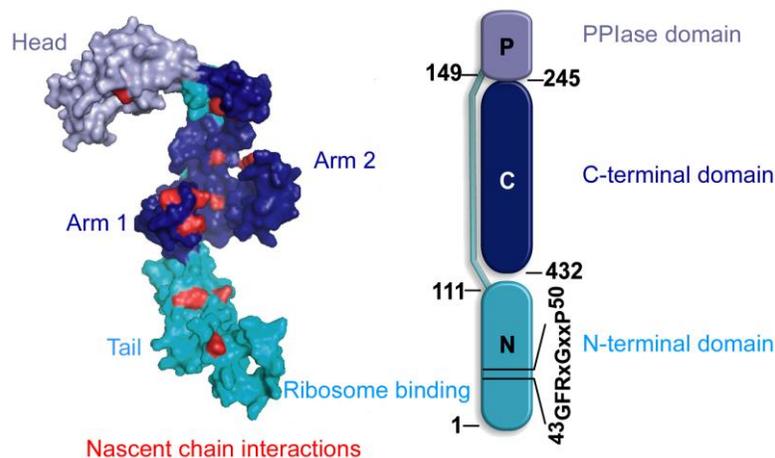


Figure 11: Structural features and domain organization of *E. coli* TF

The N-terminal domain (cyan) forms the tail of the dragon-shaped TF molecule and the ribosome-binding motif (amino acid 43-50) is denoted. The PPIase domain is shown in light blue and the C-terminal domain with two arms is shown in dark blue. The sites involved in nascent chain interaction are denoted in red. Adapted from (Hoffmann et al., 2010).

The structure of full-length TF was solved in 2004 by the Bukau and Ban groups (Ferbitz et al., 2004). TF has a unique extended "Dragon-like" conformation with the C-terminal domain forming a backbone including two arms like structures similar to the SurA chaperone present in the periplasmic space of *E. coli*. This C-domain separates the N-terminal domain and the PPIase which are at the opposite ends of the molecule. The crystal structure of C-terminal truncated TF from *Vibrio cholerae* was solved and was found to be very similar to the *E. coli* TF (Ludlam et al., 2004).

The co-crystal structure of the N-terminal domain of *E. coli* TF with the heterologous 50S ribosomal subunit of *Haloarcula marismortui* was also solved (Ferbitz et al., 2004). This co-crystal structure defined a platform on which the complete TF structure was modeled. According to this model, TF crouches over the ribosomal exit tunnel thereby providing a "folding space" large enough to accommodate a domain of ~14 kDa. The nascent chain was proposed to interact with the entire crevice formed by the N-terminal domain and the arms of C-terminal domain and also with the PPIase domain depending on their length.

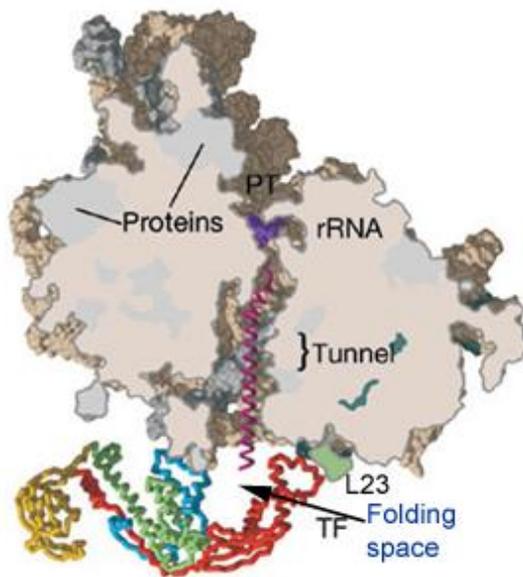


Figure 12: The “folding space” model

TF is superimposed on the ribosome-bound N-terminal fragment (residues 1-144). The ascent chain is depicted in magenta extending from the PTC and entering the folding space provided by the N-terminal domain and the arms of C-terminal domain of TF. Adapted from (Ferbitz et al., 2004).

However the model proposed by this structure may not be entirely correct because TF is absent in archaea. Later a co-crystal structure of a homologous complex of *Deinococcus radiodurans* consisting of N-terminal domain of TF and 50S subunit proposed a different model (Baram et al., 2005). It is evident from this co-crystal structure that a long extension of L24 protein (absent in archaea) would occupy the proposed "folding space" of TF, possibly hindering the folding of any nascent chain. Moreover, interactions between the L24 and TF N-domain would shift the alpha helix 2 by 40° upon ribosome binding. This large scale rearrangement of the helix would form a channel capable of guiding the nascent chains from the exit tunnel to the hydrophobic crevice.

II.5.2 Model of TF function based on various biochemical studies

Since TF is present in the cytosol in a molar excess over ribosomes, there exists an equilibrium between the ribosomal-bound monomeric form and cytosolic dimeric

form with a dissociation constant of 18 μM (Patzelt et al., 2002) and 1.8 μM (Maier et al., 2003). The function of TF domains in protein folding has been elucidated in details by two independent studies (Genevaux et al., 2004; Kramer et al., 2004b). Both the studies demonstrated that TFNC (having the N- and C-terminal domains) has the major chaperoning function *in vivo* and is enough to rescue the double knockout phenotype up to 37 °C.

The functional significance of C-terminal domain of TF was elucidated by various biochemical techniques. Limited proteolysis experiments have shown that TF protects ribosome-associated nascent chain from protease digestion (Tomic et al., 2006). According to this report the hydrophobic interactions between the C-terminal domain and the nascent chain are responsible for the protection since nascent chains of α -synuclein exposing less hydrophobic regions are not protected. TFNC provided similar levels of protease protection as the full length TF, again emphasizing that the PPIase domain has a minor role in chaperoning activity of TF.

Intramolecular FRET experiments were performed in our laboratory to understand the dynamics of TF interaction with ribosomes carrying various nascent chains (Kaiser et al., 2006). According to the model proposed in this study, TF is in rapid monomer-dimer equilibrium in the cytosol and the residence time of TF with ribosome has a half time of ~ 10 s. Upon binding to the ribosome, TF attains an open conformation as shown by the FRET experiments. This conformation prevails even after dissociation from the ribosome indicating that TF remains associated with some nascent chains with a half time upto ~ 35 s. The half time of TF relapsing back to its compact conformation depends on the properties of the nascent chain translated. It was also shown in the same study that the PPIase domain provides a secondary binding site for nascent chains, the primary site being the N- and the C- terminal domains (Kaiser et al., 2006; Lakshmipathy et al., 2007).

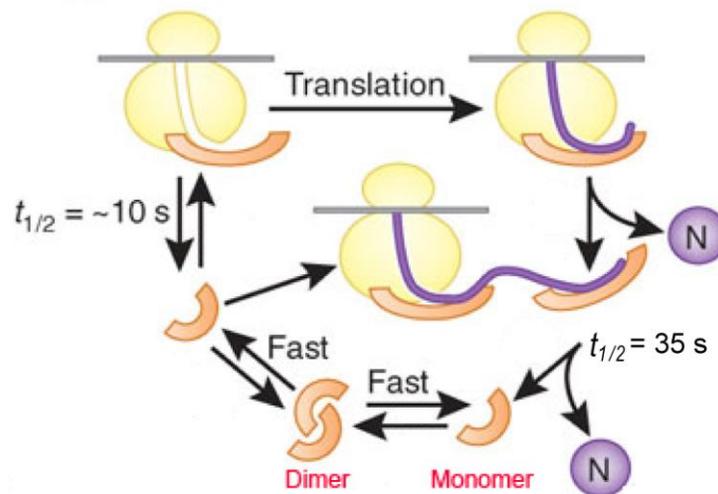


Figure 13: Model of TF function

Free monomeric TF is in fast equilibrium with the dimer form. It binds to the ribosome with a K_d of $\sim 1 \mu\text{M}$ and $t_{1/2}$ value of ~ 10 s. On binding to the ribosomes it attains an open conformation. It maintains its association with the nascent chains even after leaving the ribosome and relapses back to its compact state with a $t_{1/2}$ of ~ 35 s. Adapted from (Hartl and Hayer-Hartl, 2009).

TF-nascent chain interactions have been analyzed using site-specific photocrosslinking and real-time fluorescence experiments (Lakshmipathy et al., 2007). According to this study, the N- and the C-terminal domains interacted with all the nascent chains that were tested. The PPIase domain however showed a length-dependence in interacting with the nascent chains again indicating that it acts as a secondary binding site. Incidentally, the nascent chain binding sites also form the dimer interface (Lakshmipathy et al., 2007). Hence, it is believed that only the monomeric TF is able to bind the ribosome-nascent chain complexes whereas the dimeric TF is unable to do so since the sites required for such an interaction are buried (Kaiser et al., 2006; Lakshmipathy et al., 2007).

An alternative mechanism of TF interaction with proteins was explained in a recent study where the authors solved the structure of two *Thermotoga maritima* (*tm*) TF enclosing two folded molecules of *tm* S7 ribosomal protein (Martinez-Hackert and Hendrickson, 2009). Based on their findings, they suggest that the interactions are predominantly hydrophilic indicating a possible role of TF in ribosome biogenesis.

However, the physiological relevance of such alternative mechanisms of TF interactions with nascent chains or folded proteins remains to be explored in detail.

II.5.3 Functional significance of the PPIase domain of TF

The PPIase domain structurally resembles the FKBP (FK506 binding protein) type of PPIases (Stoller et al., 1995). It forms the “head” of TF’s “Dragon-like” conformation and is located at the opposite end of the molecule from the ribosome-binding N-domain. The functional significance of this domain has not been fully elucidated. Although the PPIase activity of TF is required for prolyl-isomerization mediated refolding of proteins like RNase T1 *in vitro* (Scholz et al., 1997; Stoller et al., 1995), it is dispensable for its chaperone function *in vivo* and its interaction with nascent chains is also independent of the presence of proline residues (Kramer et al., 2004a). Deletion of the entire PPIase domain does not significantly reduce the ability of TF to rescue the $\Delta tig \Delta dnaK$ phenotype (Genevaux et al., 2004; Kramer et al., 2004b). A few reports have suggested that the PPIase domain serves as a secondary substrate binding site in addition to the primary substrate binding site composed of N- and C-terminal domains and TFNC-nascent chain interaction is the main focus of this dissertation.

II.5.4 TF cooperates with DnaK in *de novo* folding

Unlike DnaK, the action of TF is not regulated by nucleotide binding or cochaperones. In 2004, a report examining the effects of TF and DnaK on the *de novo* folding of multi-domain proteins was published (Agashe et al., 2004). It was shown that together with TF, DnaK is involved in binding to the translating nascent chains and improve the folding yields of multi-domain proteins like Firefly Luciferase (Luc) and β -galactosidase (β -gal) in *E. coli*. Interestingly, the binding of TF and DnaK to nascent chains caused a significant delay in their folding compared to the absence of these chaperones.

Notably, DnaK and Trigger Factor (TF) have overlapping functions in chaperoning nascent chains in the bacterial cytosol. Both the chaperones recognize similar polypeptide segments on the nascent chains (Patzelt et al., 2001; Rudiger et al.,

1997). Both the chaperones can be deleted individually without causing any growth defect to the cell, however double deletion of DnaK and TF leads to severe protein aggregation and a synthetic lethal phenotype at temperatures above 30 °C (Deuerling et al., 1999). These aggregates contain mostly proteins larger than 60 kDa in size suggesting that these larger proteins ones that are known to be larger than the capacity of the GroEL chaperonin cavity rely on TF and DnaK for their proper folding (Deuerling et al., 1999). As mentioned earlier, a small ribosomal subunit S7 aggregates in the absence of TF, DnaK and DnaJ and forms a stable association with TF (Martinez-Hackert and Hendrickson, 2009).

II.6 Folding of multi-domain proteins: Co-translational and post-translational folding mechanisms

The vectorial synthesis of nascent chains on ribosomes makes the *de novo* folding pathway different from the *in vitro* refolding pathway. During the translation process the N-terminus of the nascent polypeptide is available to interact with the cytosol before its C-terminal part while during refolding the entire polypeptide is available for folding immediately after its denaturation. Thus, co-translational folding of the nascent polypeptides, especially those having multiple domains would allow them to avoid the non-native contacts with parts of the nascent chains that have yet to emerge from the ribosomal tunnel. Therefore, less aggregation of nascent chains would be predicted. On the other hand, the refolding process of such a protein is often inefficient and slow because of the presence of various unproductive intermediates. This indicates that the efficient folding of multi-domain proteins is dictated by a hierarchical folding of the individual domains.

Since eukaryotes have a slower translation rate than *E. coli* and contain more multidomain proteins, a co-translational folding mechanism would prevent non-native interactions between the incompletely folded domains. Indeed, efficient co-translational folding has been observed for some multi-domain proteins in the eukaryotic system but not in the bacterial system. It has been shown that the N-terminal 22 kDa domain of firefly luciferase (Luc) folds co-translationally when translated in eukaryotic translation

lysate (Frydman et al., 1999). This *de novo* folding of Luc proceeds through the intermediate containing the N-terminal folded domain and hence leads to faster attainment of the native structure. The *de novo* folding of Luc in *E. coli* occurs through a post-translational folding pathway and is a kinetically inefficient process (Agashe et al., 2004).

Efficient co-translational folding has also been observed for Ras-DHFR fusion protein (consisting of two domains) in the eukaryotic cell-free translation system (Netzer and Hartl, 1997). However, it folds inefficiently via post-translational folding pathway in a bacterial cell-free translation system. It has also been shown that the translation of individual Ras and DHFR domains in *E. coli* and a bacterial translation system leads to formation of native proteins.

These observations suggest that the folding machinery in the eukaryotic cell is geared towards the co-translational folding of large multi-domain proteins in order to avoid the unfavorable interactions within the translating nascent chains.

II.7 Folding of eukaryotic proteins in *E. coli*

In the past few decades, there has been enormous progress in understanding the mechanisms of various chaperones in the *E. coli* protein folding pathways. Since *E. coli* has a very well characterized genetic makeup and is amenable to large scale fermentation process, it is the most widely used expression system for production. However, heterologous protein over-expression in *E. coli* frequently leads to protein aggregation (Baneyx and Mujacic, 2004). Eukaryotic proteins are generally larger than the bacterial proteins and may require post-translational modifications to reach their native conformation which is not possible in the *E. coli* cytosol. Other reasons for recombinant protein aggregation in *E. coli* could be the differences between the cellular environment and the translation speed in *E. coli* and eukaryotes.

There are many ways proposed to alleviate the protein aggregation problem; one of them includes decreasing the translation speed by using weak promoters, decreasing the concentration of inducer and/or by using an *E. coli* strain having mutant ribosomes

that translate slowly. A recently published report explored this possibility in detail (Siller et al., 2010). The authors employed *E. coli* having ribosomes with mutations in the protein S12 of their decoding center. The translation speed in this *E. coli* strain could be modulated by the antibiotic streptomycin to enhance the folding of proteins of eukaryotic origin. They found that decreasing the translation rate in bacteria to make it similar to that of eukaryotes promotes folding of proteins like Luc, GFP, GFP-enolase and MBP-GFP (Siller et al., 2010).

Another strategy is to lower the growth temperature which not only slows down the transcription and translation rates but also decrease the extent of hydrophobic interactions, the main reason of protein aggregation. This approach however decreases the overall yield of the translated protein; hence cold inducible promoters like the main *E. coli* cold shock gene *cspA* are used for protein expression at very low temperatures. A cold induced expression system using the *cspA* promoter and the IgG-binding domain of protein A (ZZ-domain) as a fusion partner was able to improve the folding of Renilla, Gaussia and Firefly luciferases (Inouye and Sahara, 2008).

Chaperone co-expression is also a commonly used method to improve the folding of such aggregation prone proteins in *E. coli*. Chaperones like the DnaK/DnaJ/GrpE, TF and GroEL/ES have an important role in *de novo* protein folding in *E. coli*. A recently published study has showed the generation of DnaK mutants with enhanced chaperone activity by directed evolution (Aponte et al., 2010). By using a multi-parameter screening strategy, the authors improved the folding efficiency of the DnaK chaperone machinery towards two substrates; a C-terminally truncated folding deficient mutant of chloramphenicol acetyl transferase (CAT_Cd9) and Luc. One of the DnaK mutants, Cd35, improved the maximal luminescence and initial refolding rate of Luc by five and seven fold, respectively, compared to the wild type DnaK. Interestingly, most of the beneficial mutations were located in the Lid domain (LD) of DnaK. Since, LD is involved in the allosteric regulation of the DnaK cycle, the authors propose that the improved DnaK mutants have a superior ATP-dependent cycle rather than improved substrate binding, ATP binding and ATP hydrolysis.

Active aggregate solubilization is also an important process under severe protein aggregation conditions to protect the cell from the deleterious effects of these aggregates. ClpB cooperates with DnaK and the small heat shock proteins to solubilize the proteins trapped in inclusion bodies (Mogk et al., 2003). Numerous reports have demonstrated an increase in protein solubility by co-expression of chaperones individually or in combination (de Marco et al., 2007; Haacke et al., 2009). Most of these increments are protein-specific and are not general in nature. Recently, it was shown that fusion of DnaK to aggregation prone proteins like the mouse PrP, Varicella Zoster virus protein ORF21p, human GCSF and GFP resulted in soluble fusion proteins (Kyratsous et al., 2009; Ryu et al., 2008).

The reason for the inability of *E. coli* chaperones to promote efficient folding and prevent aggregation of recombinant proteins is thought to be the difference between the prokaryotic and eukaryotic chaperone mediated folding mechanisms. Since eukaryotes contain mostly multi-domain proteins, a domain-wise folding process occurring simultaneously with translation would be advantageous. We have previously shown in our lab that an artificial two-domain fusion protein requires a co-translational domain wise folding to reach its native state (Netzer and Hartl, 1997). For this reason, we set out to generate a co-translational folding environment in *E. coli* to enhance the folding of heterologous proteins.

III AIM OF THE STUDY

TF is a bacterial ribosome-associated chaperone that interacts with the nascent chains during translation and mediates their efficient folding (Agashe et al., 2004; Deuerling et al., 1999; Teter et al., 1999). To better understand the role of TF in mediating the folding of multi-domain proteins, we first analyzed in detail its direct interaction with the nascent chains using fluorescence spectroscopy in the bacterial reconstituted PURE system of translation. The role of hydrophobic motifs on the nascent chains in TF recruitment was investigated using firefly luciferase as a model substrate. TF recruitment to elongating Luc nascent chains exposing a variable number of hydrophobic motifs was also studied.

The role played by TF and its PPIase domain in the delay of folding relative to translation of multi-domain proteins such as Luc and Ras-DHFR, in *E. coli* was investigated in detail. We found that the PPIase domain acts as a secondary nascent chain binding site and contributes to the delay in folding of multi-domain proteins. Modulation of the TF-nascent chain interaction by deleting the PPIase domain (TFNC) resulted in a shift towards a co-translational folding environment in the PURE system of translation. Multi-domain proteins like Luc and Ras-DHFR that depend on domain-wise co-translational folding to attain the functional state were able to fold efficiently in the PURE system in presence of TFNC.

The second aim was to utilize the eukaryotic Hsp70 machinery comprised of Hsc70, Hdj2 and Bag1, to generate a eukaryotic-like folding environment in the PURE system of translation since these factors mediate efficient folding of Luc in RRL.

IV MATERIALS AND METHODS

IV.1 Materials

IV.1.1 Chemicals

Acetic acid	Merck
Acetone	Merck
Adenosine-triphosphate, disodium salt (ATP)	Sigma
Agarose	Biozym
Ammonium persulfate (APS)	Sigma
Ampicillin	Merck
Arabinose	Sigma
Bacto agar	Difco
Bacto trypton	Difco
Bacto yeast extract	Difco
BADAN	Invitrogen
Bovine Serum albumin (BSA)	Sigma
Bromophenol blue	Sigma
Calcium chloride	Merck
CDTA	Sigma
Chloramphenicol	Sigma
Complete EDTA-free protease inhibitor	Roche
Coomassie brilliant blue R-250	Roth
Dimethylsulfoxide (DMSO)	Merck
Dithiothreitol (DTT)	Roche
DUET plasmids	Novagen
ECL detection kit	usb

Ethanol	Merck
Ethidium bromide (EtBr)	BioRad
Ethylenediaminetetraaceticacid-sodium salt (EDTA)	Merck
Express protein labeling mix	PerkinElmer
Glucose	Sigma
Glycerol	Merck
Glycine	Roth
Guanidium hydrochloride	Sigma
HEPES	Sigma
Hydrochloric acid (37%)	Merck
IANBD	Invitrogen
Imidazol	Merck
Isopropyl- β -D-thiogalactopyranoside (IPTG)	Carl Roth GmbH
Kanamycin	Sigma
Luciferin (Potassium salt)	Promega
Magnesium chloride	Merck
Markers (DNA and protein)	Fermentas
β -mercaptoethanol	Sigma
Methanol	Merck
PIPES	Sigma
Phenyl-methyl-sulfonyl-fluoride (PMSF)	Sigma
Polyacrylamide/bisacrylamide solution 30%	Serva
Potassium hydroxide	Sigma
PURE system II classic translation system	Post Genome Institute, Japan
Puromycin	Sigma
Rabbit reticulocyte lysate (RRL)	Promega
RTS HY transcription/translation system	Roche

RNase A	Roche
Sodium chloride	Merck
Sodium dodecylsulfate (SDS)	Sigma
Sodium hydroxide	Sigma
Spectinomycin	Sigma
Sucrose	Sigma
Talon resin	Clontech
Trichloro acetic acid (TCA)	Sigma
N, N, N', N'-Tetramethylethylenediamine (TEMED)	Sigma
Tris-Base	Sigma
Triton X-100	sigma
Tween-20	Calbiochem

IV.1.2 Enzymes

Benzonase	Merck
Herculase DNA polymerase	Stratagene
Lysozyme	Sigma-Aldrich
Pfu DNA polymerase	Stratagene
Restriction endonucleases	New England Biolabs
Shrimp alkaline phosphatase	Roche
T4 DNA ligase	New England Biolabs

IV.1.3 Materials

Chromatography columns	Amersham Pharmacia
Concentration devices (Centricons)	Amicon
Millex SV filter units 0.22 μ M	Millipore

NAP desalting columns	Amersham Pharmacia
Nitrocellulose transfer membrane	Schleicher & Schuell
Steritop GP filter units 0.22 μ M	Millipore

IV.1.4 Instruments

AIDA gel imaging software version 2.31	Raytest
ÄKTA Explorer 100 chromatography system	Amersham Pharmacia
Balance AG285, PB602	Mettler Toledo
Centrifuges: Avanti J-25, Avanti J20 XP, J-6B, GS-6R	Beckmann
Centrifuges 5415C and 5417R	Eppendorf
Deionization system MilliQ plus PF	Millipore
Electrophoresis chambers MiniProtean 3	Bio-Rad
Electrophoresis power supply Power PAC 300	Bio-Rad
Fluorescence spectrometer Fluorolog 3	HORIBA Jobin Yvon
EmulsiFlex high pressure homogenizer	Avestin
Gene Pulser II electroporation system	Bio-Rad
Gilson Pipetman (2, 10, 20, 100, 200, 1000 μ l)	Abimed
Incubators Innova 4430	New Brunswick Scientific
Luminescent Image Analyzer LAS-3000	FUJIFILM
Luminometer (Lumat LB 9507)	BERTHOLD
Mini Trans-Blot Electrophoretic Transfer Cell	Bio-Rad
PCR-Thermocycler T3	Biometra
pH meter Accumet Basic	Fisher Scientific
SGD 2000 slab gel dryer	Savant
Sonicator Ultrasonic Processor XL	Misonix Inc.
Spectrophotometer DU 640 UV/VIS	Beckmann

UV/VIS Spectrometer V-560	Jasco
Thermomixer Comfort	Eppendorf
Vortex	Ikamag
Water bath	Bioblock Scientific

IV.1.5 Media and buffers

IV.1.5.1 Media

LB medium:	10 g/l bacto tryptone, 5 g/l bacto yeast extract, 10 g/l NaCl, pH adjusted to 7.0 with NaOH
LB agar:	16 g/l bacto agar dissolved in LB medium
SOC medium:	20g/l bacto tryptone, 5 g/l bacto yeast extract, 0.5g/l NaCl, 0.186 g/l KCl, 0.95 g/l MgCl ₂ . After autoclaving add 20 ml of filter sterilized 1 M glucose

IV.1.5.2 Antibiotic stock solutions

Antibiotic additives to growth media were prepared as 1000x (3000x for chloramphenicol) stock solutions and filter-sterilized before usage. The following concentrations were used for various antibiotics. Ampicilin: 100 mg/ml, Kanamycin: 25 mg/ml, Spectinomycin: 50 mg/ml, Chloramphenicol: 45 mg/ml (in 100% ethanol).

IV.1.5.3 Buffers

Buffer A	50 mM Tris-HCl pH 8.0, 20% (w/v) Sucrose
Buffer E1	PBS with 10 mM imidazole
Buffer E2	PBS with 25 mM imidazole
Buffer E3	PBS with 250 mM imidazole
Ca/glycerol buffer	10 mM PIPES, 60 mM CaCl ₂ , 15 % glycerol; pH 7.0, adjusted with NaOH, and filter-sterilized

Coomassie destaining solution	10% (v/v) Ethanol, 10% (v/v) Acetic acid
Coomassie staining solution	0.1% (w/v) Serva coomassie blue R250, 40% (v/v) Ethanol, 10% (v/v) Acetic acid
DNA loading buffer (10X)	2 g/l Orange G, 2 g/l Bromophenol Blue, 2 g/l Xylene cyanol FF, 0.37 g/l EDTA di-sodium salt dihydrate, 500 g/l sucrose
Hypo-osmotic lysis buffer	50 mM Tris-HCl pH 7.5, 20 mM MgSO ₄ , 0.2% Triton X-100, 100 U/ml Benzonase, 1 X EDTA-free complete protease inhibitor
Luciferase dilution buffer	25 mM Tris-phosphate pH 7.8, 2 mM DTT, 2 mM CDTA, 10% glycerol, 1% Triton X-100, 1 mg/ml BSA (w/v)
PBS (Phosphate buffered saline)	137 mM NaCl, 2.68 mM KCl, 10.1 mM Na ₂ HPO ₄ , 1.76 mM NaH ₂ PO ₄ , pH adjusted to 7.4 with HCl
Ponceau S stain	0.2% (w/v) Ponceau S stain, 3% (v/v) Trichloro acetic acid
SDS-PAGE electrophoresis buffer	50 mM Tris-HCl pH 8.3, 380 mM glycine, 0.1% (w/v) SDS
SDS-PAGE sample buffer (4X)	240 mM Tris (pH 6.8), 8% SDS (w/v), 40% glycerol, 1.4 M β-Mercaptoethanol, 0.02% bromphenol blue
TAE-buffer	242 g/l Tris base, 57.1 ml/l acetic acid, 50 mM EDTA
TBS (Tris buffered saline)	25 mM Tris-HCl, pH 7.2, 150 mM NaCl
TBST (TBS + Tween 20)	0.1% Tween 20 in TBS
Sucrose cushion	0.5 M Sucrose, 15 mM MgCl ₂ , 100 mM KoAc, 20 mM HEPES (pH 7.5)
Western blot transfer buffer	25 mM Tris, 192 mM glycine, 20% methanol (v/v), pH 8.4
Western blot stripping buffer	62.5 mM Tris-HCl pH 6.8, 10mM β-mercaptoethanol, 2% (w/v) SDS

IV.1.6 *E. coli* strains

BL21 (DE3)	B strain, F- <i>dcm</i> ⁺ Hte <i>ompT hsdS</i> (r _B ⁻ m _B ⁻) <i>gal l</i> <i>endA</i>
DH5α F'	F'/ <i>endA1 hsdR17</i> (r _k ⁻ , m _k ⁺) <i>glnV44 thi-1 recA1 gyrA</i> (Nal ^r) <i>relA1Δ(lacIZYA-argF) U169 deoR</i> (<i>φ80dlacΔ(lacZ)M15</i>)
MC4100Δ <i>tig</i>	F- <i>araD139 Δ(argF-lac) U169 rpsL150</i> (Str ^r) <i>relA1</i> <i>flbB5301 deoC1 pstF25 rbsR Δtig</i> Strain generated by Dr. Hung-Chun Chang

IV.1.7 Antibodies

Bag1	Lab collection
Hdj2	Sigma HPA001306
His-tag	Roche 11922416001
Hsc70	Stressgen SPA-820
Myc-tag	Santa Cruz SC 40
TF	Lab collection

IV.2 Methods

IV.2.1 General molecular biology methods

IV.2.1.1 Preparation and transformation of *E. coli* competent cells

For preparation of chemically-competent *E. coli* cells, a single colony was used to inoculate 500 ml LB medium (including antibiotic, if applicable) and grown to an optical density (OD₆₀₀) of 0.5 - 0.6 at 37 °C. The cells were then chilled on ice for 15 min and harvested at 5000 g for 10 min at 4 °C. The cell pellet was washed with 80 ml ice-cold Ca/glycerol buffer (10 mM PIPES, 60 mM CaCl₂, 15 % glycerol; pH 7.0, adjusted with NaOH, and filter-sterilized) once and incubated with additional 80 ml Ca/glycerol buffer

on ice for 30 min. Finally, the cells were pelleted and resuspended in 6 ml of Ca/glycerol buffer. 100 μ l aliquots were frozen in liquid nitrogen and stored at -80 °C.

