AGRICULTURAL UNIVERSITY OF ATHENS
DEPARTMENT OF BIOTECHNOLOGY
NATIONAL HELLENIC RESEARCH FOUNDATION
SYNTHETIC AND ENZYME BIOTECHNOLOGY LABORATORY

POSTGRADUATE STUDIES PROGRAM
SYSTEM'S BIOLOGY

Postgraduate thesis

Evaluation of selected peptide macrocycles as inhibitors of the disease-related aggregation of the protein TDP-43

Vasileia Pindos C. Theocharatou

Supervisor:
Polydefkis Hatzopoulos, professor AUA

ATHENS
2021

AGRICULTURAL UNIVERSITY OF ATHENS
DEPARTMENT OF BIOTECHNOLOGY
NATIONAL HELLENIC RESEARCH FOUNDATION
SYNTHETIC AND ENZYME BIOTECHNOLOGY LABORATORY

Postgraduate thesis

Evaluation of selected peptide macrocycles as inhibitors of the disease-related aggregation of the protein TDP-43

"Αξιολόγηση επιλεγμένων μακροκυκλικών πεπτιδίων ως προς την ικανότητα αναστολής της συσσωμάτωσης της TDP-43 πρωτεΐνης που σχετίζεται με ασθένειες"

Vasileia-Pindos C. Theocharatou

Examination Committee
Polydeukis Hatzopoulos, professor AUA (supervisor)
Georgios Skretas, principal investigator NHRF
Nikolaos Labrou, professor AUA

Evaluation of selected peptide macrocycles as inhibitors of the disease-related aggregation of the protein TDP-43

MSc System's Biology
Department of Biotechnology
Synthetic and Enzyme Biotechnology Laboratory

Summary

TAR DNA binding protein 43 (TDP-43) is a vital and versatile RNA/DNA binding protein which is aggregation prone. Ubiquitinated cytoplasmic inclusions containing TDP-43 and its C-terminal fragments are pathological hallmarks of numerous neurodegenerative diseases including Amyotrophic Lateral Sclerosis (ALS). Specifically, more than 97% of patients with ALS present TDP-43 protein deposited in neuronal cytoplasmic inclusions suggesting its pivotal role in ALS pathology. The variety of TDP-43 related neurodegenerative diseases, indicate that TDP-43 has a more wide-spread and vital role in the general process of neuro-degeneration.
The main objective of this thesis, was the evaluation of the ability of selected pentapeptides to inhibit the aggregation of TDP-43 *in vitro*. To do so, TDP-43 protein with or without fusion with GFP was cloned and overexpressed in BL21(DE3) cells. TDP-43 protein was subsequently purified by nickel-nitrilotriacetic acid (Ni-Nta) affinity chromatography and tandem size exclusion chromatography (SEC). Aggregation assays, turbidity and sedimentation assays were performed in the presence or absence of two selected cyclic pentapeptides in order to monitor any changes in TDP-43 aggregation propensity. These selected pentapeptides were evaluated for their ability to inhibit the aggregation of TDP-43 protein and thus, preventing the formation of its neurotoxic inclusions that contribute to the degeneration of motor neurons and development of ALS. TDP-43 may hold the key to finding common therapeutics, applicable to a multitude of neurodegenerative diseases. It was shown that TDP-43 is indeed aggregation prone, and its Q331K mutant fused with GFP is more rapidly aggregated than wild type (WT). Furthermore, as observed, there was no inhibition of aggregation by the two cyclic pentapeptides selected, TAFDR and TTYAR.

Scientific area: Synthetic biotechnology

Key words: ALS, TDP-43, neurodegenerative diseases

Αξιολόγηση επιλεγμένων μακροκυκλικών πεπτιδίων ως προς την ικανότητα αναστολής της συσσωμάτωσης της TDP-43 πρωτεΐνης που σχετίζεται με ασθένειες

ΠΜΣ Βιολογίας Συστημάτων
Τμήμα Βιοτεχνολογίας
Εργαστήριο Ενζυμικής και Συνθετικής Βιοτεχνολογίας

Περίληψη

Η TAR πρωτεΐνη δέσμευσης DNA 43 (TDP-43) είναι μία πολυδιάστατη πρωτεΐνη ζωτικής σημασίας η οποία παρουσιάζει τάση συσσωμάτωσης. Ουβικουιτινιωμένα κυτταροπλασματικά συσσωματώματα που περιέχουν την πρωτεΐνη TDP-43 και τα καρβοξυτελικά της θραύσματα αποτελούν σήμα κατατεθέν πολλών νευροεκφυλιστικών ασθενειών συμπεριλαμβανομένης της Πλάγιας Μυοτροφικής Σκλήρυνσης (ALS). Συγκεκριμένα, πάνω από το 97% των ασθενών με ALS έχουν κυτταροπλασματικά συσσωματώματα στους νευρώνες τους που περιέχουν συσσωματωμένη TDP-43 πρωτεΐνη, πράγμα που φανερώνει τον καθοριστικό της ρόλο στην παθολογία της ALS. Η πληθώρα των νευροεκφυλιστικών ασθενειών που συνδέονται με την TDP-43, δείχνουν πως η TDP-43 έχει έναν διαδεδομένο και καθοριστικό ρόλο στην γενικότερη διαδικασία του νευροεκφυλισμού.

