Universidad del País Vasco
Euskal Herriko Unibertsitatea

ZTF-FCT
Zientzia eta Teknologia Fakultatea
Facultad de Ciencia y Tecnología

FACULTY OF SCIENCE AND TECHNOLOGY. LEIOA

BACHELOR'S FINAL WORK BIOTECHNOLOGY

STUDY OF THE *IN VIVO* FOLDING OF THE POTASSIUM CHANNEL Kv7.2

Student: Marchena Hurtado, Javier
Data: June 2019

Director
Dr. César Augusto Martín Plágaro

Co-Director
Dr. Alejandra Aguado Martínez

Academic year
2018/19

Index

1. INTRODUCTION

1.1. POTASSIUM CHANNEL KV7.2

During this Bachelor's Final Work, the voltage-dependent potassium channel Kv7.2 was studied. This channel is primarily located in nervous system cells where together with Kv7.3 channel subunits, forms heteromeric channels responsible for the M current (Otto et. al., 2006).

The M current is a neuronal excitability regulatory current that avoids the repetitive firing of action potentials. Kv7.2 channels open after the firing of an action potential allowing potassium ions to flow out of the cell. This potassium movement turns the membrane potential more negative, bringing the membrane back to the resting potential and making it more difficult to fire new action potentials.

As Chung (2014) indicated, Kv7.2 channels form tetramers where each subunit consists of 6 transmembrane domains (S1-S6) and an intracellular C-terminal tail composed of 4 helices (A-D) (**Figure 1**). The S_5 and S_6 domains form the pore, while the intracellular helices are very important for the regulation of channel activity though interaction with different regulatory molecules.

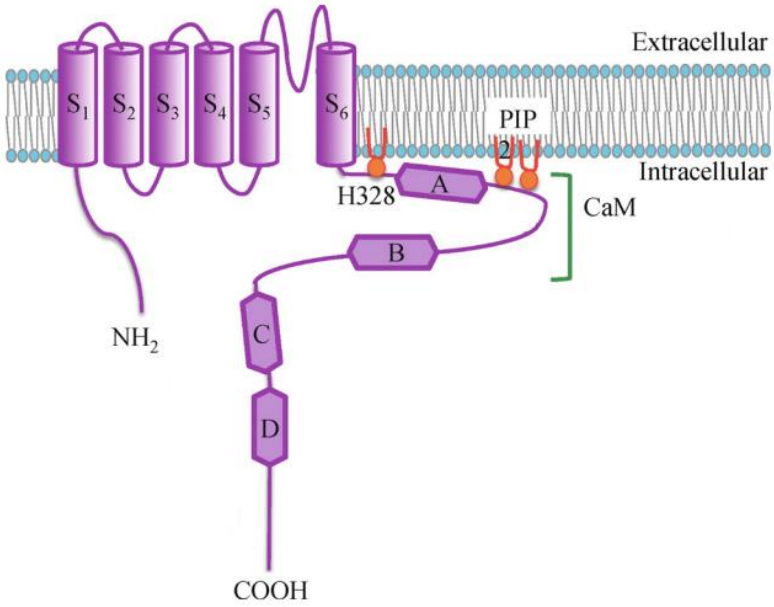

Figure 1. Structure of the potassium channel Kv7.2. The channel contains six transmembrane domains (S1-S6) and four intracellular helices (A-D). The interaction of Calmodulin (CaM) with helices A and B is crucial for channel function. Adapted from Chung (2014).

Among the different molecules implicated in channel regulation, the interaction of Calmodulin (CaM) with the channel is of particular importance. Calmodulin is a calcium sensing protein, that mediates Kv7.2 regulation by Ca^{2+} ions through CaM interaction with the channel cytosolic A and B helices. The subsequent binding of calcium ions provokes a conformation change in Calmodulin, which leads to a reorientation of helices A and B promoting channel closure and the subsequent M current inhibition (Bernardo-Seisdedos et. al., 2018). Additionally, Calmodulin assists in the trafficking of Kv7.2 to the cell membrane. When the CaM-Kv7.2 binding sites were mutated, the surface expression of the channel was reduced, and Kv7.2 accumulated in the endoplasmic reticulum (Alaimo et. al., 2009). All in all, Calmodulin interaction is necessary for the good functioning of potassium channel Kv7.2.

A proper functioning of Kv7.2 is essential for the regulation of neuronal excitability. In fact, mutations in this channel have been described to provoke serious epileptic diseases such as Benign Familiar Neonatal Convulsions (Goldberg-Stern et al., 2009), Rolandic epilepsy (Neubauer et. al., 2008) or epileptic encephalopathy (Weckhuysen et. al., 2012). The study of the structure and regulation of the potassium channel Kv7.2 is paramount in order to treat those diseases.

1.2. COTRANSLATIONAL FOLDING

The term “cotranslational folding” refers to the folding of a protein while it is being translated in the ribosome. Those first instants in which the nascent chain starts folding are crucial for the final configuration of the protein.

Most of our current understanding about protein folding comes from *in vitro* experimental approaches where folding is usually studied from purified and afterwards denatured proteins. These methods, however, substantially differ from the *in vivo* folding process, where nascent polypeptidic chains are synthesized and usually cotranslationally folded in the ribosome under complex physiological conditions. In order to approach protein folding from a more biological point of view, different technics have been developed including high resolution structural methods or the use of arrest peptides.

The structure of the potassium channel Kv7.2 has been well described, but its cotranslational folding has received little attention. Cotranslational protein misfolding could be responsible for some of the epileptic diseases caused by mutations in the Kv7.2 channel. Particularly, there is the suspicion that the mutant W344R, an epilepsy-causing Kv7.2 inactive variant harboring a tryptophan by arginine substitution in the protein helix A, could undergo some cotranslational misfolding. To decipher if this is the case, first the cotranslational folding of wild-type helices A and B must be studied. To this aim, in this work the so-called arrest peptides (APs) were used for *in vivo* protein cotranslational folding studies.

1.3. ARREST PEPTIDES

Arrest peptides are gene expression regulator sequences that interact specifically with the ribosome exit tunnel, arresting the translation of the nascent peptide. The most widely known arrest peptide is the one encoded by SecM. SecM is a prokaryotic periplasmic protein containing an arrest peptide which regulates the levels of the cotranscribed translocation motor protein SecA (Butkus et. al., 2003). Owing to the interaction of the arrest peptide with the ribosome exit tunnel, the A-site and P-site of the ribosome are altered causing the stalling of the translation (Zhang et. al., 2015).

If the arrest peptide stalls the translation, the following amino acids are not synthesized, so a shorter, truncated version of the protein is produced (Nakatogawa & Ito, 2001). In contrast, in the absence of translation arrest, the full-length protein will be synthesized.

There is a key aspect in this technique. If a force is exerted outward of the ribosome while the arresting peptide is interacting with the ribosome exit tunnel, that pulling force relieves the stall of the translation. This relief allows the translation to continue, so the full-length protein is synthesized (Ismail et. al., 2012). That force can be generated when the nascent polypeptide emerges and folds in close proximity to the ribosome exit tunnel. The protein folding generates a force that can relieve the stall of the translation (**Figure 2**).