For transformation, ~50 μ l competent cells were mixed with 0.05 - 0.2 μ g plasmid DNA or 5-10 μ l ligation reaction and incubated on ice for 15 min. The cells were heat-shocked at 42 °C for 45 s and subsequently placed on ice for 2 min. 1 ml of LB medium was added and the cells were shaken at 37 °C for 1 h. The cell suspension was then plated on selective plates and incubated at 37 °C, until colonies had developed (typically 10-16 h).

Alternatively, electroporation was applied to improve the transformation efficiency and avoid the heat shock process for certain bacterial strains (e.g. MC4100 Δ *tig* strain). Electrocompetent cells were prepared as follows: 500 ml bacterial culture was grown to an optical density (OD_{600}) of 0.5-0.6 in LB medium at the appropriate temperature (30 °C for MC4100 Δ *tig* strain). The cells were washed carefully with 250 ml ice-cold sterilized water twice and finally the cells were pelleted and resuspended in 2 ml of 10% glycerol. 50 μ l aliquots were frozen in liquid nitrogen and stored at -80 °C. For electroporation transformation, 50 μ l of competent cells were mixed with 1-2 μ l plasmid DNA (or ligation product) and transferred into a 0.2 cm Gene Pulser cuvette. The electroporation was done at 2.5 kV, 25 μ FD and 200 Ω settings with a Gene Puser II electroporation device. The transformed cells were allowed to recover in 1 ml of SOC medium with 225 rpm shaking at appropriate temperature for 1 h. The cell suspension was then plated on selective plates and incubated until colonies had developed (Dower et al., 1988).

IV.2.1.2 Plasmid purification

LB medium containing the appropriate antibiotic was inoculated with a single *E. coli* colony harboring the DNA plasmid of interest and shaken 8 - 14 h at 37 °C. Plasmids were isolated using Miniprep Kit (Promega) according to the manufacturer's instructions.

IV.2.1.3 PCR amplification

PCR (polymerase chain reaction)-mediated amplification of DNA was performed according to a standard protocol with minor modifications:

DNA Template:	10-20 ng (plasmid DNA) 250 ng or less (bacterial genomic DNA)
Primers:	20 pmol each
dNTPs:	200 μ M each
Pfu DNA Polymerase:	2.5 U
Polymerase buffer:	1 x
Additives:	4 % DMSO if GC content was >50 %,
Final volume:	50 μ l
Cycling conditions (25-35 cycles):	
Initial denaturation:	94 °C, 5 min
Cycle denaturation:	94 °C, 30-60 s
Annealing:	~55 °C, 30-60 s
Extension:	72 °C, duration dependent on template length: 1 kbp/min.
Final Extension:	72 °C, 10 min.

Completed PCR reactions were stored at 4 °C or -20 °C. When performing colony PCR, a single bacterial colony was picked with a sterile pipette and utilized as a source of DNA instead of purified DNA. PCR products were further purified using the PCR purification and gel extraction kit (Promega) according to the manufacturer's instructions.

IV.2.1.4 DNA restriction digestion and ligation

DNA restriction was performed according to the manufacturer's instructions for the respective enzymes. Typically, a 50 μ l digestion reaction contained 1-2 μ l of each restriction enzyme and 0.5-2 μ g purified PCR product or 1-5 μ g plasmid DNA in the appropriate reaction buffer and incubated at the recommended temperature for 2-3 h. Digested vector DNA was dephosphorylated with shrimp alkaline phosphatase at 37 °C for 1 h.

For ligation, 100-200 ng (~1-2 μ l) dephosphorylated plasmid DNA, 100-200 ng (~5-10 μ l) DNA insert and 1 μ l (100 U) T4 ligase were incubated in ligase buffer at 20 °C for 2 h or for increased efficiency at 16 °C overnight. The ligation product was transformed into competent *E. coli* DH5 α cells as described.

IV.2.1.5 DNA analytical methods

DNA concentrations were measured by UV absorption spectroscopy at $\lambda = 260$ nm. A solution of 50 μ g/ml of double stranded DNA in H₂O exhibits approximately A_{260} nm = 1. Agarose gel electrophoresis was performed in TAE buffer (40 mM Tris, 1 mM EDTA, 20 mM acetic acid) and 1 – 2 % TAE-agarose gels, supplemented with 1 μ g/ml ethidium bromide, at 4 – 6 V/cm. DNA sequencing was performed by Sequiserve (Vaterstetten, Germany).

IV.2.2 Cloning of various constructs for this study

pOFX-lac1TF and pOFX-lac1TFNC

The ORFs for the TF and TFNC were PCR amplified using 5'-primer encoding XbaI and NdeI (coding for initiator Methionine) sites and 3'-primer encoding BamHI site from pProEX-TF and pProEX-TFNC plasmids (Kaiser). The reactions were purified and digested with XbaI and BamHI. The digested inserts were purified and eluted in NF water. The empty plasmid pOFX-lac1 was digested with XbaI and BamHI and subsequently dephosphorylated with calf intestinal alkaline phosphatase (CIAP). The vector backbone was purified and eluted with NF water. For the ligation reaction, both

the purified plasmid and insert were incubated with 400 units of T4 DNA ligase at 16 °C for 16 h in a ratio of 1:3. In a control reaction NF water was added instead of the insert.

The whole ligation reaction was transformed into chemically competent DH5 α cells and plated on LB^{Spec} plates and incubated overnight at 37 °C. Single colonies were inoculated in LB^{Spec} for overnight cultures. Cultures were harvested and plasmids were prepared. The presence of the insert was confirmed by restriction digestion with XbaI and BamHI. Positive clones were confirmed additionally by sequencing the DNA.

pBAD18-Ras-DHFR-His₆

The ORF of Ras-DHFR-His₆ was excised from pET3a-Ras-DHFR- His₆ using XbaI and HindIII restriction enzymes and purified. pBAD18-Luc was digested with XbaI and HindIII to release Luc and the plasmid backbone was purified from Agarose gel. For the ligation reaction, both the purified plasmid backbone and insert were incubated with 400 units of T4 DNA ligase at 16 °C for 16 h in a ratio of 1:3. In a control reaction NF water was added instead of the insert.

The whole ligation reaction was transformed into chemically competent DH5 α cells and plated on LB^{Amp} plates and incubated overnight at 37 °C. Single colonies were inoculated in LB^{Amp} for overnight cultures. Cultures were harvested and plasmids were prepared. The presence of the insert was confirmed by restriction digestion with XbaI and HindIII. Positive clones were confirmed additionally by sequencing the DNA.

Plasmids for the co-expression system

The ORF of Luc was excised from pBAD18-Luc using XbaI and HindIII restriction enzymes and purified. pBAD33 was digested with XbaI and HindIII, followed by dephosphorylation with CIAP and purification. For the ligation reaction, both the purified plasmid backbone and insert were incubated with 400 units of T4 DNA ligase at 16 °C for 16 h in a ratio of 1:3. In a control reaction NF water was added instead of the insert. The whole ligation reaction was transformed into chemically competent DH5 α cells and plated on LB^{Cm} plates and incubated overnight at 37 °C. Single colonies were

inoculated in LB^{Cm} for overnight cultures. Cultures were harvested and plasmids were prepared. The presence of the insert was confirmed by restriction digestion with XbaI and HindIII. Positive clones were confirmed additionally by sequencing the DNA.

Human Hsc70 (hHsc70) was PCR amplified using 5'-primer encoding NcoI (coding for initiator Methionine) site and 3'-primer encoding NotI site from the pET11a-Hsc70 plasmid (Lab collection). The reaction was purified and digested with NcoI and NotI. The digested inserts were purified and eluted in NF water. The empty plasmid pCDFDUET (Novagen) was digested with NcoI and NotI and subsequently dephosphorylated and purified. For the ligation reaction, both the purified plasmid and insert were incubated with 400 units of T4 DNA ligase at 16 °C for 16 h in a ratio of 1:3. In a control reaction NF water was added instead of the insert. The whole ligation reaction was transformed into chemically competent DH5α cells and plated on LB^{Spec} plates and incubated overnight at 37 °C. Single colonies were inoculated in LB^{Spec} for overnight cultures. Cultures were harvested and plasmids were prepared. The presence of the insert was confirmed by restriction digestion with NcoI and NotI. Positive clones were confirmed additionally by sequencing the DNA.

Human Hdj2 (hHdj2) was excised from the pCH-Hdj2 plasmid (Lab collection) using NdeI and NheI restriction enzymes and purified. pETDUET (Novagen) was digested with NdeI and BlnI, followed by dephosphorylation with CIAP and purification. For the ligation reaction, both the purified plasmid backbone and insert were incubated with 400 units of T4 DNA ligase at 16 °C for 16 h in a ratio of 1:3. In a control reaction NF water was added instead of the insert. The whole ligation reaction was transformed into chemically competent DH5α cells and plated on LB^{Amp} plates and incubated overnight at 37 °C. Single colonies were inoculated in LB^{Amp} for overnight cultures. Cultures were harvested and plasmids were prepared. The presence of the insert was confirmed by restriction digestion with BamHI. Positive clones were confirmed additionally by sequencing the DNA.

TFNC was PCR amplified using using 5'-primer encoding NdeI (coding for initiator Methionine) site and 3'-primer encoding BlnI site from the pPROEX-HTa-

TFNC plasmid (Lab collection). The reaction was purified and digested with NcoI and NotI. The digested inserts were purified and eluted in NF water. The empty plasmid pCOLADUET (Novagen) was digested with NdeI and BlnI and subsequently dephosphorylated and purified. For the ligation reaction, both the purified plasmid and insert were incubated with 400 units of T4 DNA ligase at 16 °C for 16 h in a ratio of 1:3. In a control reaction NF water was added instead of the insert. The whole ligation reaction was transformed into chemically competent DH5 α cells and plated on LB^{Kan} plates and incubated overnight at 37 °C. Single colonies were inoculated in LB^{Kan} for overnight cultures. Cultures were harvested and plasmids were prepared. The presence of the insert was confirmed by restriction digestion with NdeI and BlnI.

This plasmid was further used for cloning of human Bag1 (hBag1) in the MCS1 by digesting with NcoI and Ecl136II. Bag1 was excised from the pROEX-HTa-Bag1 plasmid (Lab collection) using NcoI and MscI restriction enzymes and purified. For the ligation reaction, both the purified plasmid backbone and insert were incubated with 400 units of T4 DNA ligase at 16 °C for 16 h in a ratio of 1:3. In a control reaction NF water was added instead of the insert. The whole ligation reaction was transformed into chemically competent DH5 α cells and plated on LB^{Kan} plates and incubated overnight at 37 °C. Single colonies were inoculated in LB^{Kan} for overnight cultures. Cultures were harvested and plasmids were prepared. The presence of the insert was confirmed by restriction digestion with NcoI and HindIII. Positive clones were confirmed additionally by sequencing the DNA.

IV.2.3 Site directed mutagenesis

Site directed mutagenesis was employed to replace the threonine (150) codon ACC by the cysteine codon TGT in the ORF of TF and TFNC in the pProEX vector. Primers were designed such that they incorporate the mutations in the middle of their sequence and they were complementary to each other. 20 ng of the template DNA per 50 μ l of PCR reaction was mixed with 20 pmol of each of the primers, 200 μ M dNTPs, 1X enzyme buffer and 1 unit of the Pfu DNA polymerase. Elongation was performed at 72 °C for 6 min. The PCR products were digested with 1 μ l of *Dpn* I for an hour at 37 °C to

cleave the parental DNA. 10 μ l of the PCR reaction was used to transform chemically competent DH5 α cells and plated on LB^{Amp} plates. The plates were incubated overnight at 37 °C. Cells were harvested and plasmids were prepared. The presence of mutations was confirmed by DNA sequencing.

IV.2.4 Protein purification methods

IV.2.4.1 Purification of TF and TFNC and their cysteine mutants

BL21 (DE3) chemically competent cells were transformed with pPROEX-HTa plasmids encoding TF, TFNC and cysteine mutants and plated on LB^{Amp} plates. Secondary cultures in 2 l of LB^{Amp} were grown until an OD₆₀₀ of 0.6 was reached, induced with 1 mM IPTG and grown for 4 h post induction. Cells were harvested by centrifugation, resuspended in PBS and flash frozen in liquid nitrogen. The cell resuspension was thawed at 37 °C, incubated with a tablet of EDTA free Complete Protease Inhibitor (CPI), 10 U/ml benzonase and 0.5 mg/ml lysozyme for 30 min on ice. Cells were passed at least three fold through the Emulsiflex C5 Homogenizer to ensure complete lysis. The lysate was centrifuged and cleared by centrifugation at 50,000 g, for 60 min at 4 °C.

Cleared lysate was passed over a HiTrap chelating column charged with Ni²⁺ and pre-equilibrated with PBS. The column was washed with 10 mM and 25 mM imidazole in PBS and the bound protein was eluted with 250 mM imidazole in PBS. Fractions of purified protein were pooled, desalted in a HiPrep desalting column to remove imidazole.

Purified TF in PBS was subjected to TEV protease digestion in the presence of 2 mM DTT for removal of the N-terminal His₆ sequence. The reaction mix was passed onto the HiTrap chelating column charged with Ni²⁺ to remove the uncleaved protein and the His₆ tagged TEV protease, which would bind to the column by virtue of its His₆ tag. Purified protein without the His₆ tag was collected in the flowthrough, pooled, concentrated and aliquoted in small fractions and stored at -80 °C.

IV.2.4.2 Purification of eukaryotic chaperones

BL21 (DE3) chemically competent cells were transformed with pET11a-Hsc70 plasmids encoding His₆-tagged hHsc70 and plated on LB^{Amp} plates. Secondary cultures in 2 l LB^{Amp} were grown until an OD₆₀₀ of 0.6 was reached, induced with 1 mM IPTG and grown for 4 h post induction. Cells were harvested by centrifugation, resuspended in PBS and flash frozen in liquid nitrogen. The cell resuspension was thawed at 37 °C, incubated with a tablet of EDTA free Complete Protease Inhibitor (CPI), 10 U/ml benzonase and 0.5 mg/ml lysozyme for 30 min on ice. Cells were passed at least three fold through the Emulsiflex C5 Homogenizer to ensure complete lysis. The lysate was centrifuged and cleared by centrifugation at 50,000 g, for 60 min at 4 °C.

Cleared lysate was passed over a HiTrap chelating column charged with Ni²⁺ and pre-equilibrated with PBS. The column was washed with 10 mM and 25 mM imidazole in PBS and the bound protein was eluted with 250 mM imidazole in PBS. Fractions of purified protein were pooled, desalted in a HiPrep desalting column to remove imidazole. Purified protein with the His₆ tag was concentrated and aliquoted in small fractions and stored at -80 °C.

IV.2.5 Protein analytical methods

IV.2.5.1 Determination of protein concentration

Purified, unmodified proteins were quantified by absorbance spectroscopy at 280 nm in 6 M Guanidinium-HCl, 20 mM Na₂PO₄, pH 6.5 using calculated extinction coefficients from the ProtParam tool at the ExPASy proteomics server.

IV.2.5.2 SDS-PAGE (sodium-dodecylsulfate polyacrylamide gel electrophoresis)

SDS-Polyacrylamide gels were prepared as follows:

Chemicals	Stacking gel	Separating gel		
		10 %	12 %	15 %
	4 %	10 %	12 %	15 %
30 % Acrylamide (0.8% bis)	6.5 ml	16.7 ml	20 ml	25 ml
0.5 M Tris, pH 6.8	12.5 ml	–	–	–
1.5 M Tris, pH 8.8	–	12.5 ml	12.5 ml	12.5 ml
10 % SDS	0.5 ml	0.5 ml	0.5 ml	0.5 ml
2M Sucrose	–	12.5 ml	12.5 ml	12 ml
H ₂ O (up to 50 ml)	30.5 ml	7.8 ml	4.5 ml	–
TEMED	50 µl	25 µl	25 µl	25 µl
10% APS	500 µl	500 µl	500 µl	500 µl

SDS-PAGE was performed using a discontinuous buffer system (Laemmli, 1970) in BioRad Mini-Protean 3 electrophoresis chambers employing a constant current of 30 mA/gel in 50 mM Tris-Base, 380 mM glycine, 0.1 % SDS (pH 8.3). SDS loading buffer was added to the protein samples (final concentration: 60 mM Tris-HCl, pH6.8, 1% SDS, 10 % glycerol, 0,01% Bromophenol blue, 0,1 mM β -mercaptoethanol). Gels were fixed and stained in 0.1 % Coomassie brilliant blue R-250, 40 % ethanol, 7 % acetic acid for 1 h or longer and de-stained in 20 % ethanol, 7 % acetic acid.

IV.2.5.3 Autoradiography

Samples containing radiolabeled proteins were subjected to SDS-PAGE. After electrophoresis, gels were fixed, stained with Coomassie brilliant blue R-250 and destained. Afterwards the gels were briefly rinsed in water and dried on Whatman paper in a Slab Gel Dryer SGD 2000 for 50 min at 76 °C. Dried gels were exposed to a phospho-imaging plate (Fuji) overnight. The imaging plate was analyzed using an FLA

200 imaging system (Raytest). Band intensities were quantified using the AIDA software version 2.31 (Raytest).

IV.2.5.4 Western blotting

Proteins were separated by SDS-PAGE and transferred to nitrocellulose membranes in a Mini Trans-Blot Electrophoretic Transfer Cell (Bio-Rad) in 25 mM Tris, 192 mM glycine, 20 % methanol, pH 8.4 at constant current of 300 mA/gel for 1 h (Towbin *et al.*, 1979). Nitrocellulose membranes were blocked in 5 % skim milk powder in TBST (50 mM Tris-Cl, pH8.0, 150 mM NaCl, 0.05% Tween 20) for 1 h. The membranes were then incubated with a 1:1000 – 1:10000 dilution of primary antibody in TBST and extensively washed in TBST before incubation with a 1:5000 (for anti-mouse IgG) or 1:10000 (for anti-rabbit IgG) dilution of secondary antibody in TBST (anti-rabbit IgG and anti-mouse IgG, whole molecule – horseradish peroxidase (HRP) conjugate, Sigma-Aldrich). After extensive washing, protein bands were detected by incubating the membranes with ECL chemi-luminescence solution and developed with a Luminescent Image Analyzer (LAS-3000 system).

IV.2.6 *In vitro* assays

IV.2.6.1 Site specific labeling of single cysteine TF proteins

For site-specific labeling, NBD (*N*-((2-(iodoacetoxy)ethyl)- *N*-methyl)amino-7-nitrobenz-2-oxa- 1,3-diazole (IANBD ester)) and BADAN (6-bromoacetyl-2-dimethylaminonaphthalene) were used. Desalting steps in the course of site-specific labeling procedures were performed using Sephadex G-25 desalting columns (Nap-5 columns).

Typically, ~75 nmol of TF or TFNC single cysteine mutant protein was incubated with 150 nmol tris-(2-carboxyethyl) phosphine (TCEP) in 150 μ l PBS for 10 min at 25 °C to completely reduce cysteine thiol groups. A 4-fold molar excess of the fluorescent dye and 400 μ l of 20 mM Tris-HCl, pH 7.5 were added. The reaction was allowed to proceed for 90 min at 25 °C in the dark. β -mercapto-ethanol (β -ME) was added to a final concentration of 10 mM to quench the reaction. Excess dye and β -ME were removed by

desalting the protein into PBS. The eluate was concentrated to ~ 500 μ l in centricon tubes and desalted into PBS again. The eluate was again concentrated 2- to 3-fold and small aliquots of the labeled protein were flash-frozen in liquid nitrogen and stored at -80 °C in the dark.

The extent of labeling was calculated based on the molar absorptivity coefficient of $21000 \text{ M}^{-1} \text{ cm}^{-1}$ for BADAN and $23000 \text{ M}^{-1} \text{ cm}^{-1}$ for NBD. The concentration of the total protein was calculated by the Bradford method (Bradford, 1976). The ratio of the amount of the labeled protein to the total protein will yield the labeling efficiency. For example the extent of the NBD labeling is calculated by,

$$\text{Moles dye per mole protein} = \text{OD}_{397 \text{ or } 472} \times \text{dilution factor} / 21000 \text{ or } 23000 \times \text{protein concentration (M)}$$

where 21000 and 23000 are the approximate molar extinction coefficients of BADAN and NBD at 397 and 472 nm, respectively.

IV.2.6.2 *In vitro* translation in the PURE system

In vitro translations were performed in the PURE system (Shimizu et al., 2001; Shimizu et al., 2005) according to the manufacturer's protocol. Template DNA at the final concentration of 10 ng/ μ l was added to the reconstituted PURE system along with chaperones (unlabeled or NBD-labeled TF/TFNC, Hsc70, Hdj2, Bag1 and RAC at the mentioned concentrations), 0.8 μ Ci/ μ l of ^{35}S -methionine and ^{35}S - Cysteine (Express Protein Labeling Mix, PerkinElmer) and incubated at 30 °C for 50-60 min.

IV.2.6.3 Separation of ribosome-chaperone complexes

Typically 15-25 μ l of the PURE system reaction was layered over a 100 μ l sucrose cushion and centrifuged at 100,000 rpm for 20 min at 4 °C in Optima TLX ultracentrifuge. The supernatant was aspirated from pellet by gentle pipetting. The total protein in the supernatant was precipitated by the addition of an equal volume of 50% trichloroacetic acid (TCA) (v/v) followed by centrifugation at 14,000 rpm for 15 min.

The supernatant was gently aspirated with vacuum and the pellet was washed with 1 ml of 100% acetone. The supernatant was again gently aspirated, the pellet was air dried and resuspended in SDS loading buffer.

IV.2.6.4 Post-translational folding of Luc

Luc was translated from the pET3a-Luc plasmid in PURE system in the absence or presence of 5 μ M TF/TFNC at 30 °C. After 30 min of translation, 50 μ g/ml of RNase A was added. Luc activity was measured immediately before the addition of RNaseA and at regular intervals after that. Activities prior to the addition of RNase A were set to 1.

IV.2.6.5 Limited proteolysis of Ras-DHFR

Ras-DHFR was translated in the absence (control) or in presence of 5 μ M TF or TFNC in the PURE system at 30 °C. After translation initiation, aliquots were taken at different time points. These aliquots were treated with 12.5 μ g/ml Proteinase K for 8 min on ice followed by its inactivation by addition of EDTA-free protease inhibitors (Roche). Samples were precipitated with 50% TCA and then washed with acetone. They were then separated by SDS-PAGE and visualized by autoradiography.

IV.2.7 *In vivo* co-expression experiments

MC4100 *Atig E. coli* was generated by inserting Cm cassette in the *tig* gene and then eliminating Cm cassette as previously described (Datsenko and Wanner, 2000). It was cotransformed with pBAD18-Luc or pBAD18-Ras-DHFR-His₆ and pOFXlac-1 or pOFXlac-1TF or pOFXlac-1TFNC was grown in LB media to an OD_{600nm} of 0.5 at 30 °C. TF or TFNC were overexpressed by induction with 1 mM IPTG for 30 min before induction of Luc or Ras-DHFR with 0.2% arabinose for 30 min. Cells were pellet and resuspended in 100 μ l of Buffer A. Spheroplasts were produced by treating the resuspended cells with 20 μ l of 5mg/ml lysozyme, 40 μ l of 10mM EDTA (pH 8) and 40 μ l 50 mM Tris-HCl (pH 8) and lysed with 200 μ l of Hypo-osmotic lysis buffer containing 0.2% Triton X-100, 100 U/ml Benzonase (Merck) and EDTA-free protease inhibitors (Roche).

IV.2.8 Luciferase specific activity and solubility measurements

pET3a-Luc vector encoding the Firefly luciferase gene was translated in the PURE system in the absence or presence of either 5 μ M TF or TFNC and 0.8 μ Ci/ μ l of 35 S-methionine and 35 S-Cysteine (Express Protein Labeling Mix, PerkinElmer) for 60 min at 30 °C. To measure enzymatic activity 2 μ l aliquots of either the translation reactions or *E. coli* lysate were diluted into 200 μ l of Luciferase dilution buffer. 2 μ l of the above mixture was added to 48 μ l of Luciferase assay buffer (Promega) and firefly luminescence was measured in the luminometer. Simultaneously a fraction of the translation reactions was loaded on SDS-PAGE and the amount of Luciferase protein synthesized was calculated by autoradiography or Western blotting against myc-tag. Specific activity of Luciferase was calculated by dividing the enzymatic activity with the amount of protein synthesized. To determine solubility, the translation reaction or *E. coli* lysate was pelleted by centrifugation at 20,800 g for 30 min and analyzed by SDS-PAGE followed by Western blotting against myc-tag.

IV.2.9 Fluorescence measurements

Kinetic fluorescence measurements were performed on a Fluorolog 3 fluorometer (Jobin Yvon) at 30 °C with excitation wavelength of 472 nm. Emission spectra were collected at 500 nm to 650 nm. The bandwidth of the excitation and emission light was adjusted to yield a maximum of $\sim 10^6$ counts per second (cps). To correct for the fluctuations in the excitation light source, the reference signal (R) was recorded along with the actual signal (S) in the course of all the measurements. S/R was used for data evaluation. BADAN and NBD fluorescence changes were monitored in the presence of 250 nM of labeled TF proteins in the PURE system. The reactions were transferred to a cuvette prewarmed to 30 °C. Once a steady signal was reached, \sim after 3 min from the start of measurement, translation was initiated upon addition of DNA at a final concentration of 10 ng/ μ l. The change in fluorescence was monitored as translation proceeded.

Once the fluorescence of the labeled proteins had reached steady state, they were displaced by addition of excess unlabeled TF (competitor) at a final concentration of 20

μM . Depending on the best fit, data from the competition experiments were analyzed using a three parameter single exponential function or a five parameter double exponential function using Equation 1 and Equation 2, respectively.

$$y = y_0 + ae^{-bx} \quad (\text{Equation 1})$$

$$y = y_0 + ae^{-bx} + ce^{-dx} \quad (\text{Equation 2})$$

where y_0 is the initial value, a represents the amplitude and b and d represent the time constants. The time constant b is used to calculate the half-time in the single exponential reactions and time constants b and d are used to calculate the half-time in the double exponential reactions by the following equations.

$$t_{1/2} = \ln(2)/b \quad (\text{Equation 3})$$

$$t_{1/2} = \ln(2)/d \quad (\text{Equation 4})$$

V RESULTS

V.1 Analysis of TF-nascent chain interactions using real-time fluorescence spectroscopy

TF is the first chaperone to interact with nascent chains during their translation. We have utilized fluorescence spectroscopy in order to directly monitor this interaction in real-time. Based on previous cross-linking results (Lakshmipathy et al., 2007), the residue E326 in the arm 1 of the TF C-Terminal domain was mutated to cysteine (Figure 14 A) and labeled with a fluorophore named NBD (*N*-((2-(iodoacetoxy)ethyl)-*N*-methyl)amino-7-nitrobenz-2-oxa-1,3-diazole (IANBD ester)). The fluorescence of NBD increases on transfer into a hydrophobic environment (Lin and Struve, 1991).

Luc was used as a model substrate since TF was shown to improve its specific activity (Agashe et al., 2004) and it was shown to interact with the Luc nascent chains using photocrosslinking and fluorescence spectroscopy experiments (Kaiser et al., 2006; Lakshmipathy et al., 2007; Tomic et al., 2006). α -Syn was used as negative control for TF-nascent chain interaction since it lacks hydrophobic motifs predicted to bind TF. The primary sequences of Luc and α -Syn were analyzed by the Roseman hydrophobicity scale to predict potential binding motifs for TF. Based on the analysis of several model proteins, if a region of minimum five consecutive residues in the central region of a 15 amino acid window have a mean hydrophobicity value of < -0.5 kcal/mol, then this region was considered a potential motif for TF binding (Kaiser et al., 2006). Based on this criterion, Luc contains five hydrophobic motifs predicted to bind to TF with scores of +13 (residues 87-100), +7 (residues 238-245), +6 (residues 248-254), +5 (residues 255-260), and +6 (residues 285-291), whereas α -Syn lacks such hydrophobic motifs (Figure 14 B). To monitor TF ribosome interaction, we used the TF R14C mutant labeled with the fluorophore BADAN (6-bromoacetyl-2-dimethylaminonaphthalene) as previously described (Kaiser et al., 2006).

