Chapter 4: Discussion

4.1 Highly conserved bacterial factor LepA

The highly conserved structure of LepA present in all bacteria, mitochondria and probably also chloroplasts implies a highly conserved domain structure as well. LepA has five domains, the first four have a conspicuous similarity to domains I to III and V of EF-G while the complete EF-G domain IV is lacking. Furthermore, the unique LepA CTD (domain V) has no analog in the known protein world.

We used LepA domain III and IV (corresponding to III and V from EF-G) as well as the unique CTD as probes in order to avoid false hits of EF-G or the corresponding factor EF2 in archaea and cytoplasm of eucarya. We found LepA orthologs in all bacteria and eukaryotes with mitochondria, but not in archaea. Therefore, this observation suggests that LepA does not contribute to eukaryotic cytoplasmic translation, but is probably essential for correct mitochondrial translation. LepA is probably also ubiquitous in chloroplasts, since we found LepA in chloroplasts (nuclear encoded) of all three known plant genomes, *viz*. in the dicotyledon *Arabidopsis thaliana*, the monocotyledon *Oryza sativa* and the red alga *Cyanidioschyzon merolae* (Figure. 3.1_1, page 76).

The evolutionary tree shown in Figure 3.5.1(page 103) allows a number of interesting conclusions. (i) Since the total branch lengths of the EF-G and LepA families is comparable, EF-G and LepA seem to be originated at about the same time. (ii) In both EF-G and LepA families the prospective mitochondrial eukaryotic proteins are monophyletic. (iii) The fly EF-G and LepA proteins cluster with deuterostome sequences to the exclusion of worm sequences. This lends support to the coelomata hypothesis versus the ecdysozoa hypothesis (insects and nematodes form a common clade). (iv) Prospective plastid LepA sequences of rice, mouse-ear cress, and a red algae cluster with cyanobacterial sequences. This means that nuclear-encoded chloroplast LepA was once encoded in the plastid precursor genome. Plastid LepA genes were transferred to the nuclear genome where they acquired signal sequences for targeting the LepA protein back to plastids (Figure 3.5.1, page 103).

4.2 LepA is stongly related to the cytomembrane

It has been reported that most of overproduced LepA was found in the periplasmic membrane (March and Inouye, 1985a). And also quite recent there is a paper (Butland et al., 2005) reporting that LepA is mostly involved in two classes of protein networks, namely ribosomal proteins and membrane associated protein (Table 4.2_1). When LepA was used as bait, 20 proteins were found to have strong interaction with LepA, 10 out of which are ribosomal proteins distributed on both large and small ribosomal subunits. When traced back these proteins onto ribosome, they are located near two functional centers of the ribosome: the decoding center and the exit of the tunnel (Figure 4.2_1).

Consistent with the suspected link to membrane proteins, in wild-type cells grown in rich LB medium we observed that only 10% of LepA is in the cytoplasmic fraction while 90% is located in the membrane fraction (that is removed for lysate preparation; see Table 3.2.3.1, page 83). It was also found that half the cellular content of ribosomes was located in the membrane fraction, considering this the molar ratio of LepA to ribosomes in the total cell is about 0.4:1 and in the lysate (S30) preparation about 0.1:1. This is consistent with our observation that, for example, in our Mg^{2+} experiment the addition of 0.3 LepA molecules per ribosome is optimal concerning amount and quality of synthesized GFP (Figures 3.4.1.1_1, page 98). Surprisingly, when there was high Mg^{2+} concentration in the LB medium, wild type cells had a strikingly different LepA distribution: half of it was found in the cytoplasm and half in the membrane fraction. In this case, the 70S distribution was also changed: in the membrane fraction 60% of the ribosomes were located and only 40% in cytoplasm (Table 3.2.3.2, page 84). This means that under ionic stress the cell apparently requires more LepA in the cytoplasm probably to serve better the ribosomes maltreated by the ionic stress. Nevertheless, the total amount of both LepA and 70S did not change much and the total molar ratio of LepA to 70S is still around 0.5.