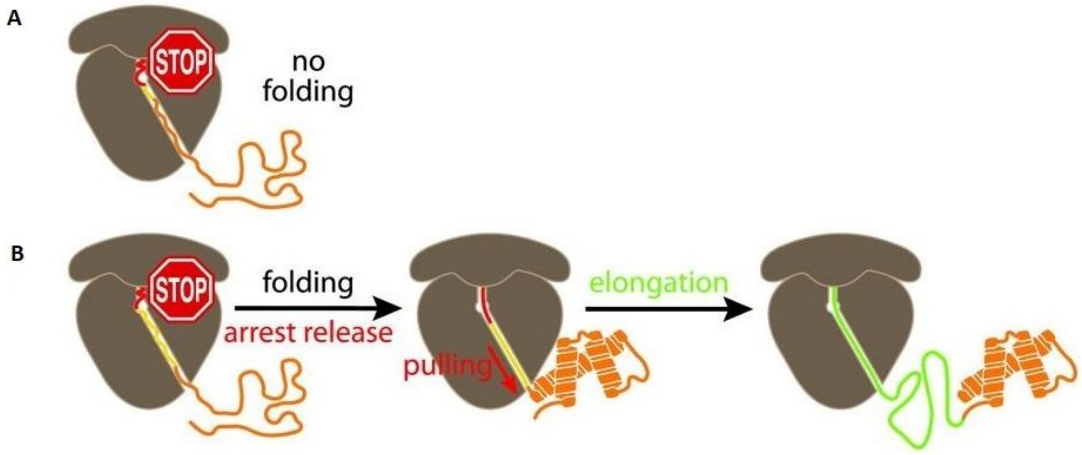

Figure 2. Relief of the arrest peptide stalling. A) If there is no folding force acting when the arrest peptide interacts with the ribosome exit tunnel, the translation is stalled. **B)** If the protein folds at the exact point in which the arrest peptide is interacting with the ribosome exit tunnel, the force generated by protein folding can relieve the stall of the translation. Adapted from Goldman et. al. (2015).

In this work, the pulling force will be caused by the folding of helices A and B. Folding has to take place exactly when the arrest peptide is interacting with the ribosome exit tunnel in order to relieve the stall and produce the full-length protein. If folding takes place earlier or later than the moment when the arrest peptide sequence interacts with the ribosome exit tunnel, the stall will not be relieved and a truncated form of the protein will be produced.

Hence, arrest peptides can be used as "*in vivo* force sensors" (Cymer & von Heijne, 2013). Comparison of full-length and truncated protein fractions will provide information about *in vivo* protein folding, as it will indicate if the helices are cotranslationally folded.

2. OBJECTIVES

The goal of this work is to study the cotranslational folding of Kv7.2 cytosolic A and B helices. It has been hypothesized that the epilepsy-causing mutant W344R (or even other mutants) could undergo some form of cotranslational misfolding. In this project, we aim to establish an arrest peptide-based method that in the future allows studying the *in vivo* cotranslational folding of those epileptic variants.

Hence, the objective of the work is twofold:

1) To test the use of arrest peptides as a technique to study cotranslational folding.

2) To determine if the A and B helices are cotranslationally folded in the ribosome vestibule.

3. MATERIALS AND METHODS

3.1. CONSTRUCTS USED IN THIS WORK

In order to study the *in vivo* folding of Kv7.2 cytosolic A and B helices, the designed constructs were composed by an arrest peptide separated from the AB helices by a variable-length linker and flanked by reporter mTFP1 and mcpVenus173 fluorescent proteins in the N- and C-terminal ends respectively (**Figure 3**).

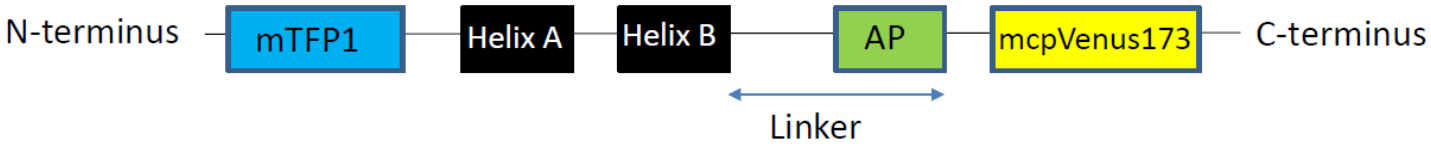

Figure 3. Constructs used in this work. 5 constructs were designed with different linker length, varying the amino acid number between Helix B and the arrest peptide (AP). The fluorescence of mTFP1 will account for the total protein, whereas the fluorescence of mcpVenus173 will account for the full-length protein.

The linker length is of especial importance for protein translation arrest efficiency. If the linker is too short, the protein will still be confined inside the ribosome exit tunnel, so it will not be able to fold. In the absence of folding force, translation will be stalled. If the linker is adequate (linkers with an intermediate length), the protein will fold in the ribosome vestibule, at the exact point in which the arrest peptide is interacting with the ribosome exit tunnel, thereby relieving the stall. Finally, if the linker is too long,

the protein will already be folded by the time the arrest peptide starts interacting with the ribosome exit tunnel, so the translation will be stalled (**Figure 4**).

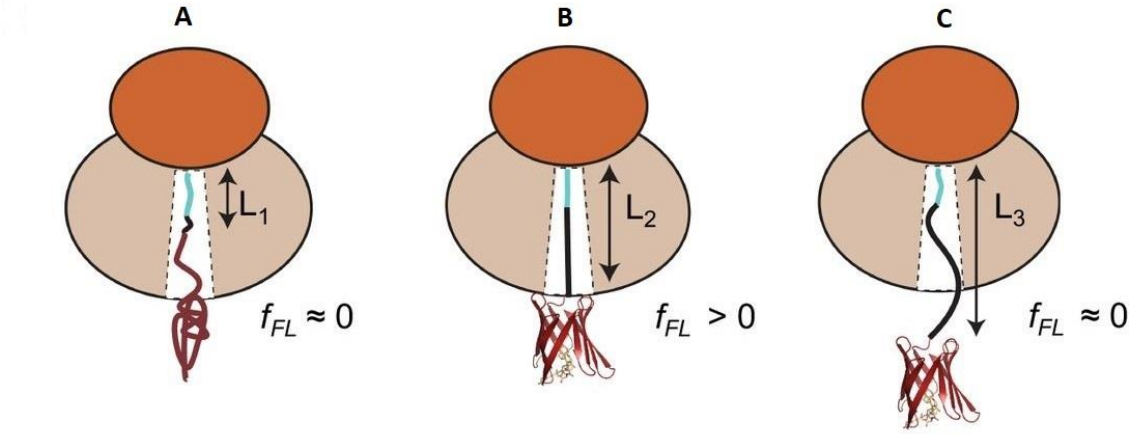

Figure 4. Importance of the linker length. The arrest peptide is depicted in blue, the linker in black and the protein in red. **A)** If the linker is too short, the protein will still be inside the ribosome exit tunnel. As protein folding does not occur in such a confined space, the arrest peptide stalls the translation. **B)** If the linker is adequate, the protein will fold at the exact point in which the arrest peptide is interacting with the ribosome exit tunnel, generating a force that relieves the stall. **C)** If the linker is too long, the protein will be already folded by the time the arrest peptide interacts with the ribosome exit tunnel. As the folding force will not be transmitted to the arrest peptide in the appropriate moment, protein translation will be stalled. Adapted from Farias-Rico et. al. (2018).

Due to its impact on protein translation arrest, a 49 residue-long linker containing restriction enzyme sites was used as template to obtain linkers of 21, 26, 33 and 39 residues long. The arrest peptide sequence length was included in the length of the linkers (**Figure 3**).

Fluorescent proteins were used in order to quantify the full-length and truncated protein fractions. mTFP1 and mcpVenus173 fluorescent proteins were located at the construct N and C-terminus respectively (**Figure 3**), so the truncated version will not have mcpVenus173. Therefore, mcpVenus173 will account for the full-length peptide while mTFP1 will be present both in the full-length and in the truncated protein forms, so it will indicate the total protein.