V.1.1 TF is recruited to nascent chains exposing hydrophobic motifs

To study the interaction of labeled TF326-NBD with Luc and α -synuclein (α -Syn) nascent chains in a chaperone-free environment, we used the PURE system of translation. The PURE system is a reconstituted translation reaction consisting of purified components from *E. coli*, specifically 32 translation factors, purified ribosomes, 46 tRNAs, the 20 amino acids and factors required for the energy regeneration system (Shimizu et al., 2001; Shimizu et al., 2005). It is superior to conventional *E. coli* translation systems since it lacks chaperones and contaminating proteases and nucleases.

During translation of Luc in the presence of TF326-NBD, a significant increase NBD fluorescence was observed corresponding to a relative change in NBD fluorescence ($F-F_0/F_0$) of ~ 0.7 (Figure 14 C). This indicates the high affinity of TF towards the hydrophobic segments present in Luc. The fluorescence saturated when the translation reaction reached steady-state. Similar experiments performed with the ribosome-binding deficient mutant (FRK/AAA) of TF326-NBD did not result in a significant increase in NBD fluorescence. This confirms that ribosome binding is a prerequisite for TF-nascent chain interaction as was shown before (Kaiser et al., 2006). The interaction of TF326-NBD with α -Syn nascent chains was weak, consistent with the absence of hydrophobic regions in α -Syn nascent chains. Lack of TF326-NBD interaction with α -Syn is not due to reduced translation of α -Syn since it is translated to higher levels than Luc (Figure 14 D).

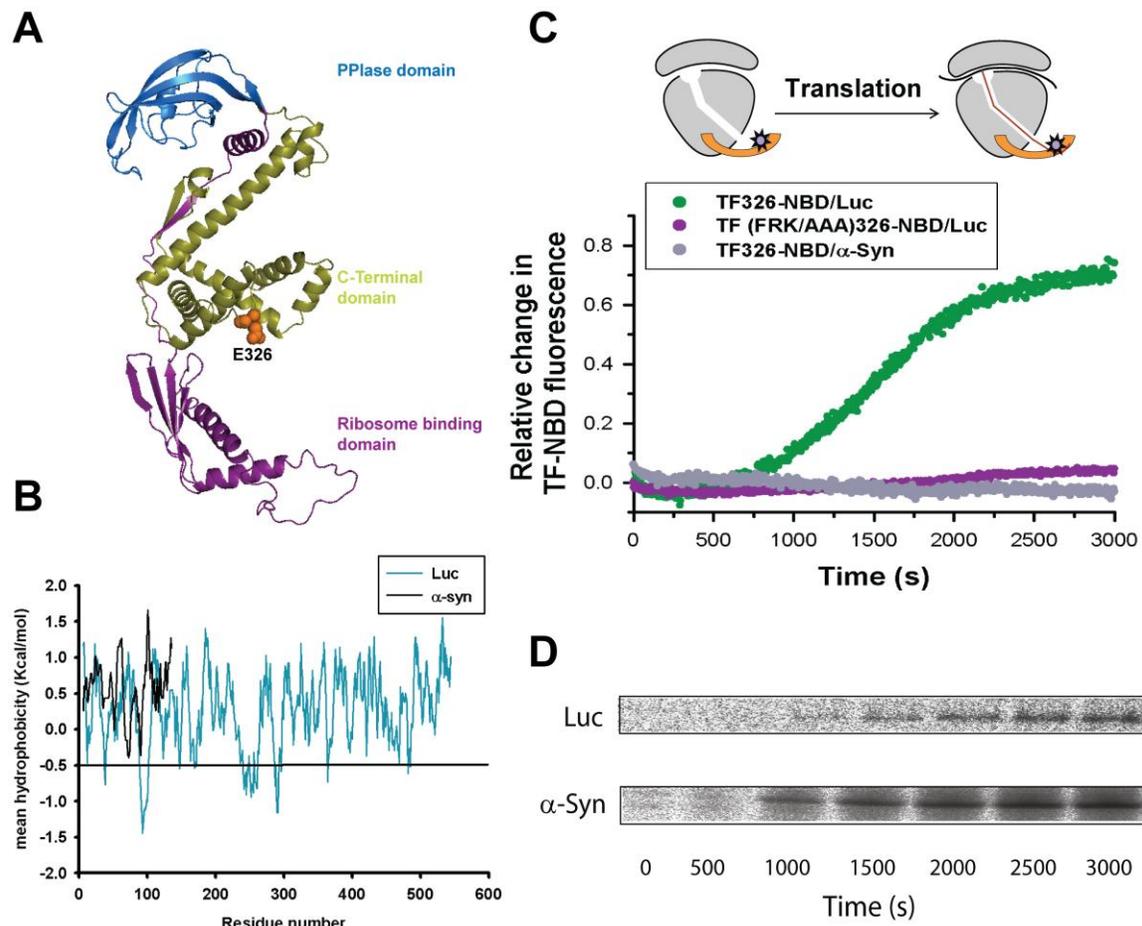


Figure 14: Real-time fluorescence measurements of TF interaction with Luc and α -Syn nascent chains.

(A) Crystal structure of *E. coli* TF (PDB code 1W26). Residue E326 which was changed into cysteine for NBD labeling is indicated in orange spheres. (B) Hydrophobicity analysis of Luc and α -Syn was performed as previously published (Kaiser). The predicted value of threshold hydrophobicity for TF binding is shown as a black line. (C) Luc was translated in presence of either 1 μ M TF326-NBD (green) or TF(FRK/AAA)326-NBD (purple) and NBD fluorescence was monitored. α -Syn was translated in presence of TF326-NBD (gray) and NBD fluorescence was monitored. (D) Autoradiograph of Luc and α -Syn translation in PURE system with aliquots taken at indicated time points.

We also monitored the recruitment of TF to ribosomes translating Luc and α -syn nascent chains. To this end, the TF mutant R14C was labeled with a conformationally-sensitive fluorophore BADAN (TF-B) (Figure 15 A) and the decrease in its fluorescence upon ribosome binding was studied. It has previously been shown that the fluorescence of TF-B decreases upon recruitment of TF to Luc RNCs. The fluorescence of TF-B during

translation of Luc saturated at around 1200 s (Figure 15 B) whereas the fluorescence of TF326-NBD saturated around 2500 s (Figure 14 C) confirming that TFs interaction with nascent chains occurs later in the translation process than its interaction with ribosomes.

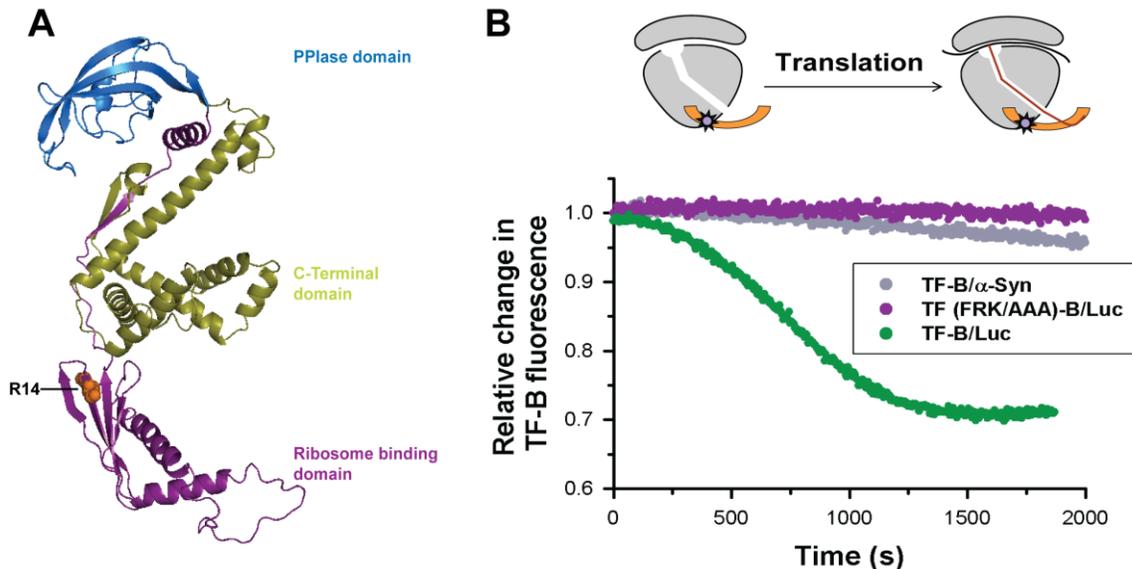


Figure 15: Real-time fluorescence measurements of TF interaction with ribosomes translating Luc and α -Syn nascent chains.

(A) Crystal structure of *E. coli* TF (PDB code 1W26). Residue R14 which was changed into cysteine for BADAN labeling is indicated as an orange sphere. (B) Luc was translated in presence of either 1 μ M TF-B (green) or TF(FRK/AAA)-B (purple) and BADAN fluorescence was monitored. α -Syn was translated in presence of TF-B (gray) and its fluorescence was monitored.

V.1.2 Additional TF molecules are recruited towards elongating Luc nascent chains exposing hydrophobic segments

Next we wanted to analyze whether additional TF molecules are recruited to longer Luc nascent chains. As mentioned earlier, Luc contains five hydrophobic motifs predicted to bind to TF (Figure 16 A). Different lengths of ribosome-arrested Luc nascent chains were generated by PCR and translated in the PURE system. These nascent chains exposed variable numbers of hydrophobic motifs outside the ribosomal exit tunnel, considering that 30-40 amino acids are accommodated inside the tunnel (Malkin and Rich, 1967). Luc nascent chains were translated in presence of TF-B and TF326-NBD

and the relative change in fluorescence was normalized for the amount of radiolabeled protein synthesized.

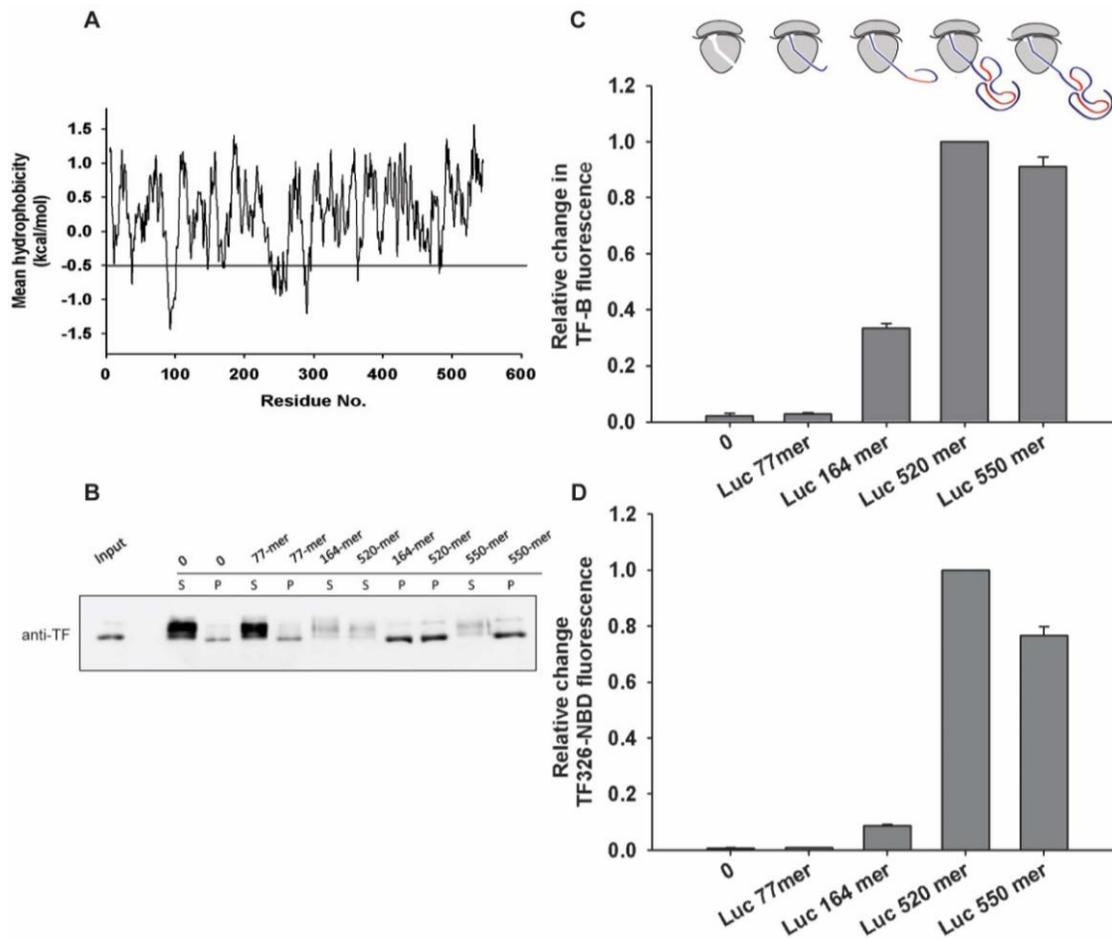


Figure 16: Increased recruitment of TF326-NBD to Luc nascent chains exposing hydrophobic motifs

(A) Hydrophobicity analysis of Luc was performed as previously published (Kaiser). The predicted value of threshold hydrophobicity for TF binding is shown as a black line. (B) Sedimentation analysis of TF-RNC interaction via sucrose cushion. The samples were separated by SDS-PAGE and immuno-blotted using anti-TF antibody. Different chain lengths of Luc were translated in presence of (C) 1 μ M TF-B and the decrease in BADAN fluorescence was monitored or in presence of (D) 1 μ M TF326-NBD and the increase in NBD fluorescence was monitored. The relative changes in BADAN and NBD fluorescence were normalized for the amount of radiolabeled protein synthesized.

During the translation of Luc 77-mer which exposed none of the hydrophobic motifs, no increase in TF-B and TF326-NBD fluorescence was observed (Figures 16 C

and D), suggesting that TF does not bind to Luc 77-mer. Interestingly, translation of Luc 164-mer caused a significant recruitment of TF-B to the ribosome, whereas only a modest increase was measured in the fluorescence of TF326-NBD (Figures 16 C and D). Luc 164-mer exposed one hydrophobic motif between 87 to 100 amino acids outside the ribosomal-exit tunnel. A 10-fold increase in NBD fluorescence was obtained when Luc 520-mer (exposing all the hydrophobic segments) was translated whereas only a 3-fold change was observed in TF-B fluorescence during the translation of this nascent chain (Figures 16 C and D). This 10-fold increase in NBD fluorescence during translation of longer nascent chains would also indicate that such nascent chains recruit multiple TF molecules. Sedimentation analysis of TF interaction with these Luc RNCs (Ribosome nascent chain complexes) also revealed an increased association of TF with longer Luc nascent chains (Figure 16 B). The NBD fluorescence was lower during the translation of Luc 550-mer (corresponding to full length Luc) than Luc 520-mer (Figure 16 D). This suggests that a minor proportion of TF-nascent chains interaction was lost due to partial hydrophobic collapse.

V.1.3 Dissociation of TF from Luc RNCs depends on the location of the fluorophore

In order to get more insights into the residence time of TF on nascent chains, we studied the dissociation of TF labeled at different locations from FL Luc RNCs. Residues E326 in arm 1, S376 in arm 2 and T150 at the interface between the PPIase domain and C-terminal domain were changed into cysteines and labeled with NBD (Figure 18 A). Positions 326 and 376 have been shown to be in close proximity to nascent chains by photocrosslinking experiments (Lakshmipathy et al., 2007).

FL (full length) Luc was translated in PURE system in presence of these labeled proteins and the increase in NBD fluorescence was monitored. The net fluorescence increase of TF376-NBD upon binding to FL Luc was similar to that of TF326-NBD (Figure 17 A). However, the fluorescence increase of TF150-NBD was ~ 3-fold lower than that with TF326-NBD and TF376-NBD (Figure 17 A). We suggest that since the PPIase domain acts a secondary binding site for nascent chains (Kaiser et al., 2006;

Lakshmipathy et al., 2007; Merz et al., 2008; Tomic et al., 2006), only a fraction of FL Luc nascent chains are in close proximity to the residue 150. The three NBD-labeled TF proteins, TF326-NBD, TF376-NBD and TF150-NBD were as efficient as wild type TF in supporting the *de novo* folding of Luc in the PURE system (Figure 17 B).

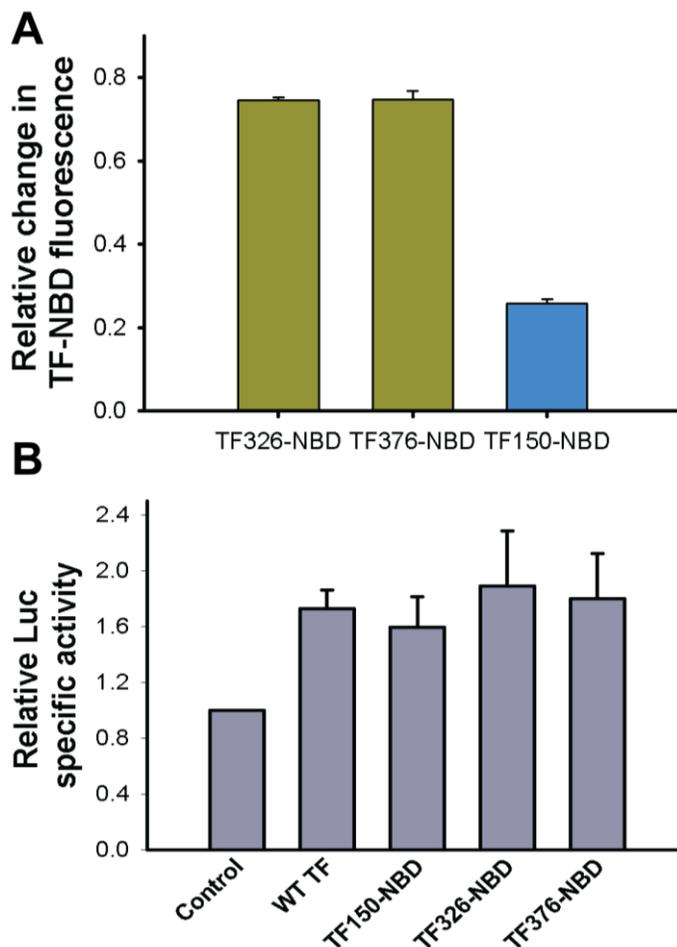


Figure 17: Properties of TF326-NBD, TF376-NBD and TF150-NBD

(A) The relative change in fluorescence of TF326-NBD, TF376-NBD or TF150-NBD upon interaction with FL Luc was plotted. (B) FL Luc was translated in presence of WT TF, TF326-NBD or TF376-NBD or TF150-NBD in PURE system and Luc specific activities were calculated by dividing the activity by amount of Luc translated.

Next, we studied the dissociation kinetics of TF150-NBD, TF326-NBD and TF376-NBD from FL Luc nascent chains. The reactions were prepared as mentioned previously (Figure 17 A) and upon saturation of NBD fluorescence, excess unlabeled TF

(20 μM) was added and the dissociation of the NBD-labeled TF from FL Luc RNC was measured. Interestingly, the dissociation of TF326-NBD from FL Luc followed biphasic kinetics with fast and slow phases (Figure 18 B) with $t_{1/2}$ values of 14 ± 2 s and 102 ± 16 s, respectively (Figure 18 C). The fast phase was identical to the dissociation of TF-B ($t_{1/2}$ value 14 ± 0.5 s) from FL Luc RNCs (Figure 18 C). When ribosome-binding deficient TF(FRK/AAA) was used as a competitor instead of TF, no displacement of TF326-NBD was observed, indicating that ribosome binding is essential for the displacement to occur (Figure 18 B).

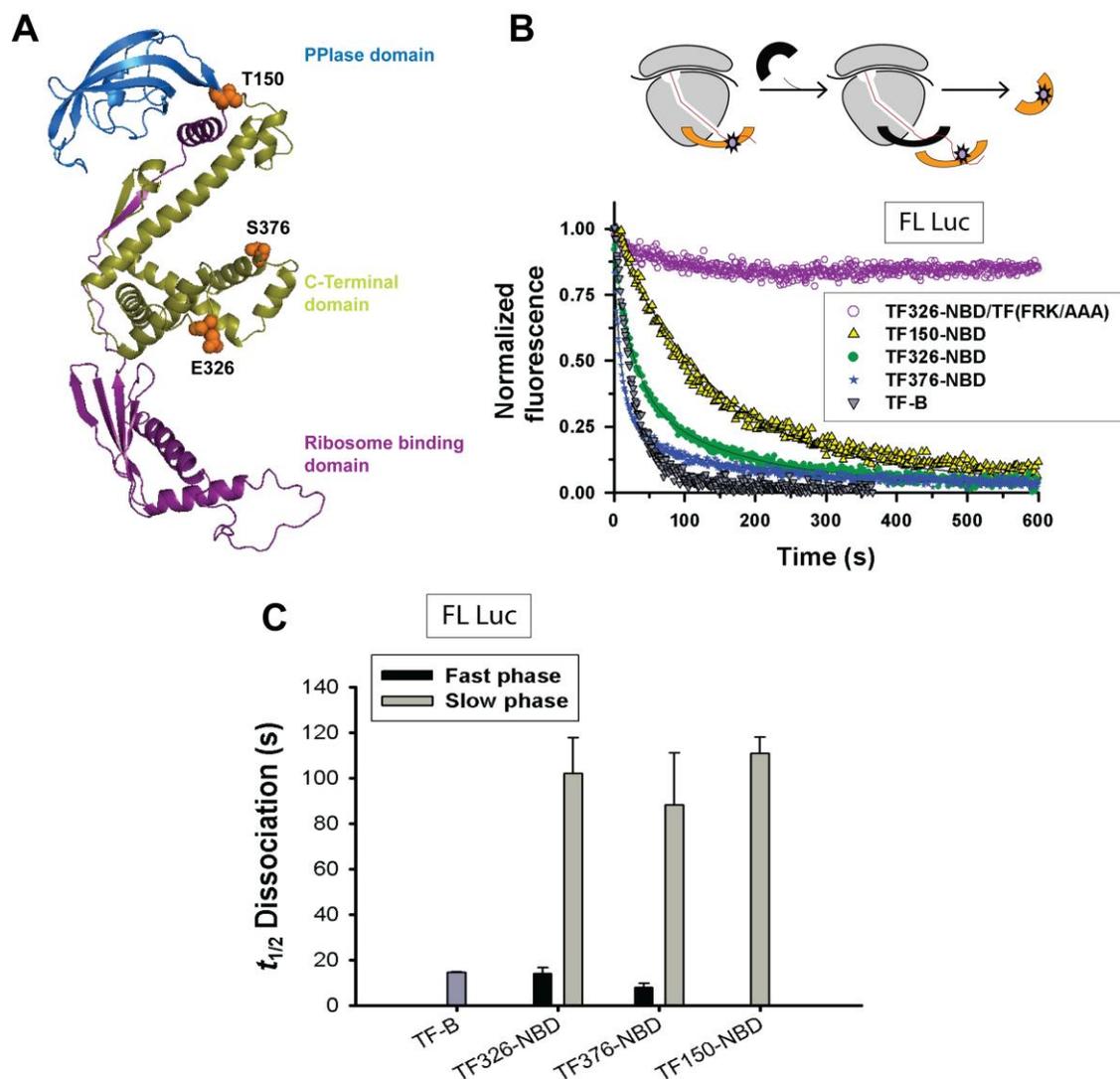


Figure 18: Dissociation kinetics of TF-B, TF326-NBD, TF376-NBD and TF150-NBD from FL Luc

(A) Crystal structure of *E. coli* TF (PDB code 1W26). Residues E326, S376 and T150 which were changed into cysteine for NBD labeling are indicated as orange spheres. (B) FL Luc was translated in presence of either 1 μ M of TF-B or TF326-NBD or TF376-NBD or TF150-NBD and upon saturation of fluorescence the complexes were dissociated by addition of either WT TF or TF(FRK/AAA). The decrease in NBD fluorescence was monitored. (C) The dissociation curves of (B) were fitted to either single exponential or double exponential equations depending on the best fit and the $t_{1/2}$ values were plotted.

We then analyzed whether the biphasic dissociation of TF from Luc was dependent on the location of the NBD probe. To this end, we studied the TF-nascent chain interaction using TF376-NBD (probe in the arm 2) and TF150-NBD (probe at the interface between PPIase domain and C-terminal domain) (Figure 17 A). We found that the dissociation of TF376-NBD from Luc nascent chains also followed biphasic kinetics with a fast rate of $t_{1/2}$ 8 ± 2 s and a slow rate of $t_{1/2}$ 88 ± 23 s, but the dissociation of TF150-NBD had only a slower phase of $t_{1/2}$ 111 ± 7 s (Figure 18 C). Overall, the analysis of dissociation of TF labeled with NBD at different positions demonstrated that TF dissociates from the translating ribosomes with a rate of ~ 14 s but maintains its association with the nascent chains for a longer time (~ 110 s). TF ribosome interaction was measured by using TF R14C mutant labeled with the fluorophore BADAN (Kaiser et al., 2006).

V.1.4 TF dissociates from different hydrophobic regions of Luc with different rates

Knowing that additional TF molecules bind to elongating Luc nascent chains, we investigated the dissociation kinetics of NBD-labeled TF from different lengths of Luc nascent chains. Luc 550-mer was translated in presence of TF326-NBD, TF376-NBD or TF150-NBD in the PURE system and following saturation of fluorescence the complexes were dissociated by addition of excess unlabeled TF. We found that the dissociation of TF326-NBD and TF376-NBD from Luc 550-mer followed biphasic kinetics similar to their dissociation from FL Luc (Figure 19A). Also, the dissociation of TF150-NBD from Luc 550-mer had a single phase similar to its dissociation from FL Luc (Figure 19A). Since the measurements were taken with arrested nascent chains, this suggests that the translation rate *per se* does not affect TF dissociation from RNCs. Interestingly, the

dissociation of TF326-NBD and TF376-NBD from Luc 164-mer had a single phase similar to the fast phase observed during dissociation of these proteins from Luc 550-mer and the dissociation of TF-B from Luc 164-mer RNCs (Figure 19 B). No increase was observed in the fluorescence of TF150-NBD during translation of Luc 164-mer, indicating that the PPIase domain does not interact with the hydrophobic region on the shorter Luc nascent chains (Figure 19 B).

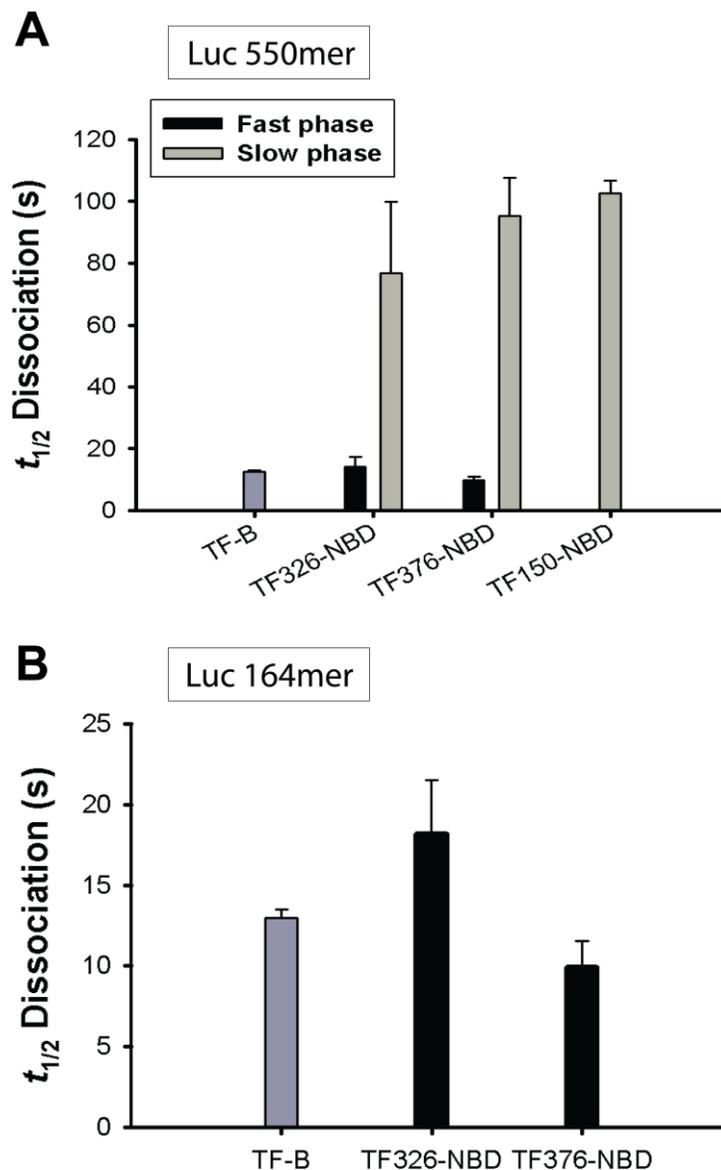


Figure 19: Dissociation kinetics of TF-B, TF326-NBD, TF376-NBD and TF150-NBD from Luc 550-mer and Luc 164-mer

(A) Luc 550-mer and (B) Luc 164-mer were translated in presence of either 1 μ M of TF-B, TF326-NBD, TF376-NBD or TF150-NBD and upon saturation of fluorescence the complexes were dissociated by addition of excess WT TF. The dissociation was analyzed using either a single or double exponential function depending on the best fit. The $t_{1/2}$ values were plotted.