Table 4.2 **LepA network in *E. coli***

Bait	Prey (Locus name)	Protein Name	function	MALDI -Tof	LC-MS/MS
b1248	lepA	lepA			Y
lepA	lepA	lepA		Y	Y
lepA	argA	Amino-acid acetyltransferase (EC 2.3.1.1) (N-acetylglutamate synthase) (AGS) (NAGS).			Y
lepA	b0878	Macrolide-specific efflux protein macA precursor	Efflux transporter for macrolide antibiotics		Y
lepA	b1342	Hypothetical protein ydaN			Y
lepA	hyfE	Hydrogenase-4 component E	Integral membrane protein		Y
lepA	recE	Restriction alleviation and modification enhancement protein			Y
lepA	rplD	L4			Y
lepA	rplI	L9			Y
lepA	rplM	L13			Y
lepA	rplV	L22			Y
lepA	rplX	L24			Y
lepA	rpsB	S2			Y
lepA	rpsE	S5			Y
lepA	rpsF	S6			Y
lepA	rpsG	S7			Y
lepA	rpsJ	S10			Y
lepA	ycjJ	Hypothetical transport protein ycjJ.	Probable amino-acid or metabolite transport protein.		Y
lepA	yeiN	Hypothetical protein yeiN			Y
lepA	yidX	Hypothetical protein yidX	transmembrane		Y

Ribosomal proteins

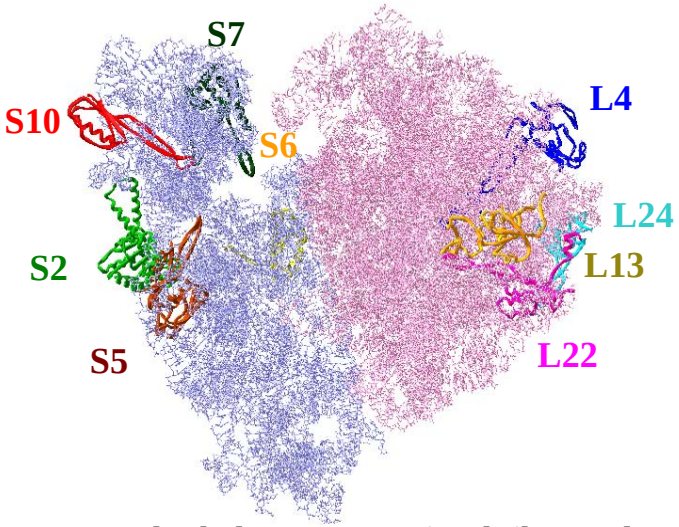

Figure 4.2 **Trace-back the LepA associated ribosomal proteins**

4.3 Structural mimicry of elongation factor

The lack of EF-G domain IV in the LepA structure is intriguing. Our finding that LepA is a back-translocator fits to the early suspicion that the domain IV of EF-G has a “door-stop” function by occupying the decoding region of the A site after the tRNAs have been translocated from A and P sites to the P and E sites, respectively, thus preventing a back movement of the tRNAs as long as EF-G is on the ribosome ((Nierhaus, 1996a), also Figures 3.1_2 and 3.1_3, page 77). In this frame LepA reduces the activation barrier between PRE and POST states similar to EF-G, but due to the absence of domain IV it catalyzes a back-translocation rather then a canonical translocation.

A closer inspection of the domain comparison between LepA and EF-G shown in Figure 3.1_3 (page 78) LepA also lacks in addition to EF-G domain IV the G’ subdomain. From the view of the *E. coli* LepA networks, lacking G’ domain might lead to a loose interaction of LepA with ribosomal GTPase associated center. Earlier, It has been speculated that the function of G’ might be to promote the GDP-GTP exchange as EF-Ts does for EF-Tu (Czworkowski et al., 1994). However, the GDP-GTP exchange on EF-G can also be explained without the help of an additional factor or G’ subdomain (Nierhaus, 1996b). The fact that LepA shows an uncoupled GTPase activity in the presence of 70S paralleling that of EF-G (Figure 3.3.2.1, page 90) - despite the absence of the G’ subdomain - argues against the assumption that this domain is involved in GDP-GTP exchange.

The first crude Cryo-EM reconstruction of LepA•POST complex also shows an additional density under L7/L12 stalk of ribosome large subunit, where is the normal location of EF-G (Figure 3.3.1.4, page 89). Put all these structural analysis together, LepA is an elongation factor with a strong similarity to EF-G except EF-G domain IV, EF-G subdomain G’ and LepA C-domain. Both factors share the same binding site so that the binding is mutually exclusive. This has an impact to Figure 3.3.1.2_2 (page 87) where we show that EF-G and LepA both bind to about 50% of the empty ribosomes under factor saturating conditions, regardless whether the binding of each factor is studied separately or together. The important conclusion is that empty ribosomes are present in two conformers, one is similar to the PRE state and thus substrate for EF-G

binding, the other might be more similar to the POST state specifically recognized by LepA. Both states seem to be present with the same population size of 50%. Evidence for two states of empty ribosomes were already reported before by other studies (Kutay et al., 1990; Mesters et al., 1994).