3.2. MOLECULAR CLONING

3.2.1. Vectors

The constructs containing AB helices with variable linker lengths were already cloned in the pProEx-HT C vector for further protein expression. However, it was very useful to clone them in a pET expression system in order to have a more restrictive and

controlled protein expression. The chosen vector was pET-14b. The most important characteristic of pET vectors is the presence of a T7 promoter. Therefore, in order to express the construct, the bacteria used must contain a copy of the T7 RNA polymerase gene inducible by IPTG. *Escherichia coli* BL21(DE3)plysS strain was used for expression. These cells contain a λDE3 lysogen that carries the T7 RNA polymerase gene. IPTG is required to induce polymerase expression and further expression of the genes cloned downstream the T7 promoter.

Another important feature is the ampicillin resistance gene, which allows the selection of positive colonies through ampicillin.

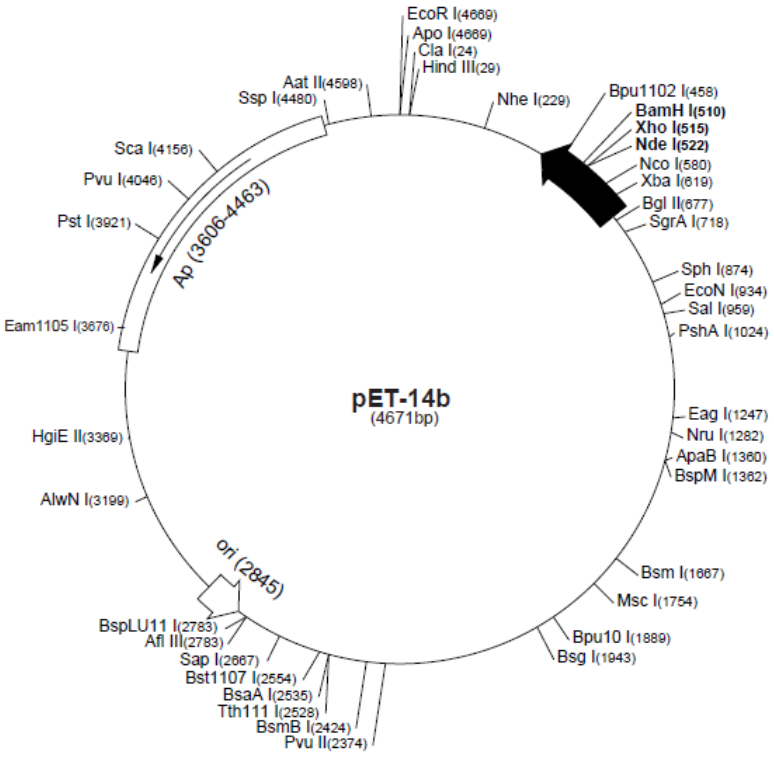

Figure 5. Features of the pET-14b vector. Adapted from Novagen manuals (Merck).

3.2.2. Digestions and ligations

In order to transfer the L49 construct from pProEx-HT C to pET-14b, restriction enzyme digestion was carried out. It was necessary to choose two enzymes that cut before and after the construct in pProEx-HT C, but also had to be present in the pET-14b polylinker, in the correct orientation. The restriction enzymes utilized were Nco I and Xho I. The protocol employed is described in **Table 1**.

Table 1. Protocol employed for DNA digestions. The enzymes and Tango Buffer were bought from Thermo Fischer Scientific. 4 µL of Tango Buffer 10X were added in order to reach a final concentration of 2X.

DNA (1 µg/µL)	Enzyme 1 (10 U/µL)	Enzyme 2 (10 U/µL)	Tango Buffer (10X)	Milli-Q H_2O	Final volume	Temperature	Time
1 µL	0.5 µL	0.5 µL	4 µL	14 µL	20 µL	37ºC	3 h

First, both the L49 construct in pProEx-HT C and the empty vector pET-14b were digested with Nco I and Xho I following the protocol (**Table 1**). The digestion was neutralized by adding 2 µL of Loading Buffer (50% glycerol, 0.2 M EDTA, 0.05% bromophenol blue, pH 8) and loaded in an agarose gel 1% (w/v), using TBE as buffer (45 mM Tris-Borate, 1 mM EDTA, pH 8). Gels were stained with GelRed™ Nucleic Acid (Thermo Fischer Scientific). As DNA length marker, λ DNA digested with BstEII was used.

Agarose gels were run for approximately 40 minutes at 120 V. Afterwards, the bands corresponding to the desired vector and fragment (construct) were excised from the gel and the DNA was purified with a DNA extraction kit (GeneJET Gel Extraction Kit, Thermo Fischer Scientific) according to the manufacturer's instructions.

Finally, the purified vector and construct were ligated with T4 DNA ligase (Thermo Fischer Scientific). A mixture was prepared with 7.5 µL DNA fragment, 1 µL of vector DNA, 1 µL of ligase buffer (50 mM Tris-HCl, 10 mM $MgCl_2$, 1 mM ATP, 10 mM DTT, pH 7.5) and 0.5 µL of T4 DNA ligase. Ligation was performed overnight at room temperature.

3.2.3. Bacterial transformation

For cloning, *Escherichia coli* BSJ strain was used, because it grows quickly. Bacterial transformation was carried out by chemical transformation. The process was executed in sterile conditions. First, 100 µL of chemo-competent bacteria (treated with $CaCl_2$) were thawed on ice. Then, a mixture was prepared with 45 µL of transformation buffer (22 mM $MgCl_2$ and 11 mM $CaCl_2$), 5 µL of ligate and 100 µL of bacteria. The mixture was first incubated 20 minutes on ice, and then 10 minutes at room temperature. Afterwards, to help cells recover from the shock, bacteria were incubated in Luria-

Bertani (LB) growth medium at 37 ºC for 45 minutes. Next, cells were harvested by centrifugation, and seeded into LB-agar plates with ampicillin. Plates were incubated at 37ºC overnight.

3.2.4. DNA extraction

After cell growing in LB-agar plates, single colonies were picked and grown overnight at 37ºC in 3 mL of LB growth medium with ampicillin. 2 mL of this culture were used for DNA extraction, while the other mL was stored as a stock in case the procedure had to be repeated. The cells were harvested from those 2 mL by centrifugation, and the bacterial pellet was resuspended in 150 µL of Tris-EDTA (10 mM:1 mM, pH 8). Then, 150 µL of lysis solution (0.2 M NaOH and 1% SDS) were added and mixed, followed by the addition of 150 µL of neutralization salts (3 M AcK^+ and 5 M AcH). After adding 80 µL of phenol/chloroform (1:1) and vortexing, the tubes were centrifuged. This centrifugation with phenol/chloroform causes lipids to precipitate, proteins to stay in the interphase between pellet and supernatant, and plasmid DNA and RNA to stay soluble in the supernatant. The supernatant was mixed with 200 µL of isopropanol, which causes nucleic acids to precipitate. After centrifugation, the supernatant was discarded and the DNA pellet was resuspended in 30 µL of Tris-EDTA with RNase. This will be the final purified vector DNA, with an approximate concentration of 1 µg/µL.

In order to prepare the aforementioned stock, 700 µL of bacteria were mixed with 300 µL of glycerol 50%. The bacterial stock was stored at -80ºC.

3.2.5. DNA checking

After extraction, the DNA was checked in order to make sure that the purified DNA was the desired one, and not the parental DNA (in this case, the construct in pProEx-HT C) or contamination. With this goal, the purified DNAs were digested with Bgl II and Xho I and analyzed by agarose gel electrophoresis as described before (3.2.2. Digestions and ligations). The Bgl II site is present in pET-14b, but absent in pProEx-HT C, so this checking will be further proof that the construct has been correctly inserted in pET-14b.

The digestions were carried out as described before, except that 0.2 μL of enzymes were used (instead of 0.5 μL) and the digestion time was 2 hours (instead of 3 hours).