The dissociation kinetics of TF from Luc 164-mer is similar to its dissociation from the ribosomes (Figure 19 B), indicating that TF interacts with the shorter Luc nascent chains while still being in contact with the ribosome. However, during elongation of Luc nascent chains, TF vacates its ribosome binding site for another TF molecule and maintains its association with the nascent chain for a longer time represented by the slower dissociation phase.

V.2 The PPIase domain of TF delays the folding of eukaryotic multi-domain proteins relative to their translation

The real-time fluorescence analysis of the TF interaction with nascent chains demonstrated that the PPIase domain of TF provides a secondary binding site for nascent chains, whereas the N- and C-terminal domains form the primary nascent chain binding site. To examine the effect of the PPIase domain in nascent chain binding and the subsequent effect of this binding on the folding of multi-domain proteins, we generated the PPIase domain deleted mutant of TF called TFNC and purified the protein similarly to TF. In this mutant the PPIase domain, comprising residues 151 to 247, was deleted and replaced by a flexible linker containing the amino acid sequence GTSAAA (Figure 20). Notably, TF and TFNC have similar ribosome-binding kinetics (Kaiser et al., 2006).

We used Luc as a model protein because it folds inefficiently in *E. coli* cells and in the *E. coli in vitro* translation systems (Agashe et al., 2004). In addition, it was shown that Luc depends on TF to a certain extent for its *de novo* folding but not for its refolding (Agashe et al., 2004). It is active as a monomer and does not depend on assembly after *de novo* synthesis. Additionally, Luc has a rapid and sensitive luminescence based enzymatic assay. This means that light emitted from the monomer is a direct measure of the functional and fully folded enzyme.

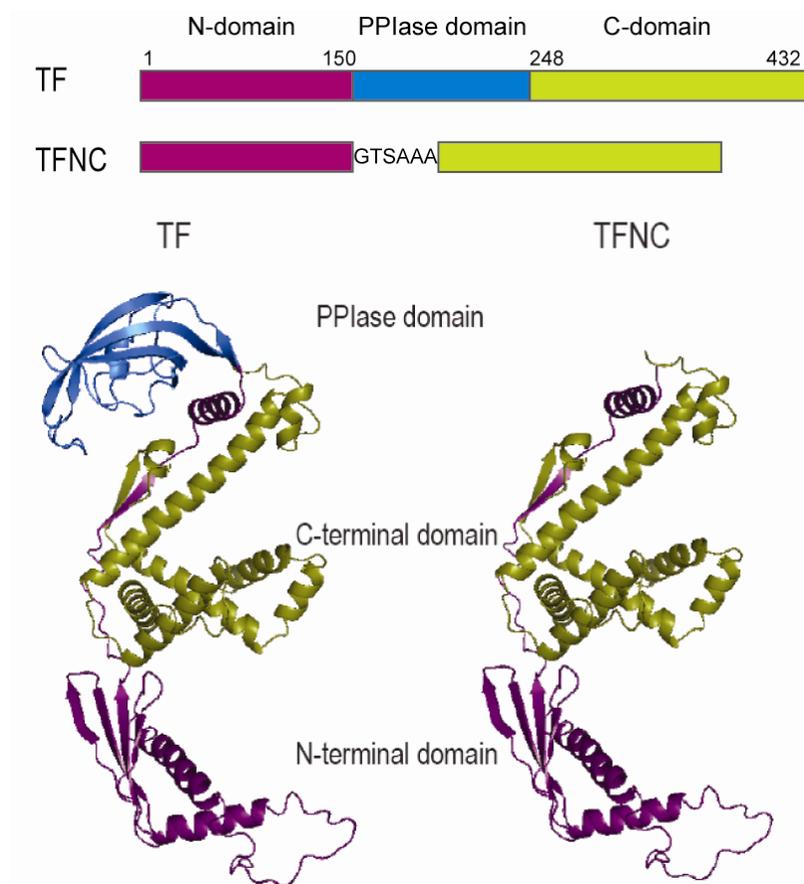


Figure 20: Representation of the linear and folded structure of TF and TFNC

TF contains three domains namely; the N-domain, PPIase domain and C-domain. TFNC contains only the N-domain and the C-domain linked by a flexible 6 amino acid long linker. Based on the structure coordinates of PDB code 1W26 (Ferbitz et al., 2004).

V.2.1 Dissociation of TFNC150-NBD from Luc-RNCs is faster than that of TF150-NBD

As demonstrated by the results above, the dissociation of TF327-NBD and TF376-NBD from FL Luc had biphasic kinetics (Figure 18 C). Thus, to obtain insights into the role of the PPIase domain in nascent chain binding, we chose the position T150 which had single phase dissociation kinetics from FL Luc. T150 of TFNC was changed to cysteine and labeled with the fluorophore NBD (*N*-((2-(iodoacetoxy)ethyl)-*N*-methyl)amino-7-nitrobenz-2-oxa-1,3-diazole) to yield TFNC150-NBD.

Translation of Luc in presence of 1 μM of the labeled proteins TF150-NBD or TFNC150-NBD led to an increase in NBD fluorescence indicating TF recruitment to the hydrophobic regions present in Luc (Figure 21 A). After the increase in fluorescence had saturated, dissociation of TF150-NBD or TFNC150-NBD from Luc nascent chain complexes was achieved by the addition of excess unlabeled TF. This led to dissociation of labeled proteins from Luc nascent chains causing a decrease in NBD fluorescence. Interestingly, we found that TFNC150-NBD dissociated faster from the Luc nascent chains than TF150-NBD, indicating that the TF residence time with Luc nascent chains is shortened by the deletion of the PPIase domain (Figure 21 B).

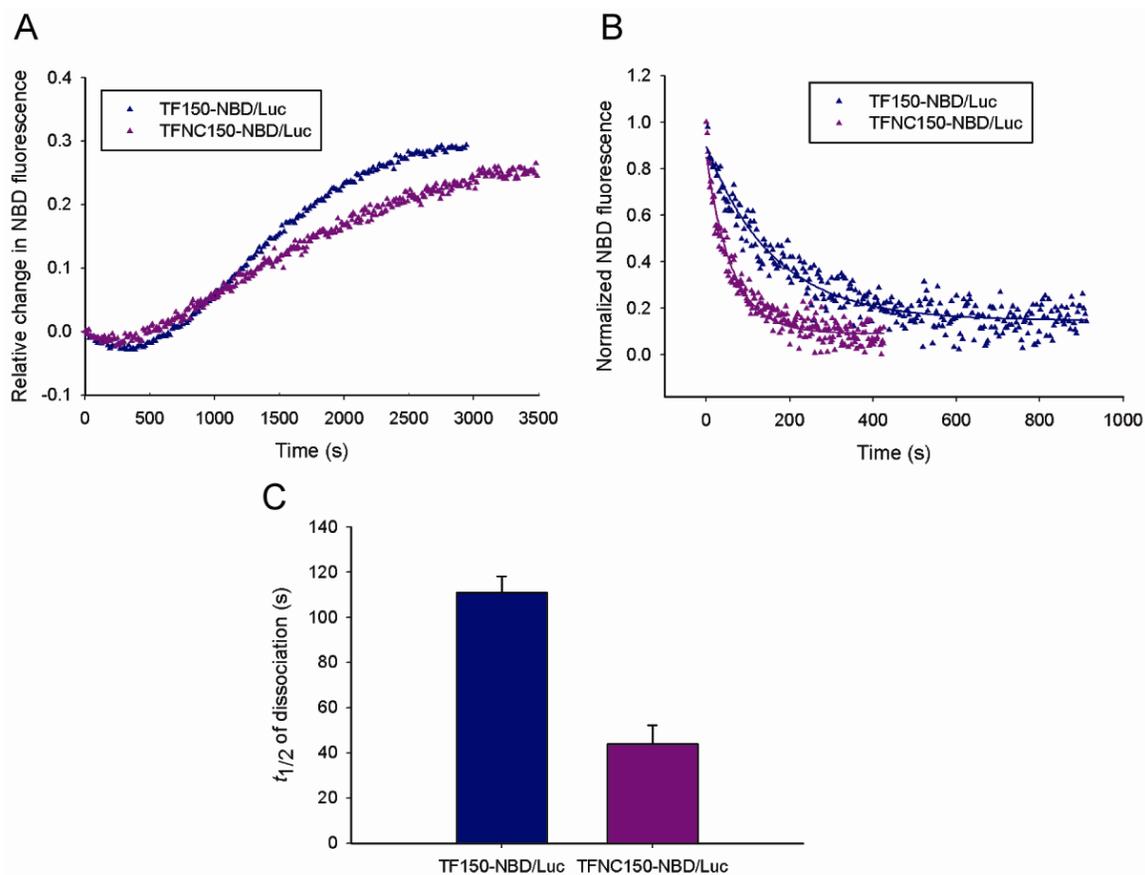


Figure 21: Real time fluorescence measurements of TFNC interaction with Luc nascent chains.

(A) Luc was translated in the PURE system in the presence of TF150-NBD or TFNC150-NBD. The relative change in NBD fluorescence during ongoing translation was monitored. After saturation of fluorescence had occurred, excess unlabeled TF was added to the reaction to dissociate the labeled TF or TFNC from ribosome-associated

Luc nascent chains. The normalized fluorescence change was plotted against time. (C) Comparison of $t_{1/2}$ values for TF150-NBD/Luc and TFNC150-NBD/Luc.

The dissociation curves were fitted to a single exponential three parameter function to calculate the mean residence time of these labeled proteins on Luc nascent chains. TFNC150-NBD dissociated with a $t_{1/2}$ of 57 ± 9 s whereas TF150-NBD dissociated with a $t_{1/2}$ of 111 ± 7 s (Figure 21 C). These competition experiments clearly show that TFNC resides for a shorter time on Luc nascent chains as compared to TF, hence relieving the delay in Luc folding relative to translation. This mode of action may provide an opportunity to nascent chains to undergo more efficient co-translational folding

V.2.2 TFNC improves the folding of Luc in PURE system

The folding pathway mediated by TFNC maybe relatively more co-translational as compared to the post-translational folding pathway mediated by TF (Agashe et al., 2004). This may improve the efficiency of folding as is observed upon translation in the eukaryotic cytosol (Frydman et al., 1994). We used the PURE system of translation to test the effect of TF/TFNC on the folding of Luc. Translation of Luc from pET3a-Luc plasmid in the presence of TF led to an improved folding yield as evidenced by a 1.7-fold increase in Luc specific activity. Interestingly, when Luc was translated in presence of TFNC, we observed an increase in the specific activity of Luc of 2.6-fold compared to the unsupplemented PURE system reaction (Figure 22). Luc was translated at comparable levels in both control and TF-supplemented reactions, whereas the level of Luc translation was 2-fold higher in the reaction supplemented with TFNC. The reason for this increased translation is not clear and remains to be further investigated.

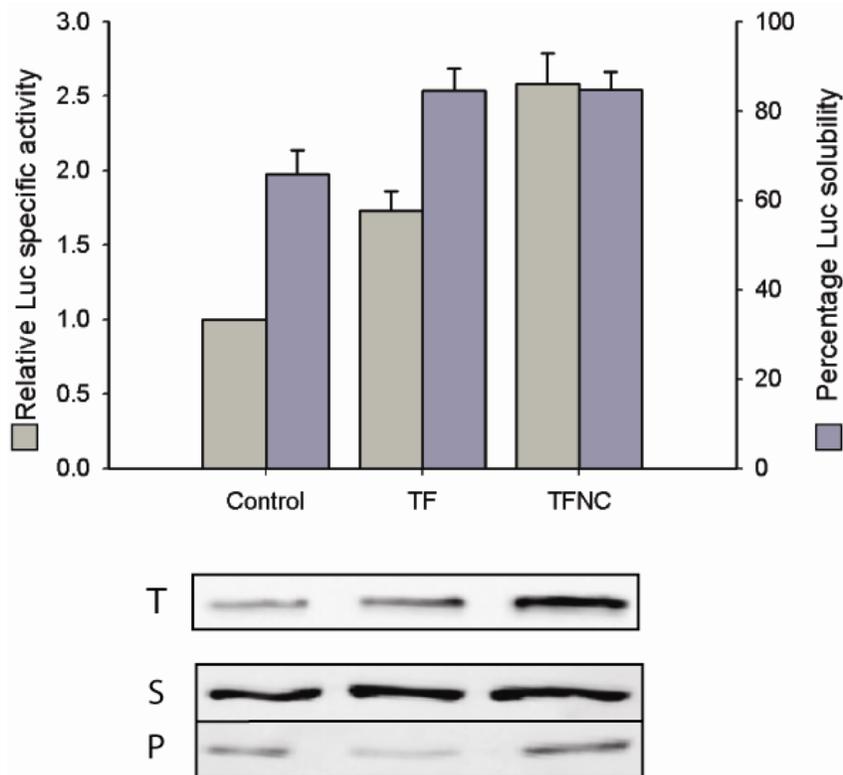


Figure 22: Folding yield of Luc in the PURE system

Luc was translated in PURE system in presence of 5 μ M TF or 5 μ M TFNC. Luc activity was measured in the luminometer after 1 h of translation. The amount of full-length Luc synthesized was determined by Western blotting against a myc-tag that was added to the C-terminus of Luc for easy quantification, followed by densitometry. The specific activity was calculated by dividing Luc activity by the total amount (T) of Luc synthesized. The Luc specific activity in the control reaction without added chaperone was set to 1 (grey bars). To determine the percentage solubility, reactions were centrifuged at 20,800 g for 30 min and separated into soluble (S) and pellet (P) fractions followed by Western blotting against myc-tag and densitometry (purple bars).

V.2.3 TFNC improves the folding of Luc in the Rapid Translation

System (RTS)

In order to understand the folding pathway mediated by TFNC, we analyzed the folding kinetics of Luc in the presence of either TF or TFNC. RTS is the system of *in vitro* translation that has previously been used to study the effect of TF on Luc folding kinetics (Agashe et al., 2004). We first analyzed the effect of TFNC on the specific activity of Luc in RTS. Consistent with the previous findings in the PURE system, TFNC

also enhanced folding of Luc in RTS. TF supplementation led to a ~1.3-fold increase in Luc specific activity compared to the unsupplemented RTS reaction, whereas addition of TFNC led to an increase of ~2.1-fold (Figure 23). These results emphasize the improved chaperone function of TFNC with respect to the eukaryotic multi-domain protein Luc, which is conserved in different translation systems.

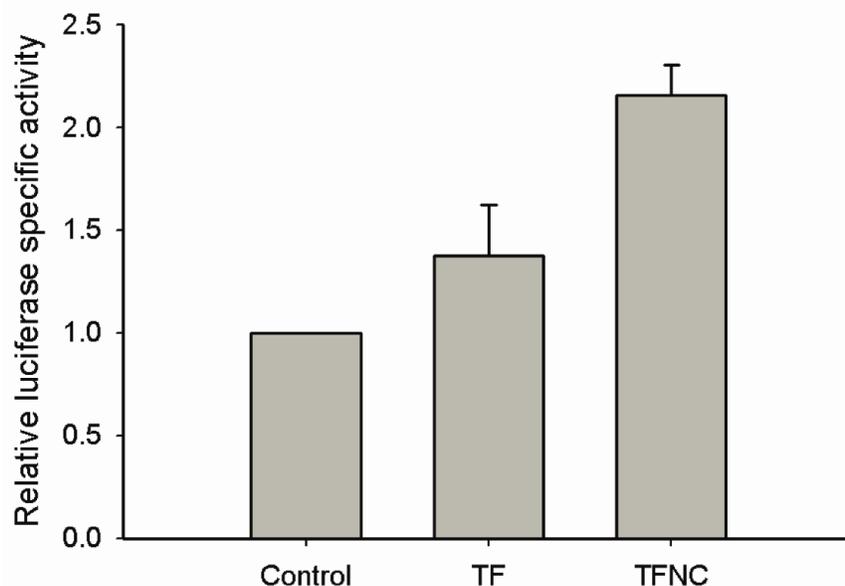


Figure 23: Folding yield of Luc in Rapid Translation System (RTS)

Luc was translated in RTS in the presence of either 5 μ M TF or 5 μ M TFNC. Luc activity was measured in the luminometer after 1 h of translation. The amount of full-length Luc synthesized was determined by Western blotting against myc-tag followed by densitometry. The specific activity was calculated by dividing Luc activity by the total amount (T) of Luc synthesized. The Luc specific activity in the control reaction without any added chaperone was set to 1.

V.2.4 TFNC-mediated folding is more co-translational

It has previously been shown that TF causes a delay in Luc folding relative to its translation whereas the default folding pathway in the absence of any chaperones is co-translational, albeit inefficient (Agashe et al., 2004). Upon translation in the unsupplemented *E. coli* RTS lysate, Luc folded concurrently with the translation of the Luc chains, indicating a co-translational mode of folding as shown previously (Agashe et al., 2004). However, upon addition of TF to the RTS reaction, Luc folding decelerated

with respect to its translation. This led to the conclusion that TF delays folding of Luc by binding and maintaining the Luc nascent chains in a folding competent and aggregation-free state until the completion of their translation. Such a folding mechanism results in an increased yield of folded Luc.

To study the effect of TFNC on the folding kinetics of Luc, we translated the protein in the presence of TF or TFNC in the RTS lysate and documented the appearance of Luc activity over time. The activities at the end of all the three reactions were set to one. In the reaction containing TF we observed a marked delay in the appearance of folded Luc compared to the unsupplemented reaction ($t_{1/2}$ ~27 min compared to ~22 min, respectively), whereas with TFNC the delay was less pronounced ($t_{1/2}$ ~24 min) (Figure 24, upper panel).

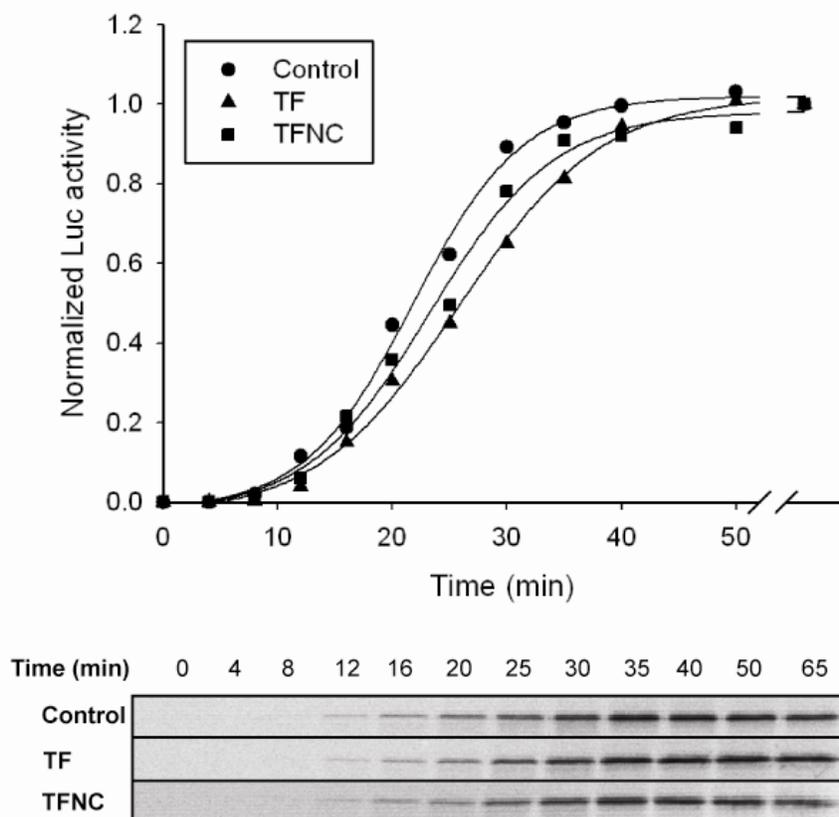


Figure 24: Kinetics of Luc folding in RTS

Luc was translated in the RTS in the presence of 5 μ M TF or 5 μ M TFNC and 35 S-methionine. Luc activity was measured in the luminometer at regular time points after translation initiation. The end point activities of all three reactions were normalized to 1

and plotted against time. Aliquots were taken at different time points, separated by SDS-PAGE and visualized by autoradiography.

This difference in appearance of Luc activity does not arise due to differences in the amount of Luc protein being translated, as the protein amounts were similar in all the reactions as demonstrated by the radio-labeling of the Luc nascent chains during translation (Figure 24, lower panel). This result indicates that the Luc folding pathway mediated by TFNC is relatively more co-translational.

V.2.5 TFNC improves the folding of Luc in Δtig *E. coli* cells

To study the physiological effect of TFNC on Luc folding, we used TF knockout *E. coli* cells. We compared the *de novo* folding of Luc during co-expression of either TF or TFNC in Δtig MC4100 cells. The chaperones were expressed from an IPTG inducible promoter prior to the expression of Luc from an arabinose inducible promoter. This strategy ensured a chaperone enriched cytosol at the time of Luc synthesis.

In the absence of any endogenous or overexpressed TF in Δtig cells, only 23% of total Luc translated was soluble and the specific activity of Luc was also reduced compared to cells in which TF was overexpressed. Similar to the observations in the PURE system, Luc folding was substantially improved with TFNC relative to TF. We observed a specific activity increase of 2.5-fold compared to the empty vector control, whereas with TF over-expression the increase was only 1.5-fold (Figure 25, grey bars). Notably, Luc was translated to similar levels in all the three transformed cells as was determined by Western blotting. TFNC also increased the fraction of soluble Luc compared to TF. The solubility of Luc was higher (~56%) in the presence of TFNC compared to TF (~39%) (Figure 25, purple bars).

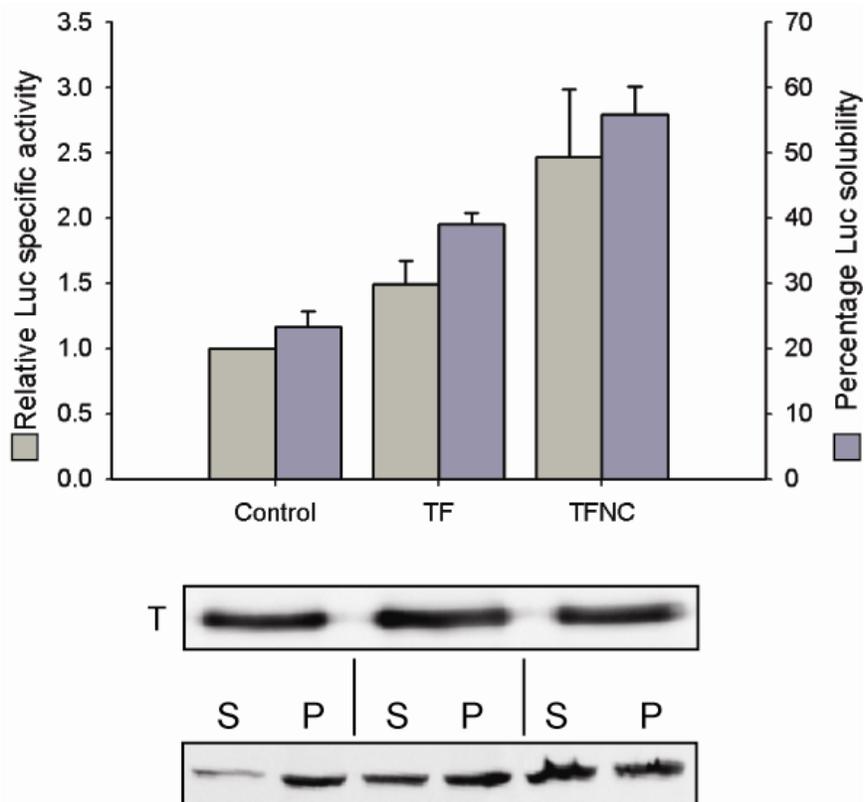


Figure 25: Folding yield of Luc in Δ tig MC4100 cells

pBAD18-Luc was cotransformed with pOFXlac-1, pOFXlac-1TF or pOFXlac1-TFNC in Δ tig MC4100 cells which were grown in LB media to an OD_{600nm} of 0.5 at 30 °C. Expression of TF or TFNC was induced by the addition of 1 mM IPTG for 30 min before the induction of Luc with 0.2% of arabinose for 30 min. Spheroplasts were produced and lysed hypo-osmotically. Luc activity was measured in the luminometer after 1 h of translation and the amount of full-length Luc synthesized was determined by Western blotting against myc-tag followed by densitometry. The specific activity was calculated as described in Figure 22 and represented by grey bars. The percentage solubility was also determined as described in Figure 22 (purple bars).

V.2.6 Folding of Ras-DHFR in presence of TF and TFNC

After establishing the efficiency of TFNC in mediating co-translational folding of Luc, the folding of another multi-domain protein which folds inefficiently in *E. coli* was studied. For this purpose we used an artificial fusion protein consisting of two single domain proteins, namely human H-Ras and mouse DHFR (Dihydro folate reductase). These two proteins were linked by a conserved 11 aa linker (Netzer and Hartl, 1997) (Figure 26). It was previously shown that this fusion protein folds inefficiently via a post-

translational folding pathway in *E. coli* cells and bacterial S30 lysate and via a co-translational mechanism in rabbit reticulocyte lysate (RRL). The individual domains however fold efficiently in *E. coli* (Netzer and Hartl, 1997).

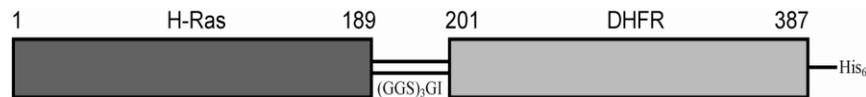


Figure 26: Linear representation of Ras-DHFR fusion protein

The Ras-DHFR fusion protein consists of human H-Ras and mouse DHFR (both single domain proteins) linked by a flexible 11 amino acids linker (GGS)₃GI. The C-terminus of the protein contains a His₆ tag (Netzer and Hartl, 1997).

V.2.6.1 TFNC improves the solubility of Ras-DHFR in Δ *tig* *E. coli* MC4100

Since the multi-domain Ras-DHFR fusion protein expressed in *E. coli* is insoluble, we tested if co-expression of the TFNC variant would improve its solubility in a way similar to Luc. To this end, TF or TFNC was co-expressed with Ras-DHFR in Δ *tig* *E. coli* MC4100. They were expressed for 30 minutes by IPTG addition prior to expression of Ras-DHFR by arabinose in order to provide a chaperone enriched cytosol prior to the the translation of Ras-DHFR nascent chains.

In the absence or presence of TF, Ras-DHFR was ~20-22% soluble whereas in presence of TFNC the solubility improved to ~31% (Figure 27), confirming the ability of TFNC to support the folding of diverse multi-domain proteins, as was observed with Luc.

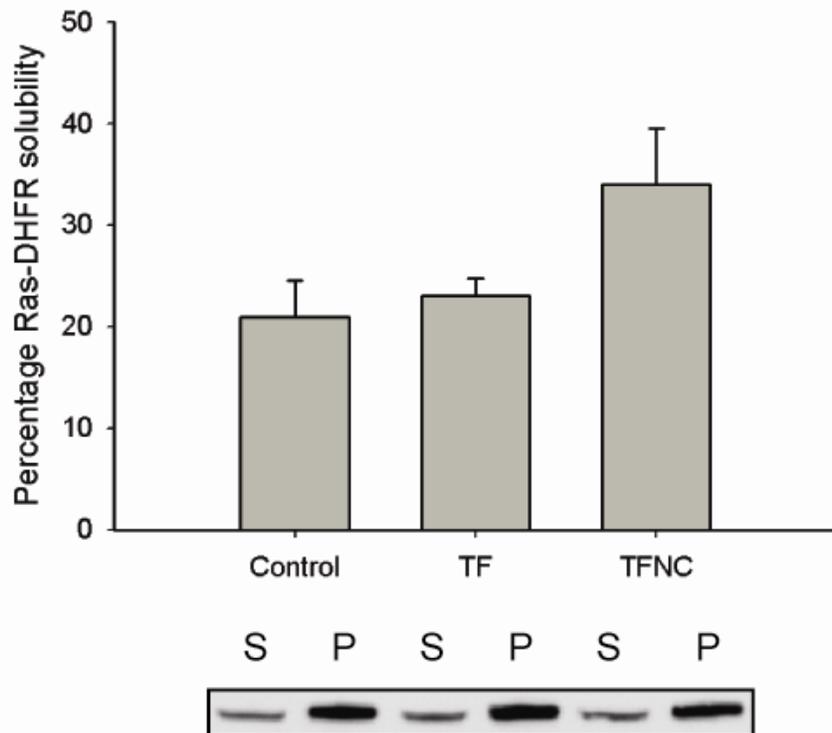


Figure 27: Folding yield of Ras-DHFR in Δ tig *E. coli* MC4100

pBAD18-Ras-DHFR-His₆ was cotransformed with either pOFXlac-1, pOFXlac-1TF or pOFXlac1-TFNC in Δ tig *E. coli* MC4100 and cells were grown in LB media to an OD_{600nm} of 0.5 at 30 °C. TF or TFNC were induced by 1 mM IPTG for 30 min before induction of Luc with 0.2% of arabinose for 30 min. Spheroplasts were produced and lysed hypo-osmotically. To determine the fraction of the soluble protein, reactions were centrifuged at 20,800 g for 30 min and separated into soluble (S) and pellet (P) fractions followed by Western blotting against His-tag and densitometry.