For the unique CTD of LepA, which might be related to the cytomembrane and co-regulate the cell irritability under high ionic stress, might has some relation with the exit tunnel interaction network.

4.4 First back-translocator

The first experimental hint for the back-translocation activity of LepA came from two separate functional tests, the puromycin reaction and dipeptide formation. Both Pi and POST states with an AcPhe-tRNA donor at the P site are usually equally good substrates for peptide-bond formation using puromycin or an aminoacyl-tRNA as acceptors at the A site. The essential point is that LepA prevents peptide-bond formation exclusively of the POST state leaving the Pi state unaffected (Figures 3.3.3_2, page 92). A likely interpretation was that LepA induces a back-translocation by shifting the tRNAs from E and P sites back to the P and A sites, respectively. Since the A site is now filled with AcPhe-tRNA, this prevents binding of both puromycin and aa-tRNA, and thus peptide-bond formation with both substrates.

This interpretation could be substantiated by three structural assays monitoring (i) the tRNA occupancy of the A site *via* protection of diagnostic rRNA bases of the A site, (ii) the functional state – PRE or POST – of the ribosome *via* conformation-specific Pb^{2+} cleavage, and (iii) the movement of the ribosome on the mRNA *via* toeprinting. Protection of residues A1408 and U531 of the 16S rRNA is diagnostic for the presence of a tRNA at the A site (Moazed and Noller, 1990). Upon administering LepA•GTP to the POST state ribosome, which usually has an empty A-site and therefore shows no A-site tRNA footprints, protection of these A-site specific positions was observed again, thus arguing for the re-occupation of the A-site by the peptidyl-tRNA (Figures 3.3.5.1_A and B, page 96). Pb^{2+} cleavages occur at distinct binding pockets of RNAs and are therefore very sensitive to conformational changes. Cleavage at position C2347 of 23S rRNA is strong in the POST and weak in the PRE states (Polacek et al., 2000), and LepA reduces the cleavage level from that of the POST to that of the PRE state (Figure

3.3.5.1_C, page 96). Eventually, the toeprinting assay (Hartz et al., 1991) demonstrated directly the back-movement of the ribosome on the mRNA by one codon upon adding LepA to the POST state (Figure 3.3.5.2, page 97). Such a back-translocation cannot take place with a Pi state, where the ribosome carries only a single tRNA, since an A site cannot be filled with a tRNA in a stable fashion, when the adjacent P site is free. The fact that LepA functions only with the POST state rather than with a Pi state means that its function depends on the ribosome having an occupied E site. This requirement is a strong indication that the E site also exists in mitochondrial ribosomes, for which the number of tRNA-binding sites has not yet been assessed.

Inside the mitochondria large changes in ionic strength are unlikely, since this intracellular organelle exists in homeostasis, *viz*. in a relatively constant environment. However, it is possible that the extremely reduced rRNAs of mitochondria (in higher animals the mitochondrial rRNA are about 30% shorter than the corresponding *E. coli* rRNAs) may constrain the translocation reaction, such that ribosomes, which become stuck due to incomplete translocation, must be healed by the back-translocation activity of LepA. We note however that this must be true only under specific and as yet unknown conditions, because a knock-out of the LepA ortholog in yeast mitochondria exhibits no phenotype (Kiser and Weinert, 1995). Be it as it is, the extreme conservation of both the domain structure and the amino-acid sequence in all currently available sequences of mitochondrial LepA orthologs signals that an important function for this protein must also exist in this organelle.

4.5 LepA is an essential factor at high ionic condition

How we can reconcile the *in vivo* and *in vitro* effects of LepA to a molecular description of its function? We think that the following scenario describes best the various *in vivo* and *in vitro* results:

EF-G dependent translocation might not be successful in 100% of the cases,

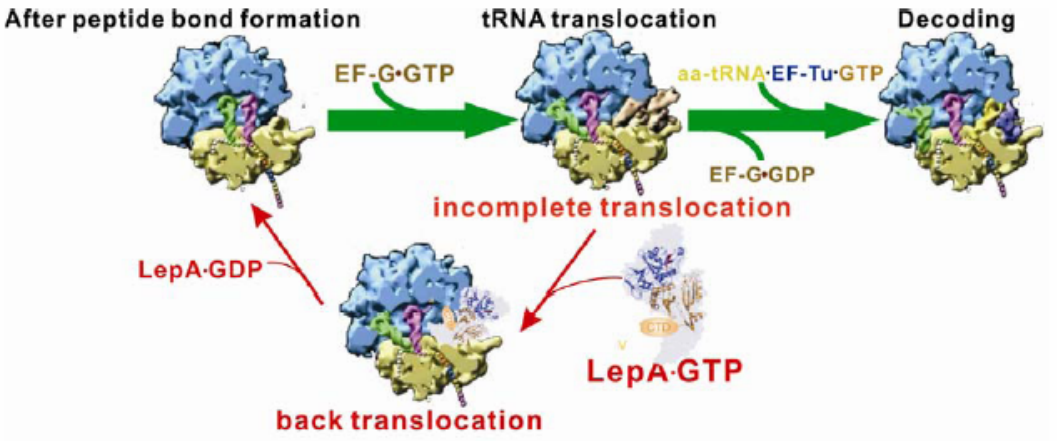

Figure 4.5 **Hypothesis of LepA function**

particularly at higher Mg2+ concentrations, where the ribosome might not reach the canonical POST state, but rather become stuck during the course of a translocation reaction. This would explain the lethal phenotype of the LepA knock-out mutant in conditions of high ionic strength. The results presented in this thesis suggest that low concentrations of LepA (≤0.3 molecules per 70S ribosome) specifically recognize ill-translocated, stuck ribosomes, back-translocates them, thus providing EF-G a second chance for a proper translocation reaction (Figure 4.5).

At higher concentration (~1 molecule per 70S) LepA looses its specificity and back-translocates every POST ribosome, thereby turning the translational machinery into a non-productive mode. This is seen by the inhibition of high concentrations of LepA in the coupled transcription-translation system (Figure 3.4.1.1, page 98) as well as explaining the toxicity of overexpressing LepA *in vivo*. From these results it is clear that the intracellular level of LepA must be precisely tuned and regulated to restrict it to the narrow beneficial concentration window. We finally note a potential application for LepA derived from results of the coupled transcription-translation system (Figure 3.4.1.1),

namely that the addition of a small, defined amount of LepA to bacterial lysates significantly improves the protein output combining both high yield and full activity of the synthesized protein. This illustrates not only the importance of LepA for the protein synthesis in the bacterial cell, but paves the way to the development of more efficient in vitro transcription-translation systems.

4.6 An application of LepA in protein in vitro *synthesis system*

The most comprehensive *in vitro* systems for protein synthesis are coupled transcription-translation systems with cell lysates, where one adds, for example, T7 polymerase and a plasmid carrying a gene under a T7 promoter. The T7 transcript programs the translational apparatus of the lysate yielding up to 4 mg of synthesized protein per ml. Such systems are offered commercially by several firms and are important tools for structural and functional studies of proteins. Examples of the usage of these systems include the synthesis of toxic proteins that might be difficult to express *in vivo*, expression of heterologous proteins from organisms that might be difficult to cultivate in order to crystallize and/or to perform functional studies, synthesis of proteins doted with deuterium, ^{13}C and ^{15}N incorporation for NMR structure determination in solution, incorporation of artificial amino acids, such as selenomethionine, at specific protein positions for crystallization or pharmaceutical applications etc.

The major drawback of the current available systems is the low accuracy with which the proteins are produced, *i.e.* the active fraction of distinct proteins can be as low as 30% of the total protein fraction for a given protein, therefore compromising the use of these protein products for subsequent molecular analysis. Here we demonstrate that addition of the detected ribosomal factor LepA improves the accuracy of the synthesized proteins to about 100% without significantly affecting the protein yield.

Addition of small amounts of LepA slightly reduces the total synthesis but increases the active fraction to virtually 100%. This is important if the structure of synthesized proteins should be determined via crystallization or after doting the synthesized protein with isotopes such as [^{13}C] or [^{15}N] for NMR analysis. Likewise, an analysis of the function of the synthesized protein becomes prohibitively difficult by a large inactive fraction of the protein under observation. A possible reason for the large inactive fraction

might be that the translocational efficiency of EF-G is not 100%. If for example the fraction of the optimally translocated ribosomes were 99% and the sub-optimally translocated ribosomes are preferentially recognized by LepA and back-translocated, then EF-G gains a second chance to translocate the tRNAs correctly, thus improving the overall translocational accuracy by a factor of 100. Sub-optimally translocated ribosomes can lead to stuck polysomes, a lethal situation of the cell. A second not probably alternative possibility exposes the A-site codon in an improper way, thus causing errors in the selection of the aminoacyl-tRNAs and generating inactive proteins. Healing these unfavorable situations might be the main function of LepA.