3.2.6. Linker deletions

Once the L49 construct was inserted and checked in pET-14b, linker deletions were executed to obtain linker lengths of 39 (L39), 33 (L33), 26 (L26) and 21 (L21) residues long. The deletions were carried out as described before (3.2.2. Digestions and ligations) using the following restriction enzyme pairs: Eco 72I (L39), Sna BI /Stu I (L33), Sma I /Eco 72I (L26) and Sna BI /Eco 72 I (L21).

Once the digestion was completed, the same cloning procedure was repeated for each linker length. They all were run in an agarose gel, ligated (recircularized, in this case), and transformed in *E. coli* BSJ for DNA extraction and subsequent checking).

3.2.7. DNA Sequencing

In order to absolutely check that the constructs had been successfully cloned in pET-14b, DNA sequencing was carried out following the Sanger sequencing method. Samples containing 1.5 μL of the desired primer at a concentration of 100 μM, 1.5 μL of DNA at an approximate concentration of 1 μg/ μL (reaching a final concentration of 0.1 μg/ μL) and 12 μL of H_2O were sent to the Secugen sequencing company. The primer used in this work hybridizes with the final nucleotides corresponding to mTFP1. Thus, the helices and linker were sequenced.

Upon the reception of the sequencing result, the obtained sequence was compared with the expected one using the BLASTn alignment tool (Madden, 2002).

3.3. PROTEIN EXPRESSION

3.3.1. Bacterial transformation

For protein expression, *E. coli* BL21(DE3)plysS strain was used. BL21 are well-known to express heterologous genes, while DE3 means that the strain has the T7 RNA polymerase inducible by IPTG, which is necessary to express pET vectors. plysS is a plasmid that expresses low levels of T7 lysozyme, an inhibitor of T7 RNA polymerase. Hence, the background expression of the target genes is inhibited. When adding IPTG,

however, big amounts of T7 RNA polymerase are produced, much more than what T7 lysozyme can inhibit, so target genes will be expressed without problems.

The transformation procedure was exactly the same as the one described in section 3.2.3. Bacterial transformation. The only difference was that, in this case, 0.5 μL of DNA were used (instead of 5 μL of ligate in section 3.2.3.), because the extracted DNA is much more concentrated than the ligate obtained in section 3.2.2.

3.3.2. Cell growth and protein expression

Two colonies were picked from every plate, and each colony was grown in 10 mL of LB growth medium with ampicillin. The cultures were incubated at 37ºC until the optical density, measured with the spectrophotometer Jenway 6300, reached 0.6 at 600 nm. At this point, cells are expected to be in the exponential growth phase, which is ideal for gene expression. Gene expression was induced by adding IPTG to a final concentration of 0.5 mM, in order to produce T7 RNA polymerase and express the constructs. After 3 hours of gene expression at 37ºC, cells were harvested by centrifugation.

3.3.3. Cell homogenization

The bacterial pellet was resuspended in 500 μL of buffer 120 mM KCl, 5 mM NaCl, 50 mM HEPES pH 7.4, 5 mM EGTA, 1 mM DTT and protease inhibitor. Cells were lysed by sonication, using the Soniprep 150 MSE (SANYO) equipment. 5 sonication cycles were used, each cycle with 3 seconds of activity, 5 seconds of rest and an amplitude of 7.5. The lysate was centrifuged for 10 minutes at 13,000 g to separate the soluble fraction from cellular debris.

3.3.4. Polyacrylamide gel electrophoresis

Protein expression was then analyzed by SDS-PAGE electrophoresis. For that purpose, 10% SDS-polyacrylamide gels of 1 mm thickness were prepared. For sample preparation 20 μL of each soluble fraction was mixed with 5 μL of electrophoresis loading buffer (62.5 mM Tris-HCl, 20% glycerol, 2% SDS, 5% β-mercaptoethanol, 0.05% bromophenol blue, pH 6.8). 15 μL of the mixture of each sample were loaded into the gel. Gels were run with electrophoresis buffer (25 mM Tris-HCl, 192 mM

glycine, 0.1% SDS, pH 8.3) at 120 V until the bromophenol blue reached the end of the gel. As molecular mass marker Protein Marker VI (PanReac AppliChem) was used.

Finally, the gels were visualized using the VersaDoc 4000 (BioRad) image system. First, the gels were excited at 488 nm, in order to excite the mTFP1 and see both full-length and truncated protein bands. Then, the gels were excited at 555 nm, in order to excite mcpVenus173 and observe the full-length protein bands.

3.3.5. Fluorometric assay

The fluorometric assays were carried out in a FluoroMax-3 (Horiba) fluorometer. Quartz cuvettes with a path length of 3 cm and a volume of 100 µL were used. For each linker length, 100 µL of the soluble sample were used to measure the emission spectra of mTFP1 and mcpVenus173 fluorescent proteins upon excitation at 458 and 515 nm respectively. mTFP1 emission spectrum was recorded at 470-570 nm while mcpVenus173 emission was monitored at 520-570 nm.

The ratio of full length (FL)/total protein was calculated as follows and used as an indicator of "*in vivo*" protein folding:

$$\text{Ratio FL/total protein} = \frac{\text{Fluorescence at 528 nm when exciting at 515 nm}}{\text{Fluorescence at 492 nm when exciting at 458 nm}} \quad (1)$$

where 492 is the peak emission wavelength of mTFP1, while 528 is the peak emission wavelength of mcpVenus173.

4. RESULTS

4.1. MOLECULAR CLONING

4.1.1. L49 construct

4.1.1.1. Digestion

First, it was necessary to insert the construct having the 49 residue-long linker in the pET-14b vector. As this construct was already cloned in a pProEx plasmid, both the empty pET-14b vector and the construct-containing pProEx plasmid were digested with Nco I and Xho I restriction enzymes. The empty pET-14b vector has a total length of 4671 bp. After digestion, it was expected to release fragments of 65 and 4606 bp.

The pProEx vector containing the sequence of interest has a total length of 6517 bp, and fragments of 1910 and 4607 bp were expected. The 1910 bp fragment is the L49 construct, which was later inserted in pET-14b.

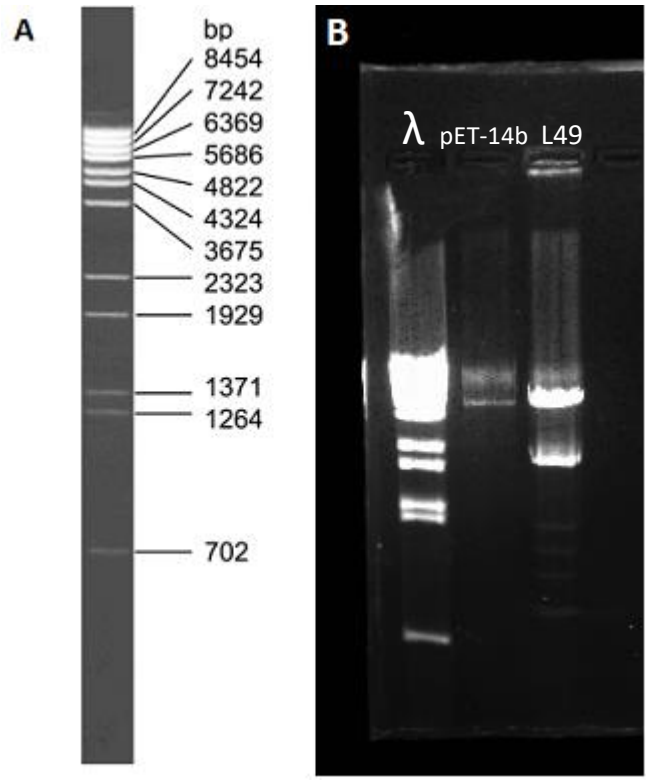

Figure 6. Digestion of the empty pET-14b vector and the L49 construct-containing pProEx vector. A) λ DNA digested with BstEII used as DNA molecular marker. Adapted from SibEnzyme webpage. B) 1x agarose gel with the empty pET-14b and the pProEx vector containing the L49 construct, digested with Nco I and Xho I. The gel was run for 50 minutes at 120 V. Lane 1: λ DNA digested with BstEII used as DNA molecular marker. Lane 2: pET-14b digested with Nco I and Xho I. Lane 3: L49 construct in pProEx digested with Nco I and Xho I.