V.2.6.2 TFNC mediated efficient co-translational folding of Ras-DHFR

To gain insight into the TF and TFNC mediated folding mechanism of Ras-DHFR, the protease protection pattern of Ras-DHFR was monitored in the PURE system. It has previously been demonstrated that the correctly folded Ras-DHFR fusion protein is susceptible to cleavage by Proteinase K at the linker region yielding the individual Ras and DHFR domains, whereas the misfolded protein is either resistant to protease cleavage or is completely cleaved by the protease depending on the flexibility of the structure (Netzer and Hartl, 1997).

In the unsupplemented PURE system, after limited proteolysis part of the total Ras-DHFR protein was misfolded and another part properly folded as evidenced by the

appearance of the two bands corresponding to the individual domain after 31 min of translation (Figure 28, lanes 6-10). In the presence of TF, folding of Ras-DHFR was delayed with respect to the translation of the full length protein (Figure 28, lanes 6-10). In the absence of protease, the band for the full length protein first appears in the 24th min aliquot, whereas in the presence of protease the folded domains appear in the 31st min aliquot (Figure 28). In contrast, TFNC was able to mediate efficient and co-translational folding of Ras-DHFR. Properly folded Ras-DHFR is cleaved by the protease at the linker region and appears as two bands of folded domains after 24 min, the same time at which the full length protein begins to appear (Figure 28, lanes 6-10).

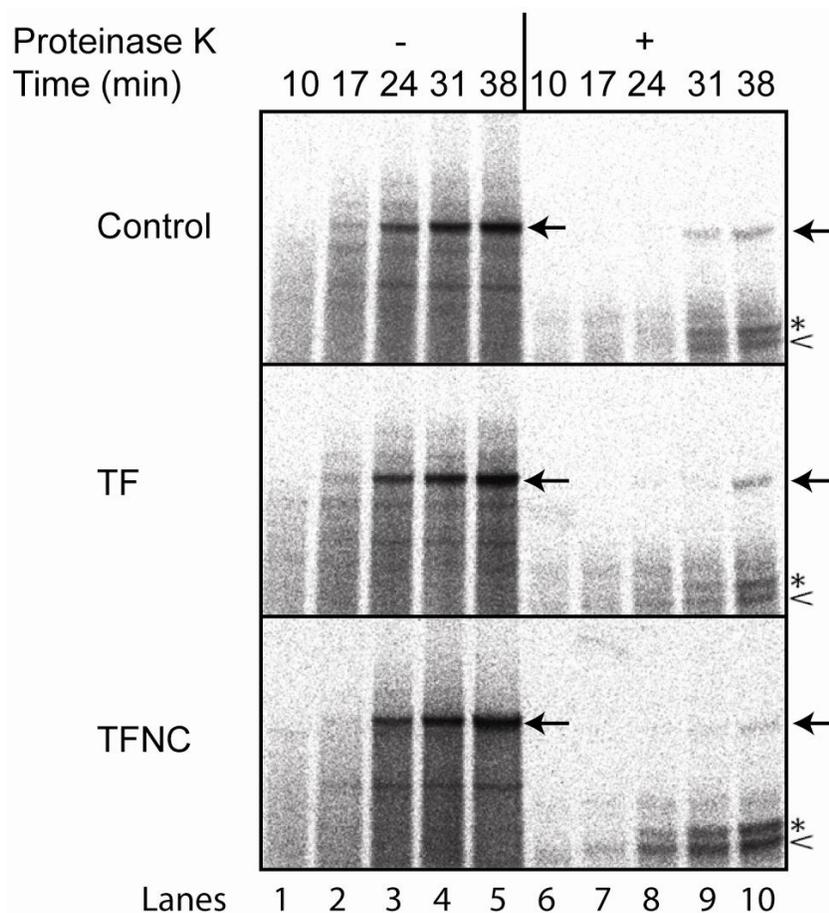


Figure 28: Limited proteolysis of Ras-DHFR

Ras-DHFR was translated in the absence or presence of TF or TFNC in the PURE system. The aliquots were incubated in the absence (lanes 1-5) or presence of Proteinase K (lanes 6-10) at different time points. Samples were TCA precipitated, separated by SDS-PAGE and visualized by autoradiography. Full-length Ras-DHFR

protein is indicated by an arrow, protease resistant DHFR domain is indicated by an asterisk and protease resistant Ras domain is indicated by an arrowhead.

To quantify the extent of Ras-DHFR folding in these reactions, the density of the DHFR band in lane 10 was represented as a fraction of density contributed by full length Ras-DHFR in lane 5. We found that in the presence of TFNC 68±8% of the fusion protein had a folded DHFR domain, compared to only 38±10% in presence of TF (Figure 29).

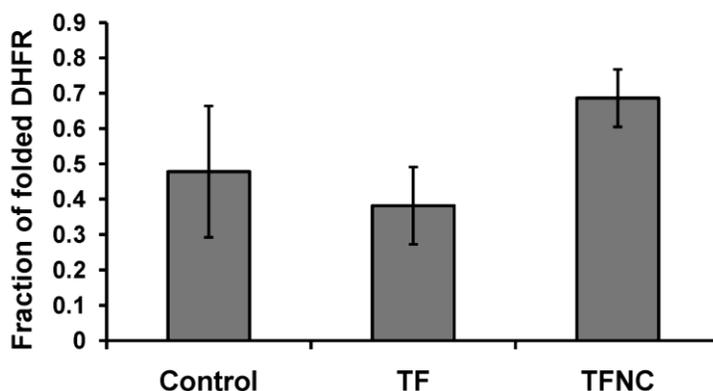


Figure 29: Quantification of folded DHFR domain in presence of TF or TFNC

The densities of DHFR bands in lane 10 of Figure 23 are presented as fraction of the densities contributed by DHFR in the full-length fusion protein of lane 5. The total number of cysteines and methionines in full length protein and DHFR is 19 and 8, respectively.

V.2.6.3 Effect of TF and TFNC on the folding of the Ras domain

To understand the effect of TF or TFNC on the folding of the individual Ras domain, a stalled Ras construct with additional 50 aa (Ras+50) from the fusion protein was utilized. This construct allows the exposure of the complete Ras domain outside the ribosome exit tunnel. In all the three PURE system reactions (unsupplemented control, TF supplemented and TFNC supplemented) the Ras domain appeared as a correctly folded, protease-resistant domain in the 17 min aliquot, i.e. at the same time as the appearance of the full-length Ras+50 nascent chain (Figure 30, lanes 6-10). This suggests that the Ras+50 nascent chains are able to undergo efficient co-translational folding

irrespective of the presence of TF or TFNC. These results demonstrate the efficiency of TFNC to mediate the productive folding of eukaryotic proteins multi-domain proteins. In contrast, single domain proteins might not require an active assistance of TF or TFNC.

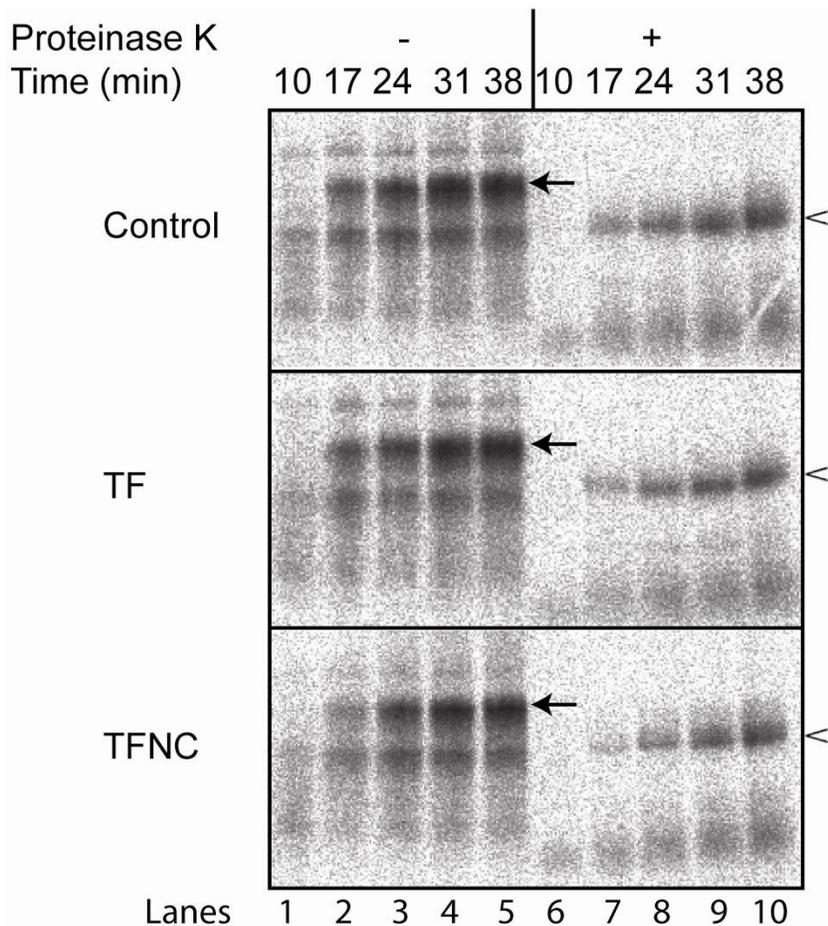


Figure 30: Limited proteolysis of Ras domain

Ras was translated in the absence or presence of TF or TFNC in the PURE system. The aliquots were incubated in the absence (lanes 1-5) or presence of Proteinase K (lanes 6-10) at different time points. Samples were TCA precipitated, separated by SDS-PAGE and visualized by autoradiography. Full-length Ras domain is indicated by an arrow and protease resistant Ras domain is indicated by an arrowhead.

V.3 Eukaryotic chaperones mediate efficient folding of Luc

It has been shown earlier that the eukaryotic cytosol is enriched in chaperones that mediate efficient Luc folding (Frydman et al., 1994), whereas bacterial chaperones are

unable to do so (Agashe et al., 2004). The recombinant over-expression of eukaryotic proteins in *E. coli* is often marked by protein instability and aggregation (Baneyx, 1999; Baneyx and Mujacic, 2004). Hence a possible strategy to improve the yield of such proteins in *E. coli* could be their co-expression with eukaryotic chaperones.

V.3.1 Eukaryotic *in vitro* translation system is capable of efficient Luc folding

We first compared the Luc folding capabilities of a bacterial and eukaryotic *in vitro* translation system. To this end, we utilized the PURE system which is composed of purified components of bacterial translation machinery and the Rabbit Reticulocyte Lysate (RRL) which contains the chaperones Hsp70, Hsp40, TRiC and Hsp90 (Frydman et al., 1994). RRL is often used to demonstrate the chaperone dependent *de novo* folding and refolding of Luc (Frydman et al., 1994; Tzankov et al., 2008). Luc was translated from the same vector pET3a-Luc in both the systems under identical conditions. The translation rate in RRL was slower as compared to the PURE system. From a comparison of specific activities of Luc produced in both systems, it was evident that the efficiency of Luc folding was ~12 fold higher in the RRL than in the PURE system (Figure 31 A).

Next, we wanted to test whether the increased Luc folding in RRL was due to the presence of chaperones or the slower translation rate in RRL, or a combination of both factors. Luc was translated in the PURE system in presence of various amounts of added RRL which provided the chaperones. The reactions were carried out in the presence of 2 mM cycloheximide to specifically inhibit translation from ribosomes present in RRL. Surprisingly, addition of RRL in PURE system did not cause a significant improvement in Luc specific activity (Figure 31 B). There could be several reasons for this observation: (1) For example, the amount of chaperones present in the RRL supplementation could be insufficient or the chaperones could be inactive in the PURE system. (2) It is also possible that the slower translation rate may be important for the efficient Luc folding in RRL.

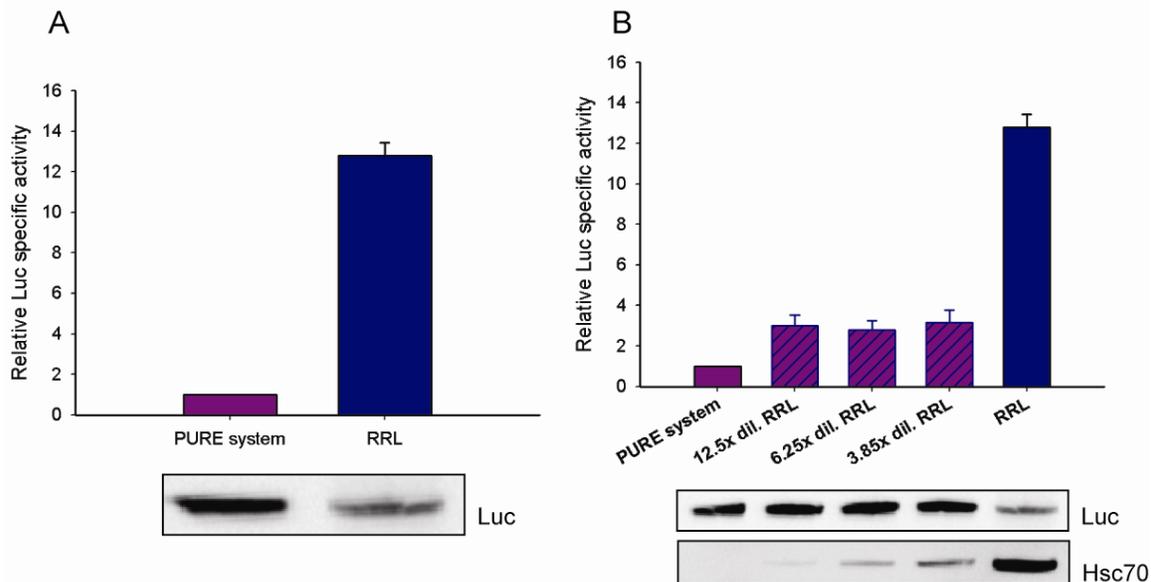


Figure 31: Comparison of Luc folding in PURE system and Rabbit Reticulocyte Lysate (RRL)

(A) Luc was translated in PURE system and RRL at 30 °C from pET3a-Luc. Luc activity was measured in the luminometer after 1 h of translation. The amount of Luc synthesized was determined by Western blotting against myc-tag followed by densitometry. The specific activity was calculated as described earlier. The Luc specific activity in the PURE system reaction was set to 1. (B) Luc was translated in the PURE system, PURE system with 12.5/6.25/3.85 fold diluted RRL (in presence of cycloheximide) and RRL at 30 °C from pET3a-Luc. Luc activity was measured in the luminometer after 1 h of translation and the amount of full-length Luc synthesized was determined by Western blotting. The specific activity was calculated as described previously. The Luc specific activity in the PURE system reaction was set to 1. The endogenous levels of Hsc70 in these reactions were detected by Western blotting.

V.3.2 Supplementation of purified eukaryotic chaperones in PURE system enhance the folding of Luc

We have shown above that TFNC increased Luc specific activity more than TF (Figure 22). However, despite this increase the efficiency of folding did not reach that of folding in RRL (Figure 31 A) which is an enriched eukaryotic *in vitro* translation system. This indicates that the chaperone machinery present in RRL is highly efficient in folding Luc. Hence, we supplemented the PURE system with members of the eukaryotic Hsp70 chaperone machinery (Hsc70, Hdj2 and Bag1). Hsc70 is the constitutively expressed Hsp70 in the eukaryotic cytosol and Hdj2 is the type A Hsp40 protein having close

homology to bacterial DnaJ (Mayer and Bukau, 2005; Minami et al., 1996; Terada et al., 1997; Tzankov et al., 2008). Bag1 is the Bcl-associated athanogene 1, which performs the nucleotide exchange function for Hsc70 (Hohfeld and Jentsch, 1997; Takayama et al., 1997; Tzankov et al., 2008). All the three proteins were of human origin and were purified upon recombinant expression in *E. coli* (Figure 32).

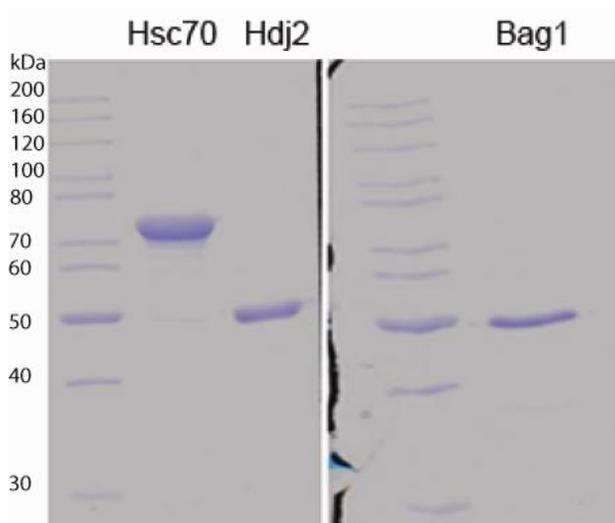


Figure 32: Purified eukaryotic chaperones Hsc70, Hdj2 and Bag1 used for supplementation in PURE system.

Human Hsc70 was purified as a His-tagged protein via metal-affinity chromatography. Hdj2 and Bag1 are from the laboratory stock.

In order to optimize the folding of Luc in the PURE system in the presence of eukaryotic chaperones, we performed titration experiments to find the optimal combination of Hsc70, Hdj2 and Bag1 that mediates efficient folding of Luc. We first performed a titration of Hsc70 in the PURE system and analyzed the specific activity of Luc. At 10 μM of Hsc70, a ~ 5 -fold improvement in Luc specific activity over the control PURE system reaction was observed (Figure 33 A). When 1 μM of Hdj2 in combination with 10 μM Hsc70 was used, the improvement was ~ 6 fold as compared to the control reaction. Addition of Hdj2 at concentrations higher than 1 μM inhibited Luc folding (Figure 33 B). Further supplementation of 6 μM of Bag1 to this combination in the PURE system led to an increase in the specific activity of Luc up to ~ 9.5 fold as compared to the control (Figure 33 C). These results suggest that a ratio of 10 μM Hsc70:

1 μM Hdj2: 6 μM Bag1 is close to optimal in mediating efficient *de novo* folding of Luc in the PURE system under our experimental conditions.

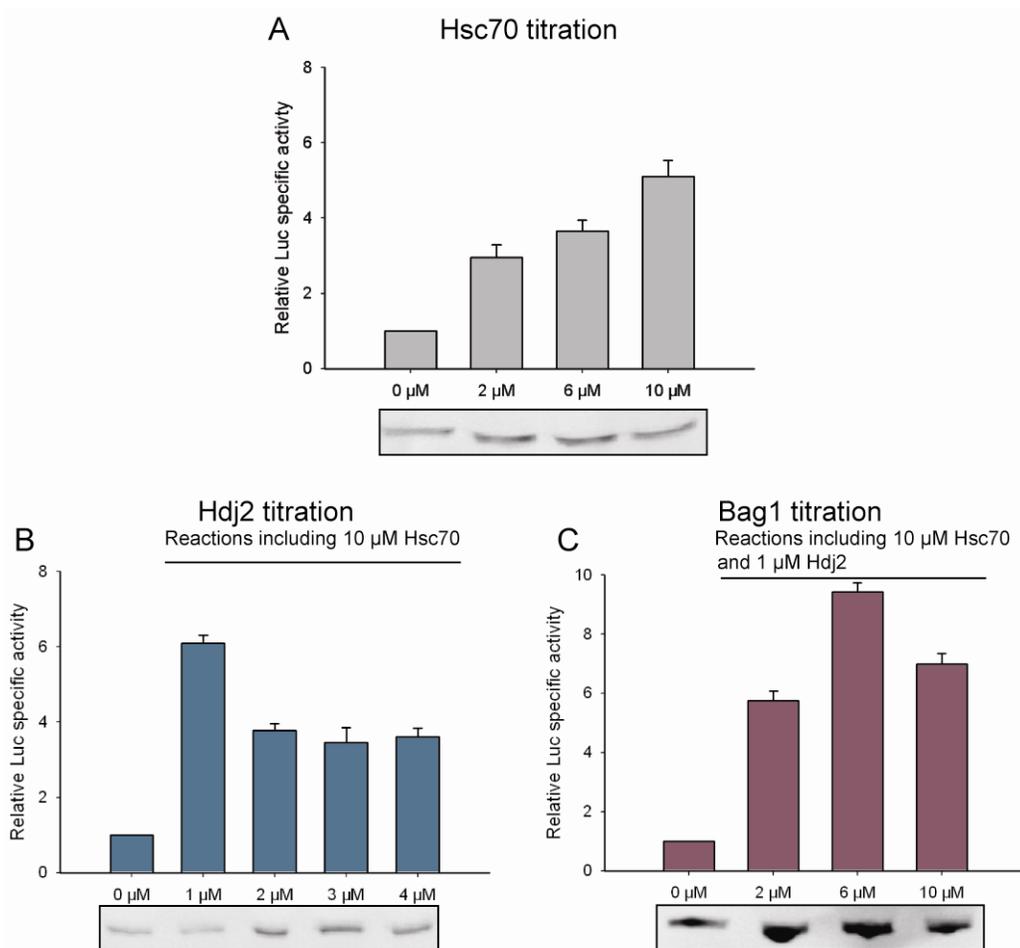


Figure 33: Optimization of Hsc70, Hdj2 and Bag1 concentrations for efficient Luc folding

(A) Luc was translated in the PURE system in presence of various concentrations of Hsc70 and Luc specific activity was calculated as described earlier. The Luc specific activity in the reaction without any added chaperone was set to 1. (B) Luc was translated in presence of 10 μM Hsc70 and various concentrations of Hdj2 and specific activity was analyzed. (C) Luc was translated in presence of 10 μM Hsc70, 1 μM Hdj2 and various concentrations of Bag1. Luc specific activity in various reactions was compared.

V.3.3 Eukaryotic Hsp70 system promotes efficient Luc folding in PURE system

Having determined the optimal concentrations of Hsc70, Hdj2 and Bag1 to mediate efficient Luc folding, we performed the Luc folding assays with individual components of the Hsp70 machinery in combination with TFNC. We achieved a ~10 fold increase in Luc specific activity compared to the unsupplemented control when 10 μ M Hsc70, 1 μ M Hdj2 and 6 μ M Bag1 were added to the PURE system. Interestingly, addition of TFNC together with Hsc70, Hdj2 and Bag1 led to further enhancement of Luc folding, increasing Luc specific activity by up to ~12 fold as compared to the unsupplemented PURE system, reaching a folding efficiency similar to that achieved upon translation in RRL (Figure 34).

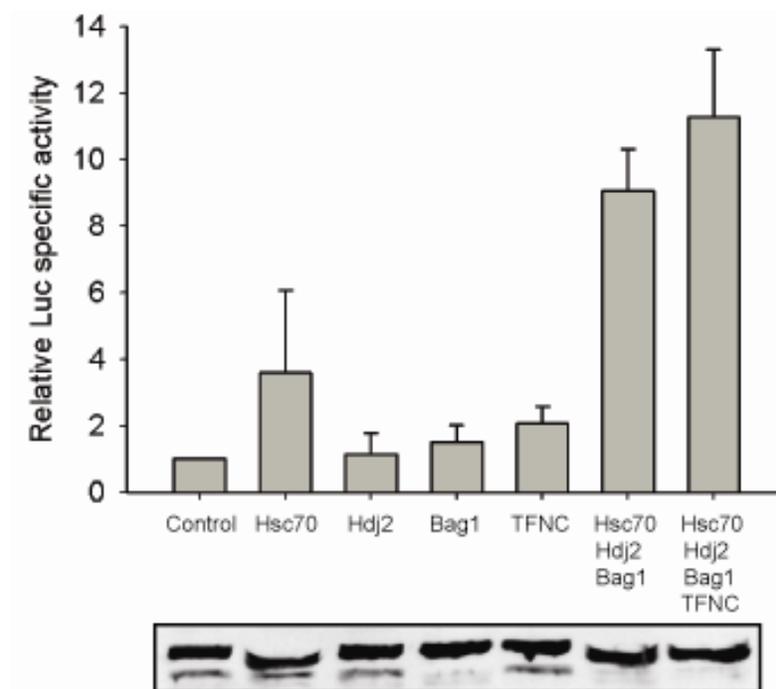


Figure 34: Effect of eukaryotic chaperones together with TFNC on Luc folding in PURE system

Luc was translated in PURE system in presence of 10 μ M Hsc70, 1 μ M Hdj2, 6 μ M Bag1 and 5 μ M TFNC, separately or in combination, as shown in the figure. Luc activity was measured in the luminometer after 1 h of translation and the amount of full-length Luc synthesized was determined by Western blotting against the myc-tag, followed by densitometry. The specific activity was calculated as described previously. The Luc

specific activity in the control reaction in the absence of any added chaperone was set to 1.

The combination of Hsc70, Hdj2, Bag1 and TFNC also improved the solubility of Luc significantly. When Luc was translated in the PURE system in the absence of any chaperones, only ~20% was soluble, whereas supplementation of Hsc70, Hdj2 and Bag1 led to an improved solubility of ~70%. The maximal solubility of more than 90% was observed when the above mentioned eukaryotic Hsp70 system was supplemented by TFNC (Figure 35). Together these results suggest that the eukaryotic chaperone machinery can create an efficient folding environment in PURE system which is similar to the folding environment in a eukaryotic *in vitro* translation system (RRL). Hence this strategy could have possible implications in recombinant protein production since most of the biotechnologically important eukaryotic proteins are modular in structure.

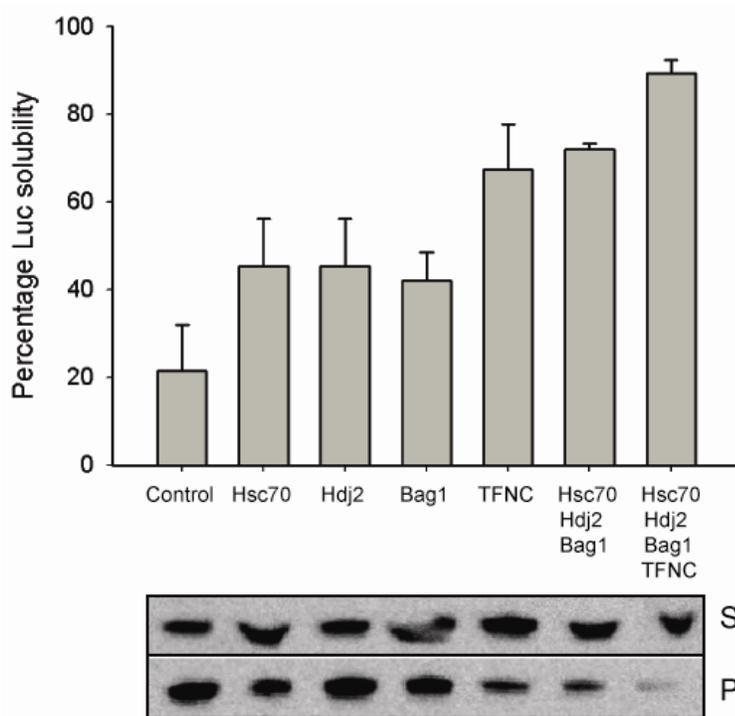


Figure 35: Effect of eukaryotic chaperones together with TFNC on Luc solubility in PURE system

Luc was translated in PURE system in presence of 10 μ M Hsc70, 1 μ M Hdj2, 6 μ M Bag1 and 5 μ M TFNC, separately or in combination as shown in the figure. For determining the percentage solubility, reactions were centrifuged at 20,800 g for 30 min and

separated into soluble (S) and pellet (P) fractions, followed by Western blotting against myc-tag and densitometry.

V.3.4 The eukaryotic Hsp70 machinery mediates efficient co-translational folding of Luc in PURE system

To determine whether supplementation with the eukaryotic chaperones resulted in co-translational folding in the PURE system, we used Luc as the model substrate and analyzed the post-translational increase in its activity after translation inhibition. Luc was translated in the PURE system with various combinations of chaperones for 30 minutes, followed by translation inhibition by addition of RNaseA. RNaseA disrupts the integrity of the ribosome by digesting the single stranded RNA present in the ribosomes and hence attenuates the translation process. If folding is co-translational, no significant increase in Luc activity is expected after translation inhibition, as shown previously (Agashe et al., 2004).

In the control reaction of unsupplemented PURE system, Luc folds in a co-translational manner by default. Translation of Luc in presence of TF led to a post-translational increase in its activity upto 1.4-fold for additional 15 min after the inhibition of its translation (Figure 36). These observations are consistent with the previously published data that Luc folds co-translationally in the upsupplemented RTS reaction but post-translationally in the presence of TF (Agashe et al., 2004). In the presence of TFNC, the eukaryotic Hsp70 components or both, Luc folding ceased immediately after RNaseA treatment, indicating a co-translational folding mechanism (Figure 36).

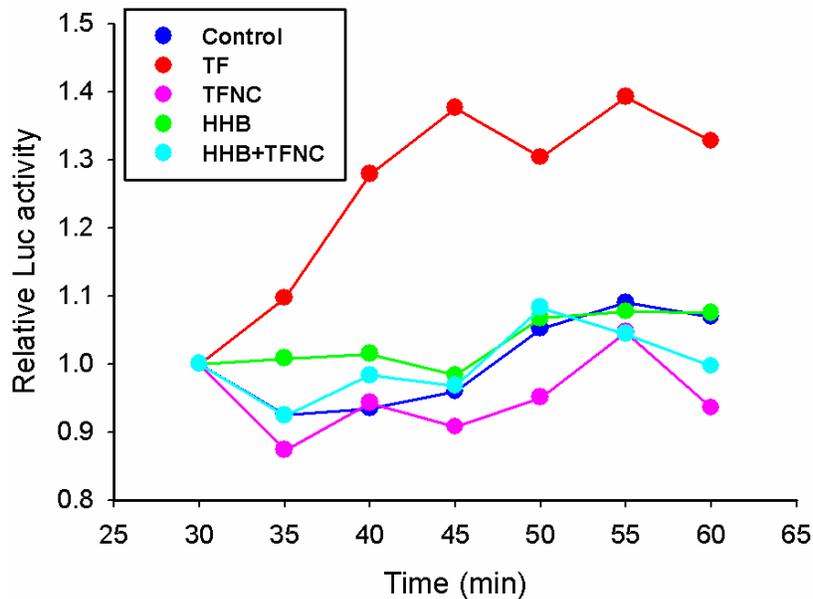


Figure 36: Co-translational Luc folding mediated by eukaryotic Hsp70 system and TFNC.

Luc was translated in the absence of added chaperone (control) or in the presence of 10 μM Hsc70, 1 μM Hdj2, 6 μM Bag1 and 5 μM TFNC in PURE system and the reactions were stopped after 30 min by addition of 50 $\mu\text{g/ml}$ RNaseA. The increase in Luc activity after translation inhibition was followed over time. HHB denotes Hsc70, Hdj2, Bag1. Representative result of 3 independent experiments.

V.4 Generation of an *in vivo* co-expression system to test folding of multi-domain proteins in presence of eukaryotic chaperones

Having shown in the previous section that the eukaryotic Hsp70 chaperone machinery and TFNC in combination can promote efficient folding of Luc in the PURE system, we wanted to analyze their chaperoning action in the cellular environment. *E. coli* BL21 (DE3) strain which contains endogenous TF was chosen as a host, since it is the most commonly used *E. coli* strain used for recombinant protein production.

In order to test the folding of a diverse set of eukaryotic multi-domain proteins in the presence of eukaryotic chaperones, we generated an *in vivo* co-expression system. A set of four vectors with compatible origins were used to clone and express the chaperones

with the substrate protein to be analyzed. All the vectors have different antibiotic markers and origins of replication. The vectors encoding the chaperones are IPTG inducible whereas pBAD33-Luc is arabinose inducible. We analyzed the expression of these chaperones with IPTG induction and found that there was no leaky expression in the absence of IPTG (Figure 37). The expression levels of eukaryotic chaperones were considerably lower than that of TFNC. First, the effect of eukaryotic chaperones on Luc folding *in vivo* was studied systematically by co-expressing chaperones individually or in various combinations.

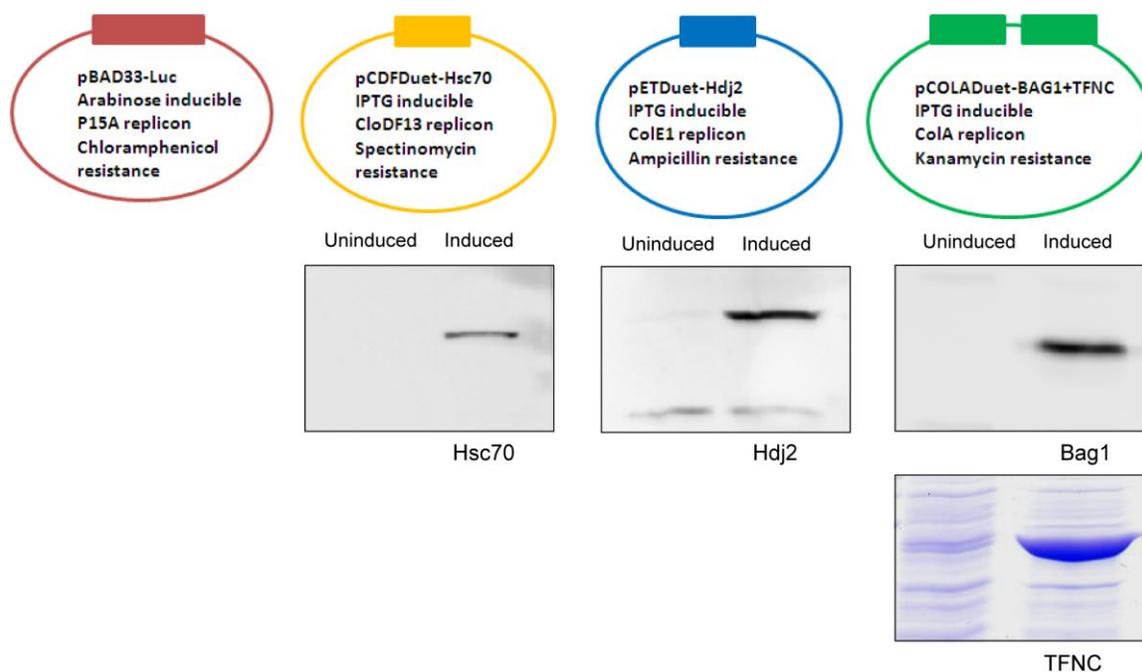


Figure 37: Co-expression system to test the folding of Luc in *E. coli* BL21 (DE3) cells

Four vectors with different antibiotic markers and origins of replication were chosen to generate a co-expression system to test Luc folding in BL21 (DE3) cells. pBAD33 encodes for Luc under arabinose induction. pCDFDuet, pETDuet and pCOLADuet encode for Hsc70, Hdj2 and Bag1 together with TFNC, respectively, under IPTG induction.

V.4.1 Effect of TFNC on the folding of Luc in BL21 (DE3) cells

In order to study the effect of TFNC on Luc folding in the presence of endogenous TF, TFNC and Luc were co-transformed in BL21 (DE3) cells. On IPTG

addition, TFNC was ~20 fold more expressed than the endogenous TF. We did not observe any improvement in the folding of Luc under such conditions (Figure 38), suggesting that the endogenous TF probably binds to the translating ribosomes with higher affinity than TFNC and, in the absence of ribosome recruitment, TFNC could not mediate Luc folding. Another reason could be that TFNC cannot properly cooperate with the endogenous *E. coli* Hsp70, DnaK.

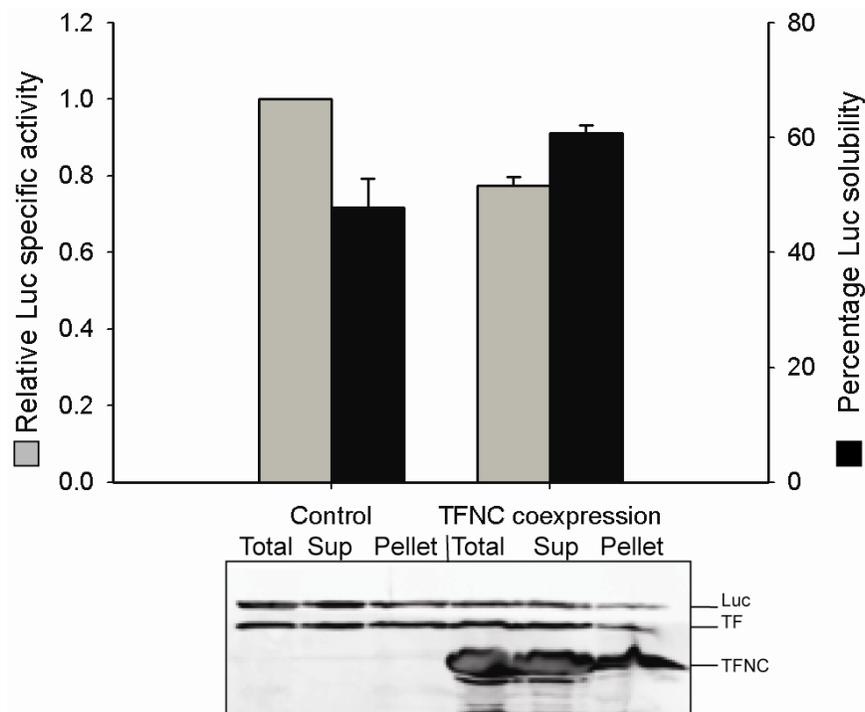


Figure 38: Effect of TFNC on the folding of Luc in BL21 (DE3) cells

BL21 (DE3) cells were co-transformed with pBAD33-Luc and pCOLADUET-TFNC and were grown in LB media till an OD_{600nm} of 0.5 at 30 °C. TFNC was induced by 1 mM IPTG for 30 min before induction of Luc with 0.2% of arabinose for 30 min. Spheroplasts were produced and lysed hypo-osmotically. Luc specific activity, solubility of the control (without chaperone expression) and TFNC co-expression samples were determined as mentioned previously in methods section.

V.4.2 Effect of Hsc70 on the folding of Luc in BL21 (DE3) cells

Next, we tested the effect of human Hsc70 co-expression on Luc folding in BL21 (DE3) cells. Hsc70 was expressed prior to the expression of Luc to ensure the availability of excess Hsc70 to the translating Luc nascent chains. We observed that the Luc

expressed under Hsc70 co-expression was fully soluble as compared to 40% solubility under normal condition. However, the specific activity of Luc decreased with Hsc70 co-expression (Figure 39).

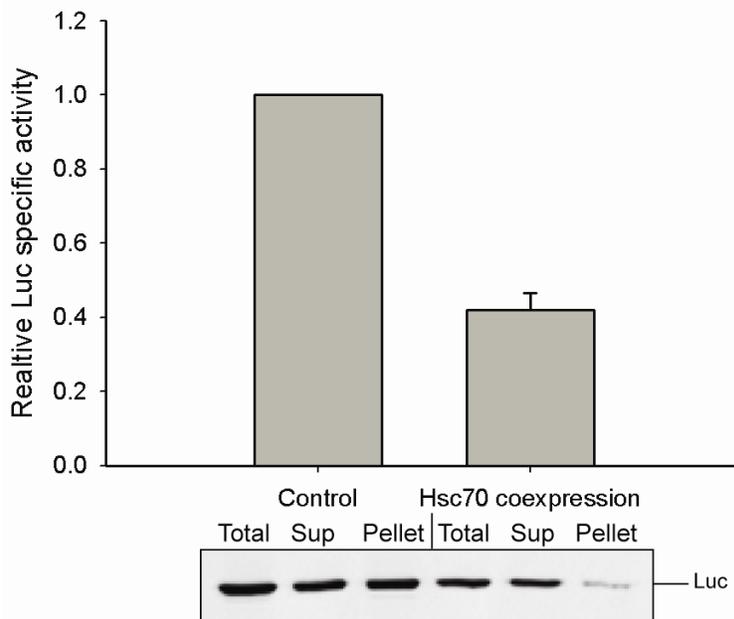


Figure 39: Effect of Hsc70 on folding of Luc in BL21 (DE3) cells

BL21 (DE3) cells were co-transformed with pBAD33-Luc and pCDFDUET-Hsc70 and were grown in LB media to an OD_{600nm} of 0.5 at 30 °C. Hsc70 was induced by 1 mM IPTG for 30 min before induction of Luc with 0.2% of arabinose for 30 min. Spheroplasts were produced and lysed hypo-osmotically. Luc specific activity of the control (without chaperone expression) and Hsc70 co-expression samples was determined as mentioned previously.

This indicates that Hsc70 probably acts as a holdase and keeps Luc in a soluble but enzymatically-inactive state. In the absence of the cochaperones of Hsc70, such a state might not be allowed to fold to the native protein. Clearly, co-expression of Hsc70 with Luc led to an incomplete Hsp70 folding cycle in the absence of its cochaperones Hdj2 and Bag1.

V.4.3 Imbalance in the expression levels of eukaryotic Hsp70 chaperone machinery with TFNC causes inhibition of Luc translation

To study the effect of the complete eukaryotic Hsp70 machinery together with TFNC, we co-transformed these chaperones with Luc in BL21 (DE3) cells. As previously described, the chaperones were expressed prior to the expression of Luc to ensure a chaperone enriched cytosol. Although there was an improvement in the solubility of Luc, the translation of Luc was severely decreased (Figure 40 A).

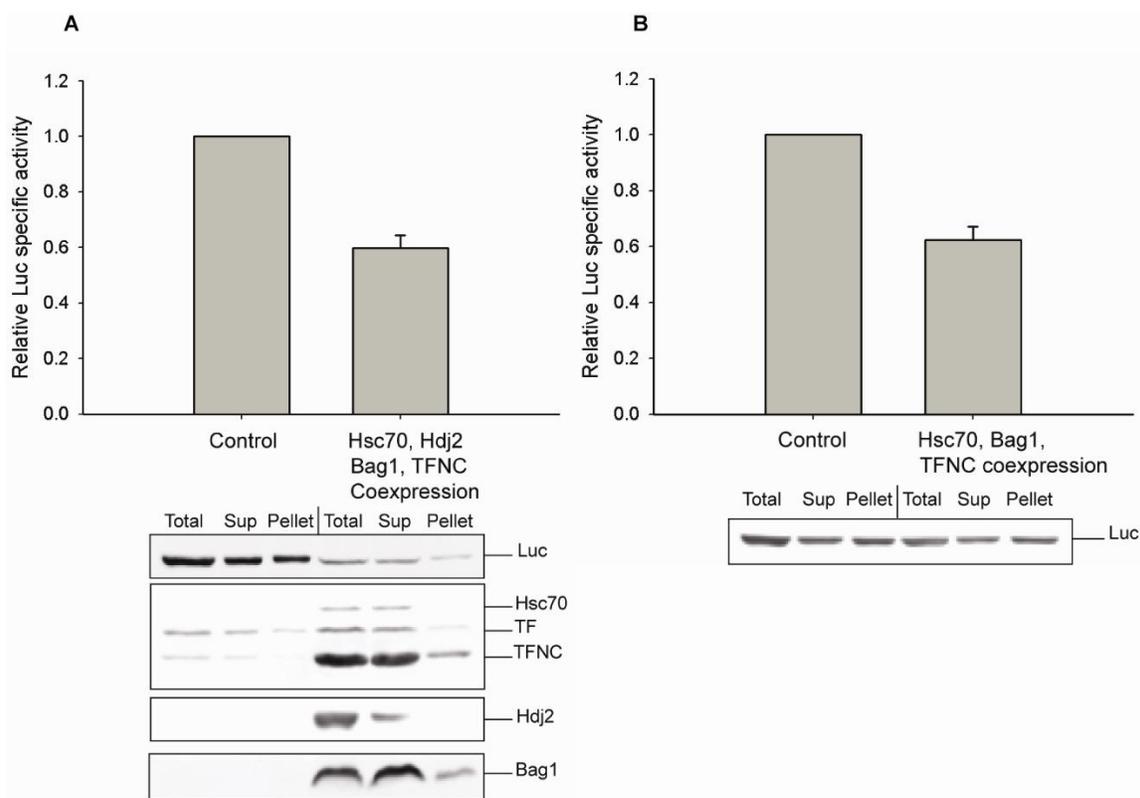


Figure 40: Effect of eukaryotic chaperones and TFNC on folding of Luc in BL21 (DE3) cells

(A) BL21 (DE3) cells were co-transformed with pBAD33-Luc, pCDFDUET-Hsc70, pCOLADUET-Bag1,TFNC and pETDUET-Hdj2 and were grown in LB media to an OD_{600nm} of 0.5 at 30 °C. (B) BL21 (DE3) cells were co-transformed with pBAD33-Luc, pCDFDUET-Hsc70 and pCOLADUET-Bag1,TFNC and were grown in LB media to an OD_{600nm} of 0.5 at 30 °C. Chaperones were induced by 1 mM IPTG for 30 min before induction of Luc with 0.2% of arabinose for 30 min. Spheroplasts were produced and lysed hypo-osmotically. Luc specific activity of the control (without chaperone expression) and chaperone co-expression samples was determined as mentioned previously.

We also observed a decrease in growth rate of the cells after induction of chaperones. The reason for this growth defect and translation inhibition is not clear, but we speculate that the co-expression of five different proteins together might lead to metabolic stress to the cell. Also, it might be possible that the over-expressed eukaryotic chaperones may interfere with the balance in the cell causing a translation inhibition. To test this hypothesis, we performed the same experiment with the omission of Hdj2 from the co-expression system as we speculated that over-expression of Hsp40 proteins might have a toxic effect on growing cells. Co-expression of Hsc70, Bag1 and TFNC with Luc did not cause a significant reduction in translation efficiency. However, this combination also did not lead to any improvement in Luc specific activity and solubility (Figure 40 B). Thus it appears that we failed to achieve the proper balance between the various components of the eukaryotic Hsp70 system necessary for efficient folding assistance.

V.4.4 Hdj2 is responsible for the inhibition of Luc translation *in vivo*

To test whether over-expression of eukaryotic Hsp40 is toxic for the *E. coli* cells, we co-expressed Hdj2 (eukaryotic type I Hsp40) with Luc in BL21 (DE3) cells. As observed in the experiments described above, Hdj2 co-expression indeed decreased the translation of Luc (Figures 41 A and 41 C). We also observed a marked decrease in the growth rate of *E. coli* cells upon Hdj2 induction (Figure 41 B).

Overall, these results indicate that the relative expression levels of members of the Hsp70 system need to be carefully adjusted in order to achieve a significant improvement in Luc folding as observed in PURE system. Moreover, the endogenous TF present in BL21 (DE3) cells might interfere with the efficient co-translational folding of Luc. Hence in future studies, the chromosomally encoded TF should be deleted in this strain to systematically analyze the role of eukaryotic chaperones on the folding of multi-domain proteins in *E. coli*.

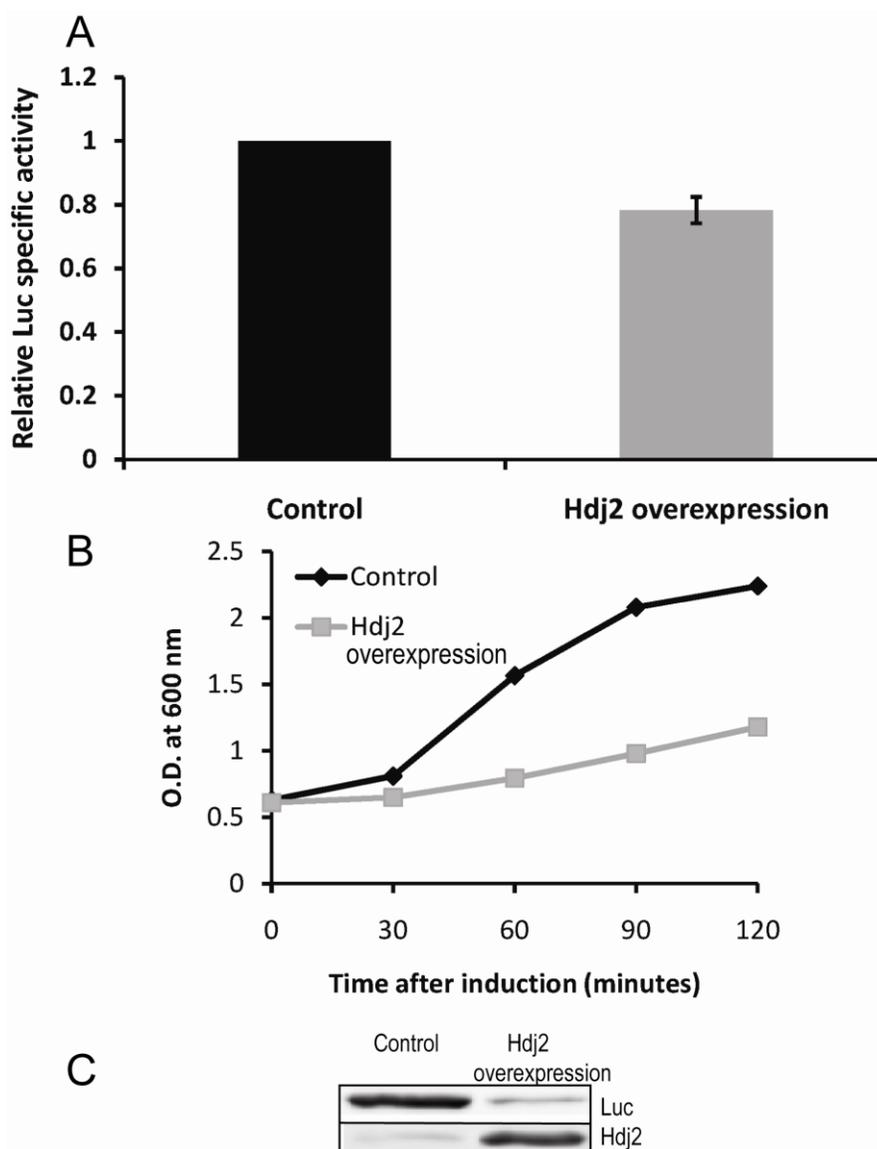


Figure 41: Inhibition of Luc translation in BL21 (DE3) cells upon Hdj2 induction

(A) Hdj2 was co-expressed with Luc in BL21 (DE3) cells from the plasmid pETDUET-Hdj2. Hdj2 was induced by 1 mM IPTG for 30 min before induction of Luc with 0.2% of arabinose for 30 min (C). In the control where no IPTG was added to induce Hdj2, some leaky expression of Hdj2 could be observed. The cells were lysed and Luc specific activity was measured as explained previously. (B) Cells co-transformed with Hdj2 and Luc were induced by addition of IPTG at OD₆₀₀ of 0.6 and the O.D. was monitored at regular time intervals after IPTG addition. In the control cell culture no IPTG was added. (C) Western blot of Luc and Hdj2 from (A).

VI DISCUSSION

Proteins have to undergo proper folding after translation in order to attain their final native structure. Molecular chaperones assist in this process without being part of a protein's final structure. The ribosome-associated chaperone TF of *E. coli* plays an important role in the folding of nascent chains during translation. The interaction of TF with non-translating and translating ribosomes have previously been studied in detail (Kaiser et al., 2006; Maier et al., 2003; Rutkowska et al., 2007).

In this thesis, we have investigated the direct interaction of TF with nascent chains using TF site-specifically labeled with the fluorophore NBD (*N*-((2-(iodoacetoxy)ethyl)-*N*-methyl)amino-7-nitrobenz-2-oxa-1,3-diazole (IANBD ester)) at various positions. We found that TF interacts with nascent chains depending on hydrophobic segments exposed in the linear sequence. Our study also showed that the PPIase domain of TF is responsible for the delay in folding relative to translation of certain multi-domain proteins in *E. coli*. Specifically, we compared the effect of TF and its PPIase domain-deletion mutant TFNC on the folding of two eukaryotic multi-domain proteins which were shown to fold post-translationally in *E. coli*. We then studied how modulating the interaction of TF with nascent chains of eukaryotic multi-domain proteins reduces their aggregation propensity and improves their folding yield and found that TFNC improves the folding yield of these proteins by shifting the folding to a relatively more co-translational pathway.

Although TFNC was able to improve the folding of firefly luciferase (Luc) in the PURE system to a certain extent, the increase was still not comparable to the highly efficient folding of Luc in a eukaryotic reticulocyte lysate (RRL) *in vitro* translation system. We therefore concluded that the chaperone machinery present in the RRL could fold Luc more efficiently than the modified prokaryotic chaperone, TFNC. Upon systematic titration of the purified eukaryotic chaperones into the bacterial PURE system, we found that the eukaryotic Hsp70 system significantly increased the yield of folded Luc. The results of this study help to better understand the mechanisms of co-

translational folding and suggest strategies for improving recombinant protein production in biotechnology.

VI.1 TF interaction with nascent chains is based on the presence of hydrophobic motifs in the primary sequence

Previously published intra-molecular FRET experiments demonstrated that TF attains an open conformation upon binding to the ribosome. Upon binding to Luc nascent chains, TF relaxes back to its compact conformation with a $t_{1/2}$ of ~ 35 s whereas it dissociates from the ribosomes with a $t_{1/2}$ of ~ 10 - 12 s (Kaiser et al., 2006). This study suggested that TF maintains its association with the nascent chains even after it had dissociated from the ribosomes. However, a direct analysis of the direct interaction of TF with nascent chains was lacking. In the present study, we employed fluorescence spectroscopy to analyze the interaction of TF with nascent chains directly.

TF was site-specifically labeled with NBD (an environmentally-sensitive probe) at various positions and the binding of these labeled proteins with nascent chains exposing variable numbers of hydrophobic regions was monitored. We found that the recruitment of TF326-NBD to nascent chains, as measured by the increase in NBD fluorescence, was dependent on the hydrophobicity of the nascent chain translated (Figure 14 C). The change in NBD fluorescence ($F-F_0/F_0$) during translation of Luc was ~ 0.7 , whereas no fluorescence change was observed during translation of α -Synuclein (α -Syn) (Figure 14 C). These experiments were complemented with the analysis of TF-B binding to ribosomes translating Luc. The binding resulted in the quenching of BADAN fluorescence, which saturated at around 1200 s after inhibition of translation (Figure 15 B), whereas the fluorescence of TF326-NBD saturated at around 2500 s (Figure 14 C). From these observations we conclude that TF initially interacts with the translating ribosomes and subsequently with the nascent chains.

It has been suggested before that multiple TF molecules may bind to the nascent chains depending on their size and hydrophobicity (Agashe et al., 2004). We studied the recruitment of TF326-NBD to the various lengths of ribosome-arrested Luc nascent

chains. A significant increase in NBD fluorescence during the translation of Luc 520-mer and Luc 550-mer (both exposing all the five predicted TF-binding motifs) was observed, when compared to Luc 164-mer exposing only one such hydrophobic motif (Figure 16 D). The increased recruitment of TF to nascent chains exposing an increasing number of hydrophobic motifs would be a mechanism to prevent their aggregation during translation.

VI.2 Complex kinetics of TF dissociation from Luc nascent chains

The dissociation kinetics of TF labeled with NBD from Luc RNCs was found to be dependent on the location of the probe. Both TF326-NBD and TF376-NBD showed biphasic dissociation kinetics from Luc nascent chains. For TF326-NBD, the fast phase occurred with a $t_{1/2}$ of $\sim 14 \pm 2$ s and the slow phase occurred with a $t_{1/2}$ of $\sim 102 \pm 16$ s (Figure 18 C). TF376-NBD also had similar fast and slow phases of dissociation, whereas TF150-NBD had only a single slow phase of dissociation from Luc nascent chains (Figure 18 C). Interestingly, the dissociation of TF326-NBD and TF376-NBD from Luc 164-mer exhibited only a fast phase similar to the dissociation of TF from the ribosome (Figure 19 B). We conclude that TF interacts with the shorter nascent chains while maintaining contact with the ribosome. TF's interaction with the ribosome contributes significantly to the overall affinity for shorter RNCs whereas its interaction with longer RNCs is governed mainly by the hydrophobicity of the nascent chain.

Notably, the net fluorescence increase of TF150-NBD (probe at the interface of PPIase domain and C-terminal domain) during the translation of FL Luc was ~ 3 -fold less than the increase in the fluorescence of TF326-NBD and TF376-NBD (Figure 17 A). This confirms that the PPIase domain acts only as a secondary nascent chains binding site, the primary site of contact for nascent chains being the N- and C-terminal domains.