As shown in **Figure 6**, the expected fragment lengths were obtained. The 65 bp fragment of pET-14b is not observed because, as it is very small, its electrophoretic mobility is very high and it runs out of the gel. To follow with the cloning process, the 4606 bp fragment of pET-14b was excised and purified. On the other hand, the 1910 bp fragment, which constitutes the L49 construct, was also excised and purified. The vector and the L49 construct were ligated to obtain the L49 construct in pET-14b and transformed in *E.coli* BSJ cells.

4.1.1.2. DNA checking

In order to check if the cloning process had been successful, 6 colonies were grown in LB medium. Their DNA was extracted and digested with Bgl II and Xho I restriction enzymes. Bgl II is present in pET-14b but absent in pProEx, so this digestion should provide further proof that the construct was successfully inserted in pET-14b. The digestion was expected to yield two fragments of 2007 bp and 4509 bp.

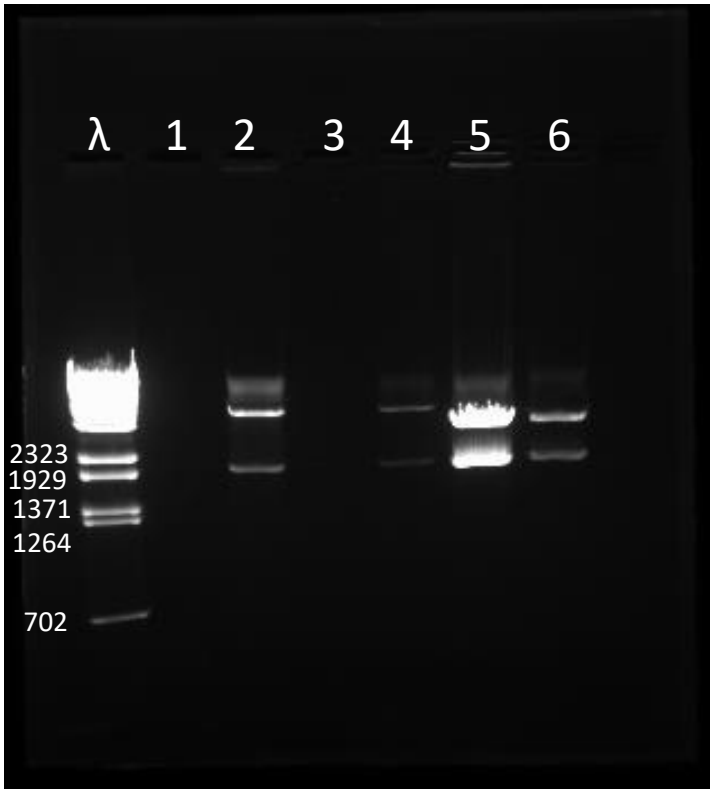

Figure 7. Agarose gel with the L49 construct in pET-14b checked with Bgl II and Xho I. The far-left lane is, as always, the molecular marker λ DNA digested with BstEII. All the other lanes correspond to the DNAs extracted from each of the 6 colonies harboring the L49 construct inserted in pET14b. The agarose concentration was 1X and the gel was run for 40 minutes at 120 V.

As shown in **Figure 7**, lanes 1 and 3 do not contain any DNA, probably because of mistakes committed during the DNA extraction. However, the rest of the checked DNAs presented the expected band pattern. The DNA corresponding to lane 2 was selected as the best one, because it contains a good quantity of DNA and seems to be clean. Thus, lane 2 DNA was used for the subsequent linker deletions.

4.1.2. Linker deletions

4.1.2.1. Digestions

As previously explained, the length of the linker determines the efficiency of protein translational arrest. Thus, the DNA of the L49 construct in pET-14b was digested with different restriction enzymes to obtain constructs with linker lengths of 39 (L39), 33 (L33), 26 (L26) and 21 (L21) residues. This produced two fragments in each digestion: a big fragment, of almost 6516 bp, and a very small fragment. The small fragments were 30 bp (L39), 48 bp (L33), 69 bp (L26) and 84 bp (L21) in length. Therefore, the big fragments were 6486 bp (L39), 6468 bp, (L33), 6447 bp (L26) and 6432 bp (L21) in length.

Figure 8. Agarose gel with the digestions carried out to obtain constructs with different linker lengths. The agarose concentration was 1X and the gel was run for 40 minutes at 120 V. Lanes 2 to 5 show the constructs in which the linker has been digested to a final length of 21 (L21), 26 (L26), 33 (L33) and 39 (L39) residues. The restriction enzymes used for each deletion were the following: Sna BI/ Eco 72 I (L21), Sma I/ Eco 72I (L26), Sna BI/ Stu I (L33), Eco 72I (L39).

As shown in **Figure 8**, the length differences are too low to be clearly appreciated. The gel should be run longer to observe them. The small fragments have too high electrophoretic mobility, and ran out of the gel. Two bands can be seen in L21, probably due to an incomplete digestion of the DNA.

The bands were excised from the gel and the DNA was purified for subsequent recircularization, giving rise to the constructs with different linker lengths in pET-14b. The DNA product was finally transformed in *E.coli* BSJ cells. Different colonies were grown in LB medium and their DNA was checked in order to be sure that linker digestions were performed properly.

3.1.2.2. DNA checking

The DNA checking was carried out with Bgl II and Xho I. The L49 construct was also digested in order to compare the five constructs. In each digestion, two fragments were obtained: a big fragment, 4509 bp long, and a smaller fragment around 2000 bp but with varying lengths. Those lengths were 2007 bp (L49), 1977 bp (L39), 1959 bp (L33), 1938 bp (L26) and 1923 bp (L21) (**Figure 9**).

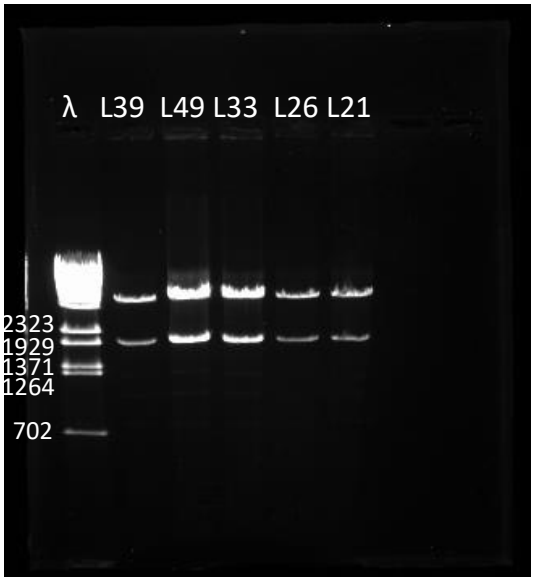

Figure 9. Agarose gel loaded with the constructs with different linker lengths, checked with Bgl II and Xho I. The agarose concentration was 1X and the gel was run for 50 minutes at 120 V.

The constructs were successfully checked. However, the length differences are barely detectable.

4.1.3. Sequencing

In order to absolutely make sure that each linker length was successfully cloned, the constructs were sequenced. The sequencing was executed by the Secugen sequencing company.