VI.3 TFNC is more efficient than TF in assisting multi-domain protein folding

Next, we studied the correlation between the role of the PPIase domain of TF in nascent chain binding and the folding of multi-domain proteins. The PPIase domain of TF is not essential for the chaperoning function *in vivo* and TFNC (the PPIase domain deleted mutant of TF) can rescue the synthetic lethality of $\Delta dnaK \Delta tig$ *E. coli* similar to wild-type TF (Genevaux et al., 2004; Kramer et al., 2004b).

By displacement experiments with unlabeled TF, we found that TFNC has a shorter residence time on Luc nascent chains as compared to TF with $t_{1/2}$ values of 52 ± 9 s and 111 ± 7 s, respectively (Figure 21 C). We then studied the effect of TFNC on the folding of multi-domain proteins such as Luc and Ras-DHFR. These proteins were used as model substrate proteins since they fold via a post-translational folding mechanism in *E. coli* but a co-translational route in eukaryotes (Agashe et al., 2004; Netzer and Hartl, 1997). An improved folding efficiency was observed for both Luc and Ras-DHFR when these proteins were translated in presence of TFNC as compared to their translation in presence of TF (Figures 22, 23 and 25). This indicated that in the absence of the additional binding event provided by the PPIase domain, multi-domain proteins were folded more efficiently by TF. We conclude that since the PPIase domain acts as an additional nascent chain binding site, the deletion of this domain decreased the interaction strength with the nascent chains enhancing their folding.

VI.4 TFNC-mediated folding pathway of Luc and Ras-DHFR is more co-translational

To obtain insight into the mechanism by which TFNC mediated efficient folding, we analyzed the folding kinetics of Luc and Ras-DHFR in the presence of TF and TFNC. It has been shown previously that Luc and Ras-DHFR fold via an efficient co-translational mechanism in the eukaryotic system (Agashe et al., 2004; Netzer and Hartl, 1997). We found that TFNC imposed less delay on Luc folding relative to translation than TF (Figure 24). When RNaseA was added to a Luc translation reaction in the

presence of TFNC to stop translation, Luc activity stopped increasing immediately, although with TF it increased for another 15 min (Figure 36). These observations indicated that the TFNC-mediated Luc folding pathway is relatively more co-translational.

Similar results were obtained from the protease protection pattern of the Ras-DHFR fusion protein when the kinetics of its native structure formation was analyzed in presence of TF and TFNC. In the presence of TFNC, the protease resistant domains of Ras and DHFR appeared simultaneously with the full length protein, indicating that Ras-DHFR folded in a domain-wise folding mechanism during translation (Figure 28). Upon quantification of the folded DHFR domain in the Ras-DHFR fusion protein, we found that in the presence of TFNC the efficiency of folding was two-fold increased relative to the presence of TF (Figure 29).

We suggest that the PPIase domain binds to the hydrophobic regions on nascent chains (Patzelt et al., 2001) and delays the folding with respect to translation. This is evident from the post-translational folding of Luc in presence of TF (Agashe et al., 2004). Here we show that the deletion of this domain relieves the delay in protein folding and allows folding to proceed co-translationally. Interestingly, a ribosome-stalled construct of Ras domain was folded to a similar extent in the presence of TF or TFNC implying its specific function towards the multi-domain proteins (Figure 30).

We conclude that the PPIase domain of TF causes a delay in the folding of Luc and Ras-DHFR nascent chains upon binding which eventually leads to reduced yields of these proteins. Deletion of this domain proves to be beneficial for the folding of these multi-domain proteins. Taken together, the results discussed in the previous sections suggest that the PPIase domain of TF is responsible for the delay in folding caused by TF. A slower dissociation rate observed with TF confirms that the PPIase domain acts as an additional binding site for the nascent chains and hence prolongs the interaction between TF and nascent chains. TFNC, however, has a lower residence time on Luc nascent chains which enables the nascent chains to undergo more co-translational folding.

VI.5 The eukaryotic Hsp70 chaperone system mediates efficient folding of Luc

Molecular chaperones have extensively been utilized to improve the folding of recombinantly expressed proteins in *E. coli* (Baneyx, 1999; Baneyx and Mujacic, 2004; Haacke et al., 2009). However, the effects of these co-expression strategies have not been generally beneficial (Chang et al., 2005; Haacke et al., 2009). Eukaryotic proteins such as Luc fold more efficiently in a eukaryotic *in vitro* translation system (RRL) than in *E. coli* based on *in vitro* translation systems such as RTS and the PURE system (Figure 31 A) (Agashe et al., 2004) and this indicates that the chaperone machinery present in RRL is efficient in folding Luc. A recent report has also suggested a role of slower translation rate in the efficient folding of Luc and certain GFP fusion proteins (Siller et al., 2010). To gain insights into the roles of eukaryotic chaperones in efficient folding of Luc in RRL, we studied the folding of Luc in the PURE system in presence of different supplementations of RRL and cycloheximide (to specifically inhibit translation from eukaryotic ribosomes). No significant improvement in Luc folding was observed (Figure 31 B). The reason for this could be that the chaperones present in RRL may become inactive upon addition to the PURE system due to differences in the buffer components between these two translation systems. It could be also be possible that Luc is able to fold efficiently during slow translation from the eukaryotic ribosomes.

We therefore systematically supplemented the PURE system with members of the eukaryotic Hsp70 chaperone machinery (Hsc70, Hdj2 and Bag1) and observed a significant improvement in both specific activity and solubility of Luc (Figures 34 and 35). Interestingly, Luc folding was further enhanced by addition of TFNC (Figures 34 and 35). We then studied the folding pathway of Luc in the presence of these eukaryotic chaperones. Translation inhibition experiments demonstrated that the eukaryotic Hsp70 system mediated co-translational folding of Luc similar to the folding pathway mediated by TFNC alone (Figure 36). We conclude that the eukaryotic chaperones are capable of mediating efficient folding of Luc via a co-translational folding pathway which allows the formation of tertiary structure elements during translation and hence prevent non-native intra- and intermolecular contacts from forming that might lead to aggregation.

VI.6 Generation of a co-expression system to test the folding of different aggregation-prone multi-domain proteins in presence of various chaperones in *E. coli*

Having found that TFNC and the eukaryotic Hsp70 chaperone machinery can assist in proper folding of aggregation-prone Luc in the bacterial *in vitro* translation system, we attempted to generate an *in vivo* co-expression system in *E. coli* BL21 (DE3) where aggregation-prone multi-domain proteins could be expressed and folded with assistance from these chaperones.

However, when Luc was co-expressed with TFNC, no improvement in Luc specific activity could be observed (Figure 38). We suggest that the endogenous TF might be recruited to the ribosome with higher affinity than TFNC *in vivo* and hence TFNC is not able to mediate productive Luc folding. Co-expression of Hsc70 with Luc also did not lead to an increase in Luc folding, although Luc solubility was improved under these conditions (Figure 39). Hsc70 might be performing only a holdase function in the absence of its co-chaperones.

Co-expression of Hsc70, Hdj2, Bag1 and TFNC with Luc caused a drastic inhibition in Luc translation (Figure 40 A). Over-expression of the five different proteins together in the cell might cause a metabolic stress and hence severely affect the translation efficiency of the ribosomes. Another explanation of this observation could be an imbalance in the relative expression of Hsc70, Hdj2 and Bag1. Since the expression of these proteins could not be regulated accurately *in vivo*, it is possible that the imbalance in their relative amounts has a negative effect on the translation machinery of the cell. Indeed, when we omitted Hdj2 from this combination, the inhibition on Luc translation was relieved (Figure 40 B). This observation was again confirmed when we co-expressed Luc with Hdj2, where we observed a drastic inhibition of Luc translation and also of cell growth (Figures 41 B and 41 C).

Overall we conclude that generation of such a co-expression system to test the effect of various chaperone combinations on folding of aggregation-prone proteins

requires careful regulation of the expression levels of the individual components. Moreover, a TF knockout strain of *E. coli* should preferably be used to perform such experiments, in order to remove the effect of endogenous.

VI.7 Perspectives

In this thesis, we have studied the effect of the *E. coli* ribosome-associated chaperone TF and its PPIase domain-deleted mutant TFNC on the folding of eukaryotic multi-domain proteins such as Luc and Ras-DHFR. We found that TFNC increased the specific activity and solubility of Luc. The eukaryotic Hsp70 machinery was also tested for its ability to enhance the folding of Luc in a reconstituted *in vitro* translation system, the so-called PURE system. In the presence of Hsc70, Hdj2 and Bag1, a significant increase in Luc specific activity and solubility was observed. TFNC further improved this effect, indicating that TFNC acts as a more efficient chaperone at the ribosomal exit tunnel and mediates initial folding of the Luc nascent chains which are later transferred to the eukaryotic Hsp70 system comprised of Hsc70, Hdj2 and Bag1.

Although, we succeeded in improving the folding of Luc in a reconstituted *in vitro* translation system by using modified *E. coli* chaperones and eukaryotic chaperones, attempts to improve Luc folding *in vivo* by co-expression of these chaperones did not yield clear results. The main reason for this is probably the imbalance in the relative expression levels of Hsc70, Hdj2 and Bag1, as it is clear from the *in vitro* experiments that only a defined ratio of these proteins leads to improved Luc folding. In the future, attempts will be made to carefully regulate the expression levels of these proteins. Also, such experiments will be performed in the absence of endogenous TF present in the BL21 (DE3) cells so that the endogenous TF does not cause any negative effect on the folding of Luc.

VII REFERENCES

Agashe, V.R., Guha, S., Chang, H.C., Genevoux, P., Hayer-Hartl, M., Stemp, M., Georgopoulos, C., Hartl, F.U., and Barral, J.M. (2004). Function of trigger factor and DnaK in multidomain protein folding: increase in yield at the expense of folding speed. *Cell* *117*, 199-209.

Agashe, V.R., Shastry, M.C., and Udgaonkar, J.B. (1995). Initial hydrophobic collapse in the folding of barstar. *Nature* *377*, 754-757.

Albanese, V., Reissmann, S., and Frydman, J. (2010). A ribosome-anchored chaperone network that facilitates eukaryotic ribosome biogenesis. *Journal of Cell Biology* *189*, 69-U105.

Anfinsen, C.B. (1972). The formation and stabilization of protein structure. *The Biochemical journal* *128*, 737-749.

Anfinsen, C.B. (1973). Principles that govern the folding of protein chains. *Science (New York, NY)* *181*, 223-230.

Aponte, R.A., Zimmermann, S., and Reinstein, J. (2010) Directed evolution of the DnaK chaperone: mutations in the lid domain result in enhanced chaperone activity. *Journal of Molecular Biology* *399*, 154-167.

Baldwin, R.L. (1989). How does protein folding get started? *Trends in biochemical sciences* *14*, 291-294.

Ballinger, C.A., Connell, P., Wu, Y., Hu, Z., Thompson, L.J., Yin, L.Y., and Patterson, C. (1999). Identification of CHIP, a novel tetratricopeptide repeat-containing protein that interacts with heat shock proteins and negatively regulates chaperone functions. *Mol Cell Biol* *19*, 4535-4545.

Baneyx, F. (1999). Recombinant protein expression in *Escherichia coli*. *Curr Opin Biotechnol* *10*, 411-421.

Baneyx, F., and Mujacic, M. (2004). Recombinant protein folding and misfolding in *Escherichia coli*. *Nat Biotechnol* *22*, 1399-1408.

Baram, D., Pyetan, E., Sittner, A., Auerbach-Nevo, T., Bashan, A., and Yonath, A. (2005). Structure of trigger factor binding domain in biologically homologous complex with eubacterial ribosome reveals its chaperone action. *Proceedings of the National Academy of Sciences of the United States of America* *102*, 12017-12022.

Barral, J.M., Broadley, S.A., Schaffar, G., and Hartl, F.U. (2004). Roles of molecular chaperones in protein misfolding diseases. *Seminars in cell & developmental biology* *15*, 17-29.

- Beck, K., Wu, L.F., Brunner, J., and Muller, M. (2000). Discrimination between SRP- and SecA/SecB-dependent substrates involves selective recognition of nascent chains by SRP and trigger factor. *The EMBO journal* *19*, 134-143.
- Bingel-Erlenmeyer, R., Kohler, R., Kramer, G., Sandikci, A., Antolic, S., Maier, T., Schaffitzel, C., Wiedmann, B., Bukau, B., and Ban, N. (2008). A peptide deformylase-ribosome complex reveals mechanism of nascent chain processing. *Nature* *452*, 108-111.
- Bradford, M.M. (1976). A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Analytical biochemistry* *72*, 248-254.
- Brinker, A., Pfeifer, G., Kerner, M.J., Naylor, D.J., Hartl, F.U., and Hayer-Hartl, M. (2001). Dual function of protein confinement in chaperonin-assisted protein folding. *Cell* *107*, 223-233.
- Buskiewicz, I., Deuerling, E., Gu, S.Q., Jockel, J., Rodnina, M.V., Bukau, B., and Wintermeyer, W. (2004). Trigger factor binds to ribosome-signal-recognition particle (SRP) complexes and is excluded by binding of the SRP receptor. *Proceedings of the National Academy of Sciences of the United States of America* *101*, 7902-7906.
- Chang, H.C., Kaiser, C.M., Hartl, F.U., and Barral, J.M. (2005). De novo folding of GFP fusion proteins: high efficiency in eukaryotes but not in bacteria. *Journal of molecular biology* *353*, 397-409.
- Clark, P.L. (2004). Protein folding in the cell: reshaping the folding funnel. *Trends in biochemical sciences* *29*, 527-534.
- Connell, P., Ballinger, C.A., Jiang, J., Wu, Y., Thompson, L.J., Hohfeld, J., and Patterson, C. (2001). The co-chaperone CHIP regulates protein triage decisions mediated by heat-shock proteins. *Nat Cell Biol* *3*, 93-96.
- Crooke, E., Brundage, L., Rice, M., and Wickner, W. (1988a). ProOmpA spontaneously folds in a membrane assembly competent state which trigger factor stabilizes. *The EMBO journal* *7*, 1831-1835.
- Crooke, E., Guthrie, B., Lecker, S., Lill, R., and Wickner, W. (1988b). ProOmpA is stabilized for membrane translocation by either purified E. coli trigger factor or canine signal recognition particle. *Cell* *54*, 1003-1011.
- Crooke, E., and Wickner, W. (1987). Trigger factor: a soluble protein that folds pro-OmpA into a membrane-assembly-competent form. *Proceedings of the National Academy of Sciences of the United States of America* *84*, 5216-5220.
- Datsenko, K.A., and Wanner, B.L. (2000). One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proceedings of the National Academy of Sciences of the United States of America* *97*, 6640-6645.

- de Marco, A., Deuerling, E., Mogk, A., Tomoyasu, T., and Bukau, B. (2007). Chaperone-based procedure to increase yields of soluble recombinant proteins produced in *E. coli*. *BMC biotechnology* 7, 32.
- Deuerling, E., Schulze-Specking, A., Tomoyasu, T., Mogk, A., and Bukau, B. (1999). Trigger factor and DnaK cooperate in folding of newly synthesized proteins. *Nature* 400, 693-696.
- Dower, W.J., Miller, J.F., and Ragsdale, C.W. (1988). High efficiency transformation of *E. coli* by high voltage electroporation. *Nucleic Acids Res* 16, 6127-6145.
- Dragovic, Z., Broadley, S.A., Shomura, Y., Bracher, A., and Hartl, F.U. (2006). Molecular chaperones of the Hsp110 family act as nucleotide exchange factors of Hsp70s. *The EMBO journal* 25, 2519-2528.
- Eisner, G., Moser, M., Schafer, U., Beck, K., and Muller, M. (2006). Alternate recruitment of signal recognition particle and trigger factor to the signal sequence of a growing nascent polypeptide. *The Journal of biological chemistry* 281, 7172-7179.
- Ellis, R.J., and Minton, A.P. (2006). Protein aggregation in crowded environments. *Biological chemistry* 387, 485-497.
- Ferbitz, L., Maier, T., Patzelt, H., Bukau, B., Deuerling, E., and Ban, N. (2004). Trigger factor in complex with the ribosome forms a molecular cradle for nascent proteins. *Nature* 431, 590-596.
- Fersht, A.R. (2008). From the first protein structures to our current knowledge of protein folding: delights and scepticisms. *Nature reviews* 9, 650-654.
- Fiaux, J., Horst, J., Scior, A., Preissler, S., Koplín, A., Bukau, B., and Deuerling, E. (2010). Structural Analysis of the Ribosome-associated Complex (RAC) Reveals an Unusual Hsp70/Hsp40 Interaction. *Journal of Biological Chemistry* 285, 3227-3234.
- Flaherty, K.M., DeLuca-Flaherty, C., and McKay, D.B. (1990). Three-dimensional structure of the ATPase fragment of a 70K heat-shock cognate protein. *Nature* 346, 623-628.
- Frydman, J., Erdjument-Bromage, H., Tempst, P., and Hartl, F.U. (1999). Co-translational domain folding as the structural basis for the rapid de novo folding of firefly luciferase. *Nature structural biology* 6, 697-705.
- Frydman, J., Nimmegern, E., Erdjument-Bromage, H., Wall, J.S., Tempst, P., and Hartl, F.U. (1992). Function in protein folding of TRiC, a cytosolic ring complex containing TCP-1 and structurally related subunits. *The EMBO journal* 11, 4767-4778.
- Frydman, J., Nimmegern, E., Ohtsuka, K., and Hartl, F.U. (1994). Folding of nascent polypeptide chains in a high molecular mass assembly with molecular chaperones. *Nature* 370, 111-117.

- Gao, Y., Thomas, J.O., Chow, R.L., Lee, G.H., and Cowan, N.J. (1992). A cytoplasmic chaperonin that catalyzes beta-actin folding. *Cell* 69, 1043-1050.
- Gautschi, M., Lilie, H., Funfschilling, U., Mun, A., Ross, S., Lithgow, T., Rucknagel, P., and Rospert, S. (2001). RAC, a stable ribosome-associated complex in yeast formed by the DnaK-DnaJ homologs Ssz1p and zuotin. *Proceedings of the National Academy of Sciences of the United States of America* 98, 3762-3767.
- Gelinas, A.D., Toth, J., Bethoney, K.A., Langsetmo, K., Stafford, W.F., and Harrison, C.J. (2003). Thermodynamic linkage in the GrpE nucleotide exchange factor, a molecular thermosensor. *Biochemistry* 42, 9050-9059.
- Genevaux, P., Keppel, F., Schwager, F., Langendijk-Genevaux, P.S., Hartl, F.U., and Georgopoulos, C. (2004). In vivo analysis of the overlapping functions of DnaK and trigger factor. *EMBO reports* 5, 195-200.
- Gross, M., and Hessefort, S. (1996). Purification and characterization of a 66-kDa protein from rabbit reticulocyte lysate which promotes the recycling of hsp 70. *The Journal of biological chemistry* 271, 16833-16841.
- Guthrie, B., and Wickner, W. (1990). Trigger factor depletion or overproduction causes defective cell division but does not block protein export. *Journal of bacteriology* 172, 5555-5562.
- Haacke, A., Fendrich, G., Ramage, P., and Geiser, M. (2009). Chaperone over-expression in *Escherichia coli*: apparent increased yields of soluble recombinant protein kinases are due mainly to soluble aggregates. *Protein Expr Purif* 64, 185-193.
- Halic, M., Blau, M., Becker, T., Mielke, T., Pool, M.R., Wild, K., Sinning, I., and Beckmann, R. (2006). Following the signal sequence from ribosomal tunnel exit to signal recognition particle. *Nature* 444, 507-511.
- Harrison, C.J., Hayer-Hartl, M., Di Liberto, M., Hartl, F., and Kuriyan, J. (1997). Crystal structure of the nucleotide exchange factor GrpE bound to the ATPase domain of the molecular chaperone DnaK. *Science (New York, NY)* 276, 431-435.
- Hartl, F.U. (1996). Molecular chaperones in cellular protein folding. *Nature* 381, 571-579.
- Hartl, F.U., and Hayer-Hartl, M. (2002). Molecular chaperones in the cytosol: from nascent chain to folded protein. *Science (New York, NY)* 295, 1852-1858.
- Hartl, F.U., and Hayer-Hartl, M. (2009). Converging concepts of protein folding in vitro and in vivo. *Nature structural & molecular biology* 16, 574-581.
- Hartl, F.U., Lecker, S., Schiebel, E., Hendrick, J.P., and Wickner, W. (1990). The binding cascade of SecB to SecA to SecY/E mediates preprotein targeting to the *E. coli* plasma membrane. *Cell* 63, 269-279.

- Ho, Y., Gruhler, A., Heilbut, A., Bader, G.D., Moore, L., Adams, S.L., Millar, A., Taylor, P., Bennett, K., Boutilier, K., *et al.* (2002). Systematic identification of protein complexes in *Saccharomyces cerevisiae* by mass spectrometry. *Nature* *415*, 180-183.
- Hoffmann, A., Bukau, B., and Kramer G. (2010) Structure and function of the molecular chaperone Trigger Factor. *Biochim Biophys Acta* *1803*, 650-661.
- Hohfeld, J., and Jentsch, S. (1997). GrpE-like regulation of the hsc70 chaperone by the anti-apoptotic protein BAG-1. *The EMBO journal* *16*, 6209-6216.
- Hohfeld, J., Minami, Y., and Hartl, F.U. (1995). Hip, a novel cochaperone involved in the eukaryotic Hsc70/Hsp40 reaction cycle. *Cell* *83*, 589-598.
- Horwich, A.L., Fenton, W.A., Chapman, E., and Farr, G.W. (2007). Two families of chaperonin: physiology and mechanism. *Annu Rev Cell Dev Biol* *23*, 115-145.
- Huang, P., Gautschi, M., Walter, W., Rospert, S., and Craig, E.A. (2005). The Hsp70 Ssz1 modulates the function of the ribosome-associated J-protein Zuo1. *Nature structural & molecular biology* *12*, 497-504.
- Hundley, H.A., Walter, W., Bairstow, S., and Craig, E.A. (2005). Human Mpp11 J protein: ribosome-tethered molecular chaperones are ubiquitous. *Science (New York, NY)* *308*, 1032-1034.
- Inouye, S., and Sahara, Y. (2008). Soluble protein expression in *E. coli* cells using IgG-binding domain of protein A as a solubilizing partner in the cold induced system. *Biochemical and biophysical research communications* *376*, 448-453.
- Jackson, S.E., and Fersht, A.R. (1991). Folding of chymotrypsin inhibitor 2. 1. Evidence for a two-state transition. *Biochemistry* *30*, 10428-10435.
- Kaiser, C.M., Chang, H.C., Agashe, V.R., Lakshmipathy, S.K., Etchells, S.A., Hayer-Hartl, M., Hartl, F.U., and Barral, J.M. (2006). Real-time observation of trigger factor function on translating ribosomes. *Nature* *444*, 455-460.
- Kim, P.S., and Baldwin, R.L. (1982). Specific intermediates in the folding reactions of small proteins and the mechanism of protein folding. *Annual review of biochemistry* *51*, 459-489.
- Kim, P.S., and Baldwin, R.L. (1990). Intermediates in the folding reactions of small proteins. *Annual review of biochemistry* *59*, 631-660.
- Koplin, A., Preissler, S., Ilina, Y., Koch, M., Scior, A., Erhardt, M., and Deuerling, E. (2010). A dual function for chaperones SSB-RAC and the NAC nascent polypeptide-associated complex on ribosomes. *Journal of Cell Biology* *189*, 57-U86.
- Kramer, G., Boehringer, D., Ban, N., and Bukau, B. (2009). The ribosome as a platform for co-translational processing, folding and targeting of newly synthesized proteins. *Nature structural & molecular biology* *16*, 589-597.

- Kramer, G., Patzelt, H., Rauch, T., Kurz, T.A., Vorderwulbecke, S., Bukau, B., and Deuerling, E. (2004a). Trigger factor peptidyl-prolyl cis/trans isomerase activity is not essential for the folding of cytosolic proteins in *Escherichia coli*. *The Journal of biological chemistry* 279, 14165-14170.
- Kramer, G., Rauch, T., Rist, W., Vorderwulbecke, S., Patzelt, H., Schulze-Specking, A., Ban, N., Deuerling, E., and Bukau, B. (2002). L23 protein functions as a chaperone docking site on the ribosome. *Nature* 419, 171-174.
- Kramer, G., Rutkowska, A., Wegrzyn, R.D., Patzelt, H., Kurz, T.A., Merz, F., Rauch, T., Vorderwulbecke, S., Deuerling, E., and Bukau, B. (2004b). Functional dissection of *Escherichia coli* trigger factor: unraveling the function of individual domains. *Journal of bacteriology* 186, 3777-3784.
- Kristensen, O., and Gajhede, M. (2003). Chaperone binding at the ribosomal exit tunnel. *Structure* 11, 1547-1556.
- Kyratsous, C.A., Silverstein, S.J., DeLong, C.R., and Panagiotidis, C.A. (2009). Chaperone-fusion expression plasmid vectors for improved solubility of recombinant proteins in *Escherichia coli*. *Gene* 440, 9-15.
- Laemmli, U.K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* 227, 680-685.
- Lakshmipathy, S.K., Tomic, S., Kaiser, C.M., Chang, H.C., Genevoux, P., Georgopoulos, C., Barral, J.M., Johnson, A.E., Hartl, F.U., and Etchells, S.A. (2007). Identification of nascent chain interaction sites on trigger factor. *J Biol Chem* 282, 12186-12193.
- Langer, T., Lu, C., Echols, H., Flanagan, J., Hayer, M.K., and Hartl, F.U. (1992a). Successive action of DnaK, DnaJ and GroEL along the pathway of chaperone-mediated protein folding. *Nature* 356, 683-689.
- Langer, T., Pfeifer, G., Martin, J., Baumeister, W., and Hartl, F.U. (1992b). Chaperonin-mediated protein folding: GroES binds to one end of the GroEL cylinder, which accommodates the protein substrate within its central cavity. *The EMBO journal* 11, 4757-4765.
- Lauring, B., Kreibich, G., and Weidmann, M. (1995). The intrinsic ability of ribosomes to bind to endoplasmic reticulum membranes is regulated by signal recognition particle and nascent-polypeptide-associated complex. *Proceedings of the National Academy of Sciences of the United States of America* 92, 9435-9439.
- Lee, H.C., and Bernstein, H.D. (2002). Trigger factor retards protein export in *Escherichia coli*. *The Journal of biological chemistry* 277, 43527-43535.
- Levinthal, C., Signer, E.R., and Fetherolf, K. (1962). Reactivation and hybridization of reduced alkaline phosphatase. *Proceedings of the National Academy of Sciences of the United States of America* 48, 1230-1237.

- Lill, R., Crooke, E., Guthrie, B., and Wickner, W. (1988). The "trigger factor cycle" includes ribosomes, presecretory proteins, and the plasma membrane. *Cell* 54, 1013-1018.
- Lin, S., and Struve, W.S. (1991). Time-resolved fluorescence of nitrobenzoxadiazole-aminohexanoic acid: effect of intermolecular hydrogen-bonding on non-radiative decay. *Photochem Photobiol* 54, 361-365.
- Ludlam, A.V., Moore, B.A., and Xu, Z. (2004). The crystal structure of ribosomal chaperone trigger factor from *Vibrio cholerae*. *Proceedings of the National Academy of Sciences of the United States of America* 101, 13436-13441.
- Macfarlane, J., and Muller, M. (1995). The functional integration of a polytopic membrane protein of *Escherichia coli* is dependent on the bacterial signal-recognition particle. *European journal of biochemistry / FEBS* 233, 766-771.
- Maier, R., Eckert, B., Scholz, C., Lilie, H., and Schmid, F.X. (2003). Interaction of trigger factor with the ribosome. *J Mol Biol* 326, 585-592.
- Malkin, L.I., and Rich, A. (1967). Partial resistance of nascent polypeptide chains to proteolytic digestion due to ribosomal shielding. *J Mol Biol* 26, 329-346.
- Martinez-Hackert, E., and Hendrickson, W.A. (2009). Promiscuous substrate recognition in folding and assembly activities of the trigger factor chaperone. *Cell* 138, 923-934.
- Mayer, M.P., and Bukau, B. (2005). Hsp70 chaperones: cellular functions and molecular mechanism. *Cell Mol Life Sci* 62, 670-684.
- Mayhew, M., da Silva, A.C., Martin, J., Erdjument-Bromage, H., Tempst, P., and Hartl, F.U. (1996). Protein folding in the central cavity of the GroEL-GroES chaperonin complex. *Nature* 379, 420-426.
- Merz, F., Boehringer, D., Schaffitzel, C., Preissler, S., Hoffmann, A., Maier, T., Rutkowska, A., Lozza, J., Ban, N., Bukau, B., *et al.* (2008). Molecular mechanism and structure of Trigger Factor bound to the translating ribosome. *Embo J* 27, 1622-1632.
- Minami, Y., Hohfeld, J., Ohtsuka, K., and Hartl, F.U. (1996). Regulation of the heat-shock protein 70 reaction cycle by the mammalian DnaJ homolog, Hsp40. *The Journal of biological chemistry* 271, 19617-19624.
- Mogk, A., Deuerling, E., Vorderwulbecke, S., Vierling, E., and Bukau, B. (2003). Small heat shock proteins, ClpB and the DnaK system form a functional triade in reversing protein aggregation. *Molecular microbiology* 50, 585-595.
- Netzer, W.J., and Hartl, F.U. (1997). Recombination of protein domains facilitated by co-translational folding in eukaryotes. *Nature* 388, 343-349.
- Onuchic, J.N. (1997). Contacting the protein folding funnel with NMR. *Proceedings of the National Academy of Sciences of the United States of America* 94, 7129-7131.

Otto, H., Conz, C., Maier, P., Wolfle, T., Suzuki, C.K., Jenö, P., Rucknagel, P., Stahl, J., and Rospert, S. (2005). The chaperones MPP11 and Hsp70L1 form the mammalian ribosome-associated complex. *Proceedings of the National Academy of Sciences of the United States of America* 102, 10064-10069.

Patzelt, H., Kramer, G., Rauch, T., Schonfeld, H.J., Bukau, B., and Deuerling, E. (2002). Three-state equilibrium of *Escherichia coli* trigger factor. *Biological chemistry* 383, 1611-1619.

Patzelt, H., Rudiger, S., Brehmer, D., Kramer, G., Vorderwulbecke, S., Schaffitzel, E., Waitz, A., Hesterkamp, T., Dong, L., Schneider-Mergener, J., *et al.* (2001). Binding specificity of *Escherichia coli* trigger factor. *Proceedings of the National Academy of Sciences of the United States of America* 98, 14244-14249.

Pauling, L., and Corey, R.B. (1951a). Atomic coordinates and structure factors for two helical configurations of polypeptide chains. *Proceedings of the National Academy of Sciences of the United States of America* 37, 235-240.

Pauling, L., and Corey, R.B. (1951b). The pleated sheet, a new layer configuration of polypeptide chains. *Proceedings of the National Academy of Sciences of the United States of America* 37, 251-256.

Pech, M., Spreter, T., Beckmann, R., and Beatrix, B. (2010). Dual binding mode of the nascent polypeptide-associated complex reveals a novel universal adapter site on the ribosome. *The Journal of biological chemistry* 285, 19679-19687.

Peisker, K., Braun, D., Wolfle, T., Hentschel, J., Funfschilling, U., Fischer, G., Sickmann, A., and Rospert, S. (2008). Ribosome-associated complex binds to ribosomes in close proximity of Rpl31 at the exit of the polypeptide tunnel in yeast. *Molecular biology of the cell* 19, 5279-5288.

Perutz, M.F., Rossmann, M.G., Cullis, A.F., Muirhead, H., Will, G., and North, A.C. (1960). Structure of haemoglobin: a three-dimensional Fourier synthesis at 5.5-Å resolution, obtained by X-ray analysis. *Nature* 185, 416-422.

Pierpaoli, E.V., Sandmeier, E., Baici, A., Schonfeld, H.J., Gisler, S., and Christen, P. (1997). The power stroke of the DnaK/DnaJ/GrpE molecular chaperone system. *Journal of molecular biology* 269, 757-768.

Poritz, M.A., Bernstein, H.D., Strub, K., Zopf, D., Wilhelm, H., and Walter, P. (1990). An *E. coli* ribonucleoprotein containing 4.5S RNA resembles mammalian signal recognition particle. *Science (New York, NY)* 250, 1111-1117.

Ptitsyn, O.B., and Rashin, A.A. (1975). A model of myoglobin self-organization. *Biophys Chem* 3, 1-20.

Raine, A., Ivanova, N., Wikberg, J.E., and Ehrenberg, M. (2004). Simultaneous binding of trigger factor and signal recognition particle to the *E. coli* ribosome. *Biochimie* 86, 495-500.

- Ramachandran, G.N., and Sasisekharan, V. (1968). Conformation of polypeptides and proteins. *Advances in protein chemistry* 23, 283-438.
- Ranson, N.A., Farr, G.W., Roseman, A.M., Gowen, B., Fenton, W.A., Horwich, A.L., and Saibil, H.R. (2001). ATP-bound states of GroEL captured by cryo-electron microscopy. *Cell* 107, 869-879.
- Rudiger, S., Germeroth, L., Schneider-Mergener, J., and Bukau, B. (1997). Substrate specificity of the DnaK chaperone determined by screening cellulose-bound peptide libraries. *The EMBO journal* 16, 1501-1507.
- Rutkowska, A., Mayer, M.P., Hoffmann, A., Merz, F., Zachmann-Brand, B., Schaffitzel, C., Ban, N., Deuerling, E., and Bukau, B. (2007). Dynamics of trigger factor interaction with translating ribosomes. *J Biol Chem*.
- Rye, H.S., Burston, S.G., Fenton, W.A., Beechem, J.M., Xu, Z., Sigler, P.B., and Horwich, A.L. (1997). Distinct actions of cis and trans ATP within the double ring of the chaperonin GroEL. *Nature* 388, 792-798.
- Ryu, K., Kim, C.W., Kim, B.H., Han, K.S., Kim, K.H., Choi, S.I., and Seong, B.L. (2008). Assessment of substrate-stabilizing factors for DnaK on the folding of aggregation-prone proteins. *Biochemical and biophysical research communications* 373, 74-79.
- Schaffitzel, C., Oswald, M., Berger, I., Ishikawa, T., Abrahams, J.P., Koerten, H.K., Koning, R.I., and Ban, N. (2006). Structure of the E. coli signal recognition particle bound to a translating ribosome. *Nature* 444, 503-506.
- Schlunzen, F., Wilson, D.N., Tian, P., Harms, J.M., McInnes, S.J., Hansen, H.A., Albrecht, R., Buerger, J., Wilbanks, S.M., and Fucini, P. (2005). The binding mode of the trigger factor on the ribosome: implications for protein folding and SRP interaction. *Structure* 13, 1685-1694.
- Scholz, C., Stoller, G., Zarnt, T., Fischer, G., and Schmid, F.X. (1997). Cooperation of enzymatic and chaperone functions of trigger factor in the catalysis of protein folding. *The EMBO journal* 16, 54-58.
- Seluanov, A., and Bibi, E. (1997). FtsY, the prokaryotic signal recognition particle receptor homologue, is essential for biogenesis of membrane proteins. *The Journal of biological chemistry* 272, 2053-2055.
- Shimizu, Y., Inoue, A., Tomari, Y., Suzuki, T., Yokogawa, T., Nishikawa, K., and Ueda, T. (2001). Cell-free translation reconstituted with purified components. *Nature biotechnology* 19, 751-755.
- Shimizu, Y., Kanamori, T., and Ueda, T. (2005). Protein synthesis by pure translation systems. *Methods (San Diego, Calif)* 36, 299-304.

- Shomura, Y., Dragovic, Z., Chang, H.C., Tzvetkov, N., Young, J.C., Brodsky, J.L., Guerriero, V., Hartl, F.U., and Bracher, A. (2005). Regulation of Hsp70 function by HspBP1: structural analysis reveals an alternate mechanism for Hsp70 nucleotide exchange. *Molecular cell* *17*, 367-379.
- Siller, E., DeZwaan, D.C., Anderson, J.F., Freeman, B.C., and Barral, J.M. (2010). Slowing Bacterial Translation Speed Enhances Eukaryotic Protein Folding Efficiency. *Journal of molecular biology* *396*, 1310-1318.
- Smith, D.F., Sullivan, W.P., Marion, T.N., Zaitso, K., Madden, B., McCormick, D.J., and Toft, D.O. (1993). Identification of a 60-kilodalton stress-related protein, p60, which interacts with hsp90 and hsp70. *Mol Cell Biol* *13*, 869-876.
- Stemp, M.J., Guha, S., Hartl, F.U., and Barral, J.M. (2005). Efficient production of native actin upon translation in a bacterial lysate supplemented with the eukaryotic chaperonin TRiC. *Biological chemistry* *386*, 753-757.
- Sternlicht, H., Farr, G.W., Sternlicht, M.L., Driscoll, J.K., Willison, K., and Yaffe, M.B. (1993). The t-complex polypeptide 1 complex is a chaperonin for tubulin and actin in vivo. *Proceedings of the National Academy of Sciences of the United States of America* *90*, 9422-9426.
- Stoller, G., Rucknagel, K.P., Nierhaus, K.H., Schmid, F.X., Fischer, G., and Rahfeld, J.U. (1995). A ribosome-associated peptidyl-prolyl cis/trans isomerase identified as the trigger factor. *The EMBO journal* *14*, 4939-4948.
- Suh, W.C., Lu, C.Z., and Gross, C.A. (1999). Structural features required for the interaction of the Hsp70 molecular chaperone DnaK with its cochaperone DnaJ. *The Journal of biological chemistry* *274*, 30534-30539.
- Takayama, S., Bimston, D.N., Matsuzawa, S., Freeman, B.C., Aime-Sempe, C., Xie, Z., Morimoto, R.I., and Reed, J.C. (1997). BAG-1 modulates the chaperone activity of Hsp70/Hsc70. *The EMBO journal* *16*, 4887-4896.
- Taniuchi, H., and Anfinsen, C.B. (1969). An experimental approach to the study of the folding of staphylococcal nuclease. *The Journal of biological chemistry* *244*, 3864-3875.
- Terada, K., Kanazawa, M., Bukau, B., and Mori, M. (1997). The human DnaJ homologue dj2 facilitates mitochondrial protein import and luciferase refolding. *The Journal of cell biology* *139*, 1089-1095.
- Teter, S.A., Houry, W.A., Ang, D., Tradler, T., Rockabrand, D., Fischer, G., Blum, P., Georgopoulos, C., and Hartl, F.U. (1999). Polypeptide flux through bacterial Hsp70: DnaK cooperates with trigger factor in chaperoning nascent chains. *Cell* *97*, 755-765.
- Theysen, H., Schuster, H.P., Packschies, L., Bukau, B., and Reinstein, J. (1996). The second step of ATP binding to DnaK induces peptide release. *Journal of molecular biology* *263*, 657-670.

- Thulasiraman, V., Yang, C.F., and Frydman, J. (1999). In vivo newly translated polypeptides are sequestered in a protected folding environment. *The EMBO journal* *18*, 85-95.
- Tomic, S., Johnson, A.E., Hartl, F.U., and Etchells, S.A. (2006). Exploring the capacity of trigger factor to function as a shield for ribosome bound polypeptide chains. *FEBS letters* *580*, 72-76.
- Tzankov, S., Wong, M.J., Shi, K., Nassif, C., and Young, J.C. (2008). Functional divergence between co-chaperones of Hsc70. *The Journal of biological chemistry* *283*, 27100-27109.
- Ullers, R.S., Houben, E.N., Brunner, J., Oudega, B., Harms, N., and Luirink, J. (2006). Sequence-specific interactions of nascent *Escherichia coli* polypeptides with trigger factor and signal recognition particle. *The Journal of biological chemistry* *281*, 13999-14005.
- Ullers, R.S., Houben, E.N., Raine, A., ten Hagen-Jongman, C.M., Ehrenberg, M., Brunner, J., Oudega, B., Harms, N., and Luirink, J. (2003). Interplay of signal recognition particle and trigger factor at L23 near the nascent chain exit site on the *Escherichia coli* ribosome. *The Journal of cell biology* *161*, 679-684.
- Valent, Q.A., de Gier, J.W., von Heijne, G., Kendall, D.A., ten Hagen-Jongman, C.M., Oudega, B., and Luirink, J. (1997). Nascent membrane and presecretory proteins synthesized in *Escherichia coli* associate with signal recognition particle and trigger factor. *Molecular microbiology* *25*, 53-64.
- Vos, M.J., Hageman, J., Carra, S., and Kampinga, H.H. (2008). Structural and functional diversities between members of the human HSPB, HSPH, HSPA, and DNAJ chaperone families. *Biochemistry* *47*, 7001-7011.
- Wang, S., Sakai, H., and Wiedmann, M. (1995). NAC covers ribosome-associated nascent chains thereby forming a protective environment for regions of nascent chains just emerging from the peptidyl transferase center. *The Journal of cell biology* *130*, 519-528.
- Wegrzyn, R.D., Hofmann, D., Merz, F., Nikolay, R., Rauch, T., Graf, C., and Deuerling, E. (2006). A conserved motif is prerequisite for the interaction of NAC with ribosomal protein L23 and nascent chains. *The Journal of biological chemistry* *281*, 2847-2857.
- Weissman, J.S., Rye, H.S., Fenton, W.A., Beechem, J.M., and Horwich, A.L. (1996). Characterization of the active intermediate of a GroEL-GroES-mediated protein folding reaction. *Cell* *84*, 481-490.
- Wiedmann, B., Sakai, H., Davis, T.A., and Wiedmann, M. (1994). A protein complex required for signal-sequence-specific sorting and translocation. *Nature* *370*, 434-440.
- Xu, Z., Horwich, A.L., and Sigler, P.B. (1997). The crystal structure of the asymmetric GroEL-GroES-(ADP)₇ chaperonin complex. *Nature* *388*, 741-750.

- Yaffe, M.B., Farr, G.W., Miklos, D., Horwich, A.L., Sternlicht, M.L., and Sternlicht, H. (1992). TCP1 complex is a molecular chaperone in tubulin biogenesis. *Nature* 358, 245-248.
- Young, J.C., Agashe, V.R., Siegers, K., and Hartl, F.U. (2004). Pathways of chaperone-mediated protein folding in the cytosol. *Nature reviews* 5, 781-791.
- Zhu, X., Zhao, X., Burkholder, W.F., Gragerov, A., Ogata, C.M., Gottesman, M.E., and Hendrickson, W.A. (1996). Structural analysis of substrate binding by the molecular chaperone DnaK. *Science (New York, NY)* 272, 1606-1614.
- Zwanzig, R., Szabo, A., and Bagchi, B. (1992). Levinthal's paradox. *Proceedings of the National Academy of Sciences of the United States of America* 89, 20-22.

VIII APPENDICES

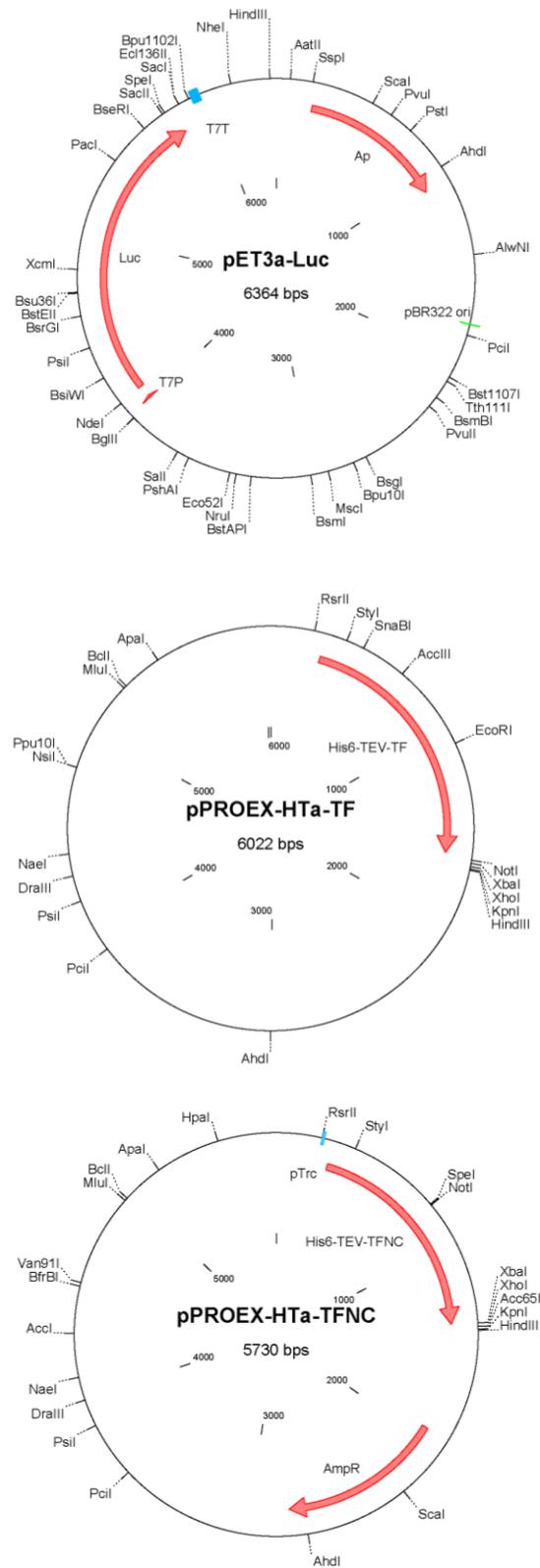
VIII.1 Abbreviations

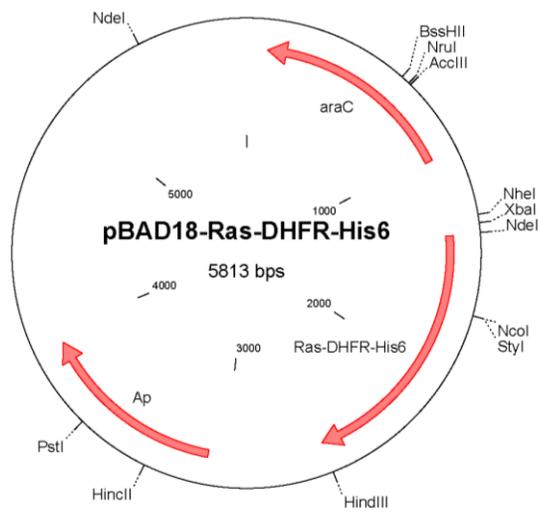
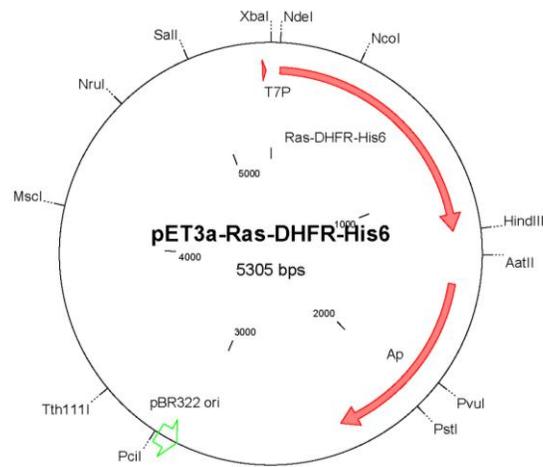
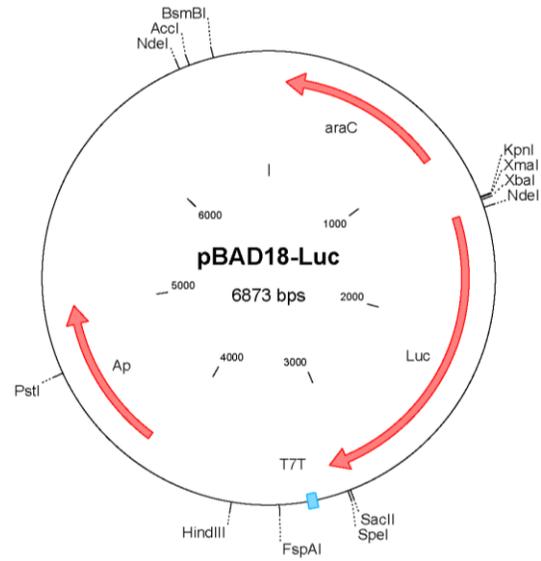
aa	Amino acid
ADP	Adenosine 5-diphosphate
Amp	Ampicillin
APS	Ammonium peroxodisulfate
ATP	Adenosine 5'-triphosphate
BSA	Bovine serum albumin
CAM	Chloramphenicol
CDTA	<i>trans</i> 1,2-diaminocyclohexane- <i>N,N,N',N'</i> -tetraacetic acid
CIAP	Calf Intestinal Alkaline Phosphatase
DHFR	Dihydrofolate reductase
DNA	Deoxyribonucleic acid
DTT	dithiothreitol
<i>E. coli</i>	<i>Escherichia coli</i>
EDTA	Ethylenediaminetetraacetic acid
FKBP	FK506 Binding Protein
FRET	fluorescence resonance energy transfer
g	acceleration of gravity, 9.81 m/s ²
GFP	Green fluorescent protein
h	Hour
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HRP	Horseradish peroxidase
Hsp	Heat shock protein
IANBD	N-((2-(iodoacetoxy)ethyl)-N-methyl)amino-7-nitrobenz-2-oxa-1,3-diazole
IPTG	Isopropyl- β -D-1-thiogalactopyranoside

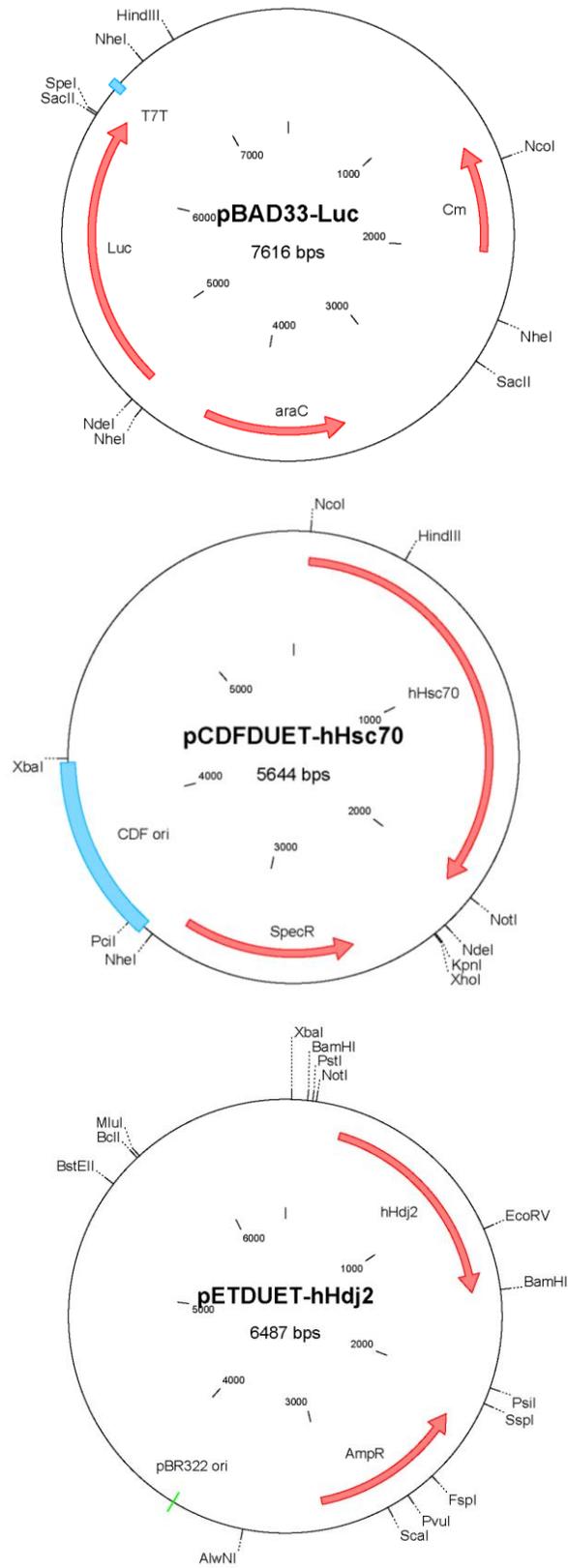
Kan	Kanamycin
LB	Luria Bertani
Luc	Firefly luciferase
MBP	Maltose binding protein
mRNA	messenger RNA
NAC	Nascent chain-associated complex
NF	Nuclease free
NTA	Nitrilo-triacetic acid
OAc	Acetate
OD	Optical density
OmpA	Outer membrane protein A
PAGE	PolyAcrylamide Gel Electrophoresis
PBS	Phosphate buffered saline
PCR	Polymerase Chain Reaction
PDB	Protein Data Bank. http://www.rcsb.org/pdb/
PPIase	Peptidyl prolyl isomerase
RAC	Ribosome-associated complex
RNA	Ribonucleic acid
RNase A	Ribonuclease A
RNase T1	Ribonuclease T1
RPL	Ribosomal protein of the large subunit
RPS	Ribosomal protein of the small subunit
rRNA	ribosomal RNA
<i>S. cerevisiae</i>	<i>Saccharomyces cerevisiae</i>
SDS	Sodiumdodecylsulfate
Spec	Spectinomycin
TCA	Trichloroacetic acid

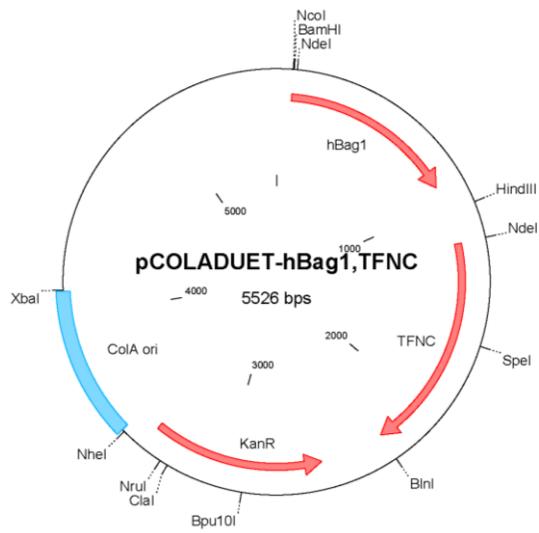
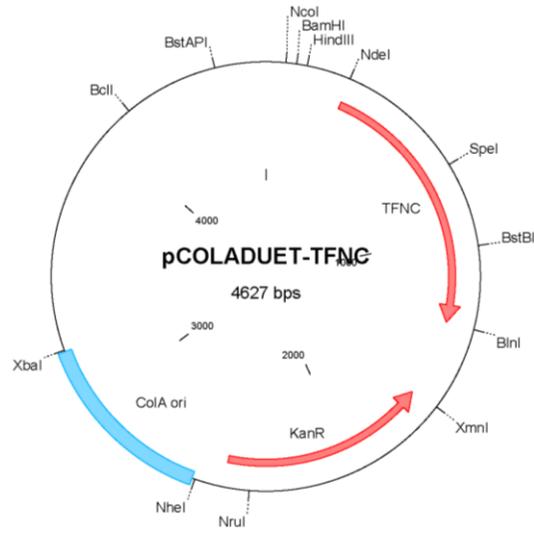
TCEP	tris-(2-carboxyethyl)phosphine
TEMED	<i>N,N,N',N'</i> -tetramethylethylenediamine
TEV	Tobacco etch virus
TF	Trigger Factor
TFNC	PPIase domain deletion mutant of Trigger Factor
<i>tig</i>	gene encoding TF
TRiC	TCP1 Ring Complex
Tris-HCl	tris(hydroxymethyl)aminomethane hydrochloride
UV	Ultraviolet
WT	Wild type

VIII.2 Plasmid maps









VIII.3 Publications

- Lakshmipathy SK, Gupta R, Pinkert S, Etchells SA and Hartl FU
Versatility of Trigger Factor Interactions with Ribosome-Nascent Chain Complexes
J. Biol. Chem. 2010 Sep 3; 285(36):27911-23
- Gupta R, Lakshmipathy SK, Chang HC, Kaiser CM, Etchells SA and Hartl FU
Trigger factor lacking the PPIase domain can enhance the folding of eukaryotic multi-domain proteins in *E. coli*
FEBS Letters 2010 Aug 20; 584(16):3620-4

VIII.4 *Curriculum vitae*

Name Rashmi Gupta

Birth date 17th October, 1982

Birth place Delhi, India

Nationality India

Dissertation

Sept 2005-Aug 2010 PhD dissertation under the supervision of Prof. Dr. F. Ulrich Hartl in the Department of Cellular Biochemistry, Max-Planck Institute of Biochemistry, Martinsried: Function of the *E. coli* chaperone Trigger factor - Role in nascent chain binding and folding delay of multi-domain proteins.

University Education

Jul 2004-May 2005 Masters thesis under the supervision of Dr. K. J. Mukherjee at School of Biotechnology, Jawaharlal Nehru University, Delhi: Metabolic Engineering of *E. coli* for improving recombinant protein production.

Jul 2003-May 2005 Master of Science in Biotechnology from School of Biotechnology, Jawaharlal Nehru University, Delhi.

Jul 2000-May 2003 Bachelor of Science (Honors) in Biochemistry from Sri Venkateswara College, University of Delhi, Delhi.

Higher Education

Jun1998-Jun 2000 High School from Kendriya Vidyalaya, Andrews Ganj, Delhi