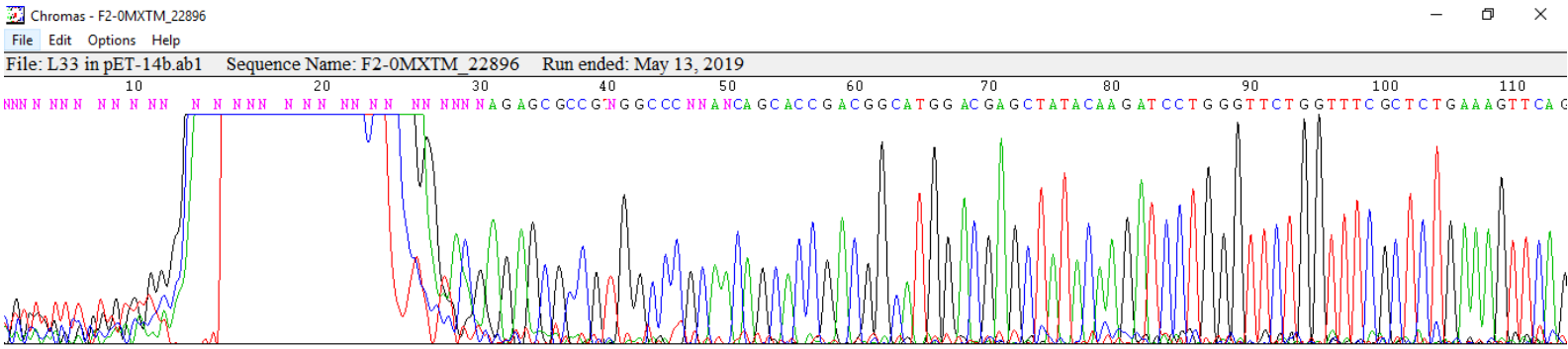

Figure 10. L33 construct in pET-14b sequencing result.

Figure 10 shows the sequencing result of one of the constructs cloned in pET14-b. The first base pairs are commonly read poorly. From there onwards, a high-quality sequencing can be seen. The received sequences were compared with the expected sequences of the constructs. They all coincided, so the constructs had been successfully cloned in pET-14b.

4.2. PROTEIN EXPRESSION

4.2.1. Polyacrylamide gel

In order to study protein expression of the five new constructs, each DNA was transformed in BL21(DE3)plysS. After gene expression and cell disruption, the soluble and insoluble fractions were separated by centrifugation and the soluble fractions were loaded in a 10% SDS-PAGE gel. Two samples were prepared for each construct, each of them extracted from two different bacterial colonies. After gel running, the gel was visualized in a VersaDoc 4000 (BioRad) imaging equipment. The mTFP1 fluorescent signal was detected upon excitation at 488 nm while mcpVenus173 was excited at 555 nm.

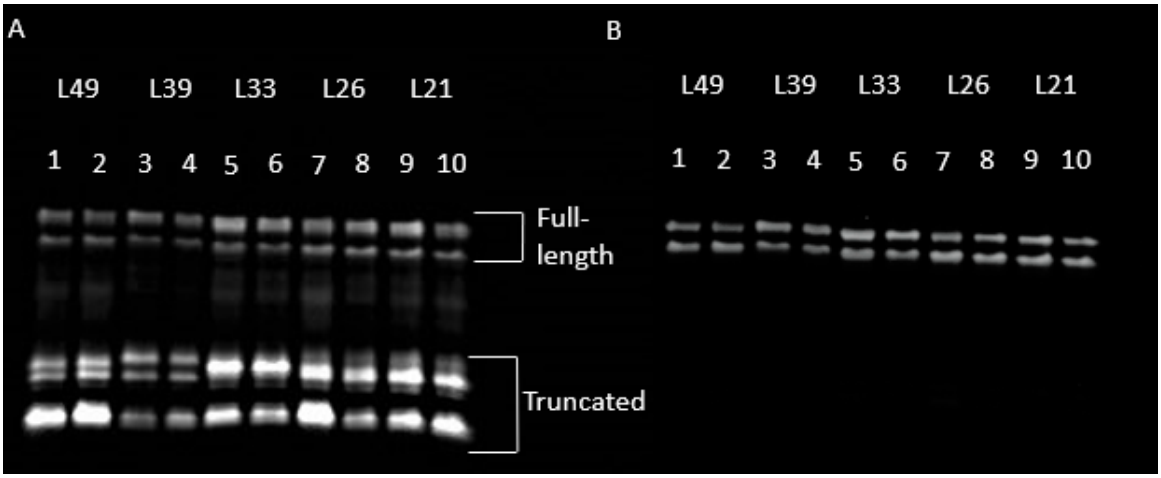

Figure 11. Fluorescently excited polyacrylamide gel with the construct samples. Two independent samples were loaded for each construct length. Constructs containing linker lengths of 49 (L49; 1, 2), 39 (L39; 3, 4), 33 (L33; 5, 6), 26 (L26; 7, 8) and 21 residues (L21; 9, 10) are shown. **A)** Gel excited at 488 nm in order to excite mTFP1. **B)** Gel excited at 555 nm in order to excite mcpVenus173.

The full-length and truncated protein bands are clearly observed and differentiated. On the one hand, full-length protein corresponds to the high molecular weight bands that can be seen in both **Figure 11A** and **11B**, as they express mTFP1 and mcpVenus173. The presence of full-length protein indicates that the some of the AB helices have folded cotranslationally, generating a force that has relieved the stall produced by the arrest peptide.

On the other hand, the truncated protein corresponds to the low molecular weight bands that can be observed just in **Figure 11 A**, since they expressed mTFP1 but not mcpVenus173 fluorescent protein. The presence of truncated protein means that some other AB helices have not been folded cotranslationally. Thus, the arrest peptide has stopped protein translation.

As it can be seen in the image excited at 488 nm (**Figure 11 A**), the truncated bands have much more fluorescence intensity than the full-length ones. This means that most of the protein has been arrested during translation suggesting that most of it has not been folded in the ribosome vestibule.

As a result of protein expression it was expected to obtain single full-length and truncated bands. However two full-length and three truncated bands are observed (**Figure 11**). It seems that the proteins could be somehow hydrolyzed into smaller peptides.

4.2.2. Fluorometric assay

Besides analyzing protein expression by SDS-PAGE electrophoresis, the emission spectra of each construct were also analyzed. Samples were excited at 458 and 515 to excite mTFP1 and mcpVenus173 respectively.

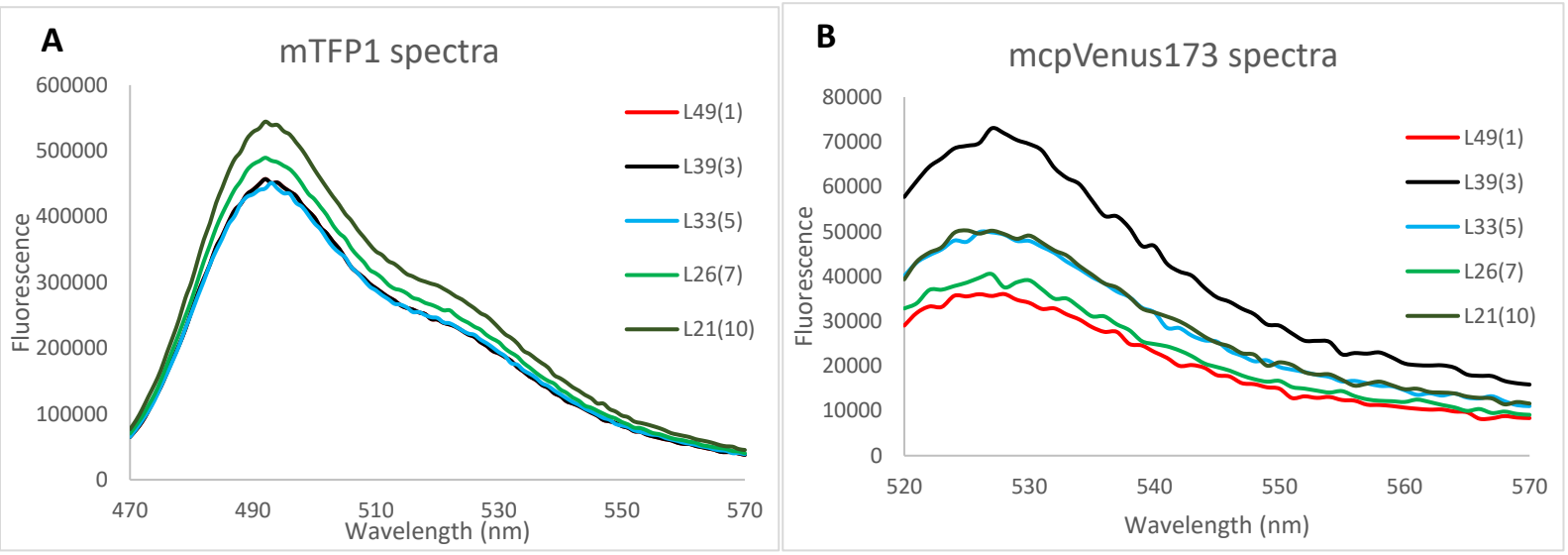

Figure 12. Emission spectra of the constructs containing different linker lengths. For each linker length, one of the two samples measured are represented, in order to make the figure clearer. **A)** Fluorescence emission when exciting mTFP1 at 458 nm. Samples were diluted so that the mTFP1 spectra coincided and differences in mcpVenus173 spectra could be clearly visible. **B)** Emission spectra when exciting mcpVenus173 at 515 nm.

First, samples were diluted with buffer so that the mTFP1 fluorescent intensity was approximately the same in all of them (**Figure 12A**). As mTFP1 fluorescence accounts for both full-length and truncated proteins, dilution of all the samples to the same mTFP1 fluorescence value ensures that in all the cases the total protein amount will be almost the same and makes it easier to detect differences in full length expression. Knowing the total protein amount is equal, then mcpVenus173 emission was recorded (**Figure 12B**), in order to determine the full-length protein fraction.

As shown in **Figure 12B**, L39 is the linker length with the highest mcpVenus173 fluorescence. Therefore, this is the construct in which there is more full-length protein, and there has been less arrest. It is the construct in which the folding of the helices has exerted most force in order to relieve the stall of the translation.

As it can be seen in **Figure 12**, the fluorescence intensity is much higher when exciting mTFP1 than when exciting mcpVenus173. This coincides with the fact that in **Figure 11A** the truncated bands had much more intensity than the full-length ones. To have

an idea about the amount of full-length protein translated, the full-length/total protein ratio was calculated by dividing the mcpVenus173 emission maximum at 528 nm by the mTFP1 emission maximum at 492 nm.

The mcpVenus173/mTFP1 ratio yielded the following results:

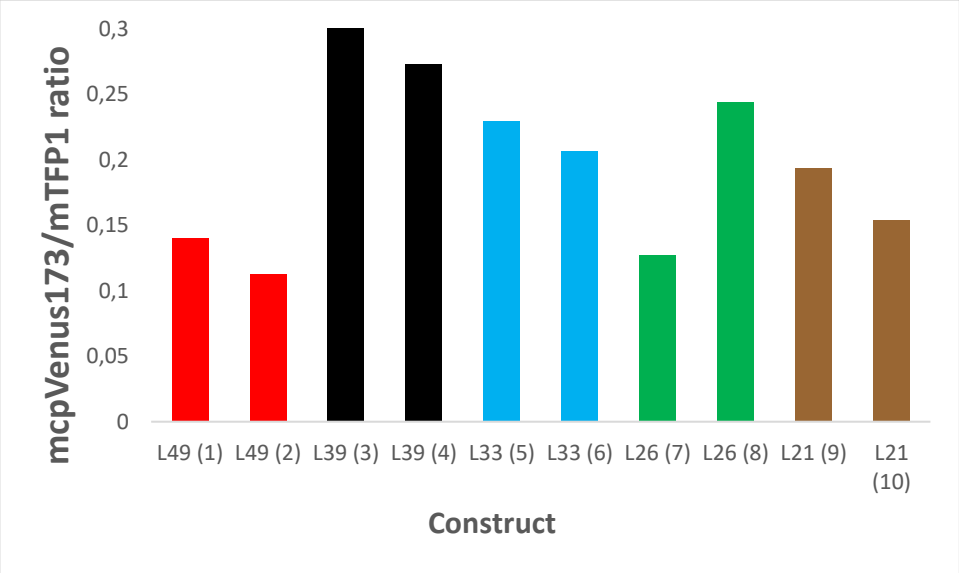

Figure 13. mcpVenus173/mTFP1 ratio for each protein construct. The relative amount of full-length protein was expressed as the mcpVenus173/mTFP1 ratio. The ratio of constructs containing linker lengths of 49 (L49; red), 39 (L39; black), 33 (L33; blue), 26 (L26; green), and 21 residues (L21; brown) are shown. Two different samples were analyzed for each construct.

As can be observed in **Figure 13**, the results are quite reproducible taking into account that the two different samples measured for each construct give similar results, except for L26. There is a high discrepancy between L26(7) and L26(8) samples. The expression of this construct should be repeated, but due to time limitations this was not possible.

As had already been noted in **Figure 12**, **Figure 13** shows that the mcpVenus173/mTFP1 ratio is highest in the L39 construct. It seems that in constructs containing linkers of 49 residues (L49) there is considerable arrest, L39 is the construct in which there is least arrest, and with shorter linker lengths there is gradually more arrest. L26(8) breaks this tendency, but I think that the discrepancy in this sample was caused by an experimental mistake during the procedure.

5. DISCUSSION

The use of arrest peptides (AP) is a relatively new technique (Ismail et. al., 2012) that overcomes the limitations of *in vitro* methods, since it allows the study of cotranslational folding *in vivo*. In this work, AP were employed to study the cotranslational folding of the cytosolic AB helices of the potassium channel Kv7.2. Mutations in this channel have been related to serious epileptic diseases. Some of them could be caused by a pathogenic cotranslational folding of the channel, as it has been hypothesized for the W344R A helix mutant.

In this work, constructs were used in which the AB helices are separated from the arrest peptide by a linker of variable length, because the distance between the point in which folding takes place and the arrest peptide is a crucial parameter to study the cotranslational folding. In addition, in order to detect the full-length and truncated protein, fluorescent proteins mTFP1 and mcpVenus173 were placed in the N-terminus and C-terminus respectively. The presence of mcpVenus173 will indicate the production of full-length protein, which means that the protein has folded cotranslationally, relieving the stall produced by the arrest peptide.

Once the constructs were successfully cloned in a pET vector, which allows a better regulation of protein synthesis, the expression of the different proteins was analyzed by SDS-PAGE. **Figure 11** shows that both full-length and truncated proteins are somehow hydrolyzed into smaller peptides. The cause of this fracture is unclear. It could happen during translation or after translation, due to sample manipulation.

As it was expected, the fluorophores permit a good detection of the full-length and truncated protein by electrophoresis (**Figure 11**). Most of the protein was synthesized as the truncated form, suggesting that in the tested conditions most of the protein does not cotranslationally fold in the ribosome vestibule and does not generate enough force to relieve the stall. The fluorometric assays support this hypothesis, as in all the constructs the fluorescence emitted when exciting mcpVenus173 was much lower than the one emitted when exciting mTFP1 (**Figure 12**).

The full-length/total protein ratio, calculated as the fluorescence of mcpVenus173/mTFP1, indicates that the linker with which most full-length protein is

produced is 39 residues long (**Figure 13**). According to Voss et. al. (2006), the ribosome exit tunnel can accommodate between 30 and 40 residues. Thus, a possible explanation for this result is that the L39 linker enables helix B to come out of the ribosome at the point in which the arrest peptide interacts with the ribosome exit tunnel, allowing the AB helices to slightly fold and relieve some of the arrest.

The L39 linker length being the one which most favors cotranslational folding is perfectly concordant with recent literature. Goldman et. al. (2015) describe that in the cotranslational folding of Top7, linker lengths of 31-38 residues are the ones which produce most full-length protein, which is quite close to the 39 residue-long linker reported in this work. Even more interestingly, Farias-Rico et. al. (2018) analyze the cotranslational folding of proteins of different size. They conclude that the protein size is crucial for the impact of the linker length. From the tested proteins, Calmodulin EF2-EF3 is the one which is closest in length to the AB helices (60 residues of Calmodulin EF2-EF3 compared to 70 residues of the AB helices). In that case, the linker length of 40 residues proved to be the one which produced most full-length protein, again very close to the 39 residues indicated in this project.

However, in this work, the observed differences between linker lengths are too low to reach an absolute conclusion about which linker length allows the most efficient cotranslational folding. Indeed, the full-length protein fraction is very scarce, which could contribute to making it difficult to detect the differences in linker length. There can be two reasons for the low full-protein production:

1) The employed arrest peptide is probably too efficient. Its interaction with the ribosome exit tunnel might be very strong, hence a poor full-length protein fraction is produced, making it problematic to observe differences between linker lengths.

2) It is possible that the presence of regulator molecules that physiologically interact with the helices will be required to promote the cotranslational folding. Calmodulin is essential for the good functioning of the channel and has been shown to help in channel trafficking to the membrane (Alaimo et. al., 2009). Thus, it would be interesting to study if Calmodulin will help in cotranslational folding as well.

Future steps to continue this work would try to improve those two areas. On the one hand, less efficient arrest peptides might be tried, so that more full-length protein is

produced and differences between linker lengths are more visible. In this direction, Cymer et. al. (2015) showed that there are key residues in the arrest peptide sequence. By mutating those residues, they obtained arrest peptides with different arrest efficiency. Thus, by modifying certain amino acids of the arrest peptide sequence, less efficient arrest peptides could be obtained, which would probably be more useful in this project.

On the other hand, the effect of Kv7.2 regulator molecules, including Calmodulin, on protein expression should be studied in order to determine if they are implicated in the cotranslational folding of the helices.

In any case, the results obtained in this project show the potential of the arrest peptides in the study of *in vivo* cotranslational folding, in conditions much closer to the physiological ones than *in vitro* experiments. This method, combined with the use of fluorescent proteins, permits the study of cotranslational folding without the need to purify the protein, which makes the technique very time and cost-effective. It can be used to study the cotranslational folding of different small protein domains, and thus offers great possibilities to analyze the effect of pathogenic mutations on protein folding.

6. CONCLUSIONS

- The results show that arrest peptides, in combination with fluorescent proteins, are a powerful tool to study cotranslational folding *in vivo*, interestingly without the need to purify the protein.

- In the tested conditions, most of the AB helices do not fold cotranslationally in the ribosome vestibule. Although linker lengths of 39 residues appear to be the most adequate ones to detect protein folding, in future projects, a less stringent arrest peptide or protein co-expression with Calmodulin might be tried to find conditions that favor protein folding.

7. BIBLIOGRAPHY

Alaimo, A., Gómez-Posada, J. C., Aivar, P., Etxeberría, A., Rodriguez-Alfaro, J. A., Areso, P., & Villarroel, A. 2009. Calmodulin activation limits the rate of KCNQ2 K+

channel exit from the endoplasmic reticulum. *Journal of Biological Chemistry*, *284*: 20668-20675.

Bernardo-Seisdedos, G., Nuñez, E., Gomis-Perez, C., Malo, C., Villarroel, Á., & Millet, O. 2018. Structural basis and energy landscape for the Ca2+ gating and calmodulation of the Kv7. 2 K+ channel. *Proceedings of the National Academy of Sciences*, *115*: 2395-2400.

Butkus, M. E., Prundeanu, L. B., & Oliver, D. B. 2003. Translocon "pulling" of nascent SecM controls the duration of its translational pause and secretion-responsive secA regulation. *Journal of bacteriology*, *185*: 6719-6722.

Chung, H. J. 2014. Role of calmodulin in neuronal Kv7/KCNQ potassium channels and epilepsy. *Frontiers in biology*, *9*: 205-215.

Cymer, F., & von Heijne, G. 2013. Cotranslational folding of membrane proteins probed by arrest-peptide–mediated force measurements. *Proceedings of the National Academy of Sciences*, *110*: 14640-14645.

Cymer, F., Hedman, R., Ismail, N., & von Heijne, G. 2015. Exploration of the arrest peptide sequence space reveals arrest-enhanced variants. *Journal of Biological Chemistry*, *290*: 10208-10215.

Farias-Rico, J. A., Selin, F. R., Myronidi, I., Frühauf, M., & Von Heijne, G. 2018. Effects of protein size, thermodynamic stability, and net charge on cotranslational folding on the ribosome. *Proceedings of the National Academy of Sciences*, *115*: 9280-9287.

Goldberg-Stern, H., Kaufmann, R., Kivity, S., Afawi, Z., & Heron, S. 2009. Novel Mutation in KCNQ2 Causing Benign Familial Neonatal Seizures. *Pediatric Neurology*, 41: 367-370.

Goldman, D. H., Kaiser, C. M., Milin, A., Righini, M., Tinoco, I., & Bustamante, C. 2015. Mechanical force releases nascent chain–mediated ribosome arrest in vitro and in vivo. *Science*, *348*: 457-460.

Ismail, N., Hedman, R., Schiller, N., & Von Heijne, G. 2012. A biphasic pulling force acts on transmembrane helices during translocon-mediated membrane integration. *Nature structural & molecular biology*, *19*: 1018.

Madden, T. 2002. The BLAST Sequence Analysis Tool. 2002 Oct 9 [Updated 2003 Aug 13]. The NCBI Handbook [Internet]. Bethesda (MD): National Center for Biotechnology Information (US).

Merck webpage. Available at: https://www.merck.com/ Accessed 10/06/2019.

Nakatogawa, H., & Ito, K. 2001. Secretion monitor, SecM, undergoes self-translation arrest in the cytosol. *Molecular cell*, *7*: 185-192.

Neubauer, B. A., Waldegger, S., Heinzinger, J., Hahn, A., Kurlemann, G., Fiedler, B., Eberhard, F., Muhle, H., Stephani, U., Garkisch, S. & Eeg-Olofsson, O. 2008. KCNQ2 and KCNQ3 mutations contribute to different idiopathic epilepsy syndromes. *Neurology*, *71*: 177-183.

Otto, J. F., Yang, Y., Frankel, W. N., White, H. S., & Wilcox, K. S. 2006. A spontaneous mutation involving Kcnq2 (Kv7. 2) reduces M-current density and spike frequency adaptation in mouse CA1 neurons. *Journal of Neuroscience*, *26*: 2053-2059.

SibEnzyme webpage. Available at: sibenzyme.com/ Accessed 10/06/2019.

Voss, N. R., Gerstein, M., Steitz, T. A., & Moore, P. B. 2006. The geometry of the ribosomal polypeptide exit tunnel. *Journal of molecular biology*, *360*: 893-906.

Weckhuysen, S., Mandelstam, S., Suls, A., Audenaert, D., Deconinck, T., Claes, L. R., Deprez, L., Smets, K., Hristova, D., & Jordanova, A. 2012. KCNQ2 encephalopathy: emerging phenotype of a neonatal epileptic encephalopathy. *Annals of neurology*, *71*: 15-25.

Zhang, J., Pan, X., Yan, K., Sun, S., Gao, N., & Sui, S. F. 2015. Mechanisms of ribosome stalling by SecM at multiple elongation steps. *elife*, *4*: e09684.