Κύριος στόχος αυτής της διπλωματικής, ήταν η αξιολόγηση της ικανότητας επιλεγμένων πενταπεπτιδίων στην αναστολή της συσσωμάτωσης της TDP-43 *in-vitro*. Προκειμένου να γίνει αυτό η πρωτεΐνη TDP-43 με ή χωρίς τη χίμαιρα με GFP κλωνοποιήθηκε και υπερεκφράστηκε σε κύτταρα BL21(DE3). Έπειτα, η πρωτεΐνη TDP-43 απομονώθηκε με χρωματογραφία συγγένειας Νίτρο-νιτριλοτριοξιού οξέος (Ni-Nta) και χρωματογραφία μοριακής διήθησης (SEC). Δοκιμές συσσωμάτωσης: θολερότητα και δοκιμές καθίζησης, έγιναν υπό την παρουσία και την απουσία δύο κυκλικών πενταπεπτιδίων ώστε να ελέγξουν τις μεταβολές στην τάση συσσωμάτωσης της TDP-43. Αυτά τα επιλεγμένα πενταπεπτίδια αξιολογήθηκαν για την ικανότητά τους να αναστείλουν τη συσσωμάτωση της TDP-43 πρωτεΐνης και έτσι να αποτρέψουν τον εκφυλισμό των κινητικών νευρώνων και την ανάπτυξη του ALS. Η TDP-43 μπορεί να αποτελεί το κλειδί στην ανακάλυψη κοινών θεραπειών που να μπορούν να εφαρμοστούν σε διαφορετικές νευροεκφυλιστικές ασθένειες. Αποδείχθηκε πως είναι πράγματι μία πρωτεΐνη που έχει την τάση να συσσωματώνεται και η Q331K μεταλλαγή της μαζί με την GFP συσσωματώνεται ακόμη γρηγορότερα από την αγρίου τύπου πρωτεΐνη (WT). Επίσης, παρατηρήθηκε πως δεν υπάρχει αναστολή της συσσωμάτωσης από τα δύο πενταπεπτίδια που επιλέχθηκαν, το TAFDR και το TTYAR.

Επιστημονικό πεδίο: Συνθετική Βιοτεχνολογία

Λέξεις κλειδιά: ALS, TDP-43, νευροεκφυλιστικές ασθένειες

Acknowledgements

This master thesis experiments were carried out at National Hellenic Research Foundation, at Synthetic and Enzyme Biotechnology Laboratory with Research Associate Professor Georgios Skretas as supervisor and Dr. Maria Tsekrekou as co-supervisor. My master was carried out at Agricultural University of Athens, Department of Biotechnology where my supervisor was Professor Polydefkis Chatzopoulos.

I would really like to thank Mr. Georgios Skretas for doing me the honor to become a member of his laboratory team, for his significant help, expertise and advice.

I am grateful to Mrs. Maria Tsekrekou for her substantial guiding, sharing of knowledge, ethical support, and patience from the first to the latest day of experiments.

I would also, like to thank Mr. Polydeukis Chatzopoulos for introducing me to the world of molecular biology and for supporting and advising me during my master.

My thanks and appreciation to the laboratory members who were willingly there for me when needed and made it a memorable experience.

To my graduate thesis committee Professor Nikolaos Labrou, I would like to thank him for reading my thesis and challenging my knowledge.

Finally, I am thankful to my family and to all my friends for their valuable presence and precious encouragement to fulfill my dreams and achieve my goals.

Contents

1. Abbreviations

- TDP-43: TAR DNA-binding protein-43 kDa
- ALS: Amyotrophic Lateral Sclerosis
- FTD: Frontotemporal Dementia
- FTLD: Frontotemporal Lobar Degeneration
- LB: Luria-Bertani broth
- min: minutes
- h: hours
- WT: wild type
- MW: molecular weight
- ATc: Anhydrotetracycline
- IPTG: Isopropyl β-D-1-thiogalactopyranoside
- His: Histidine
- GFP: Green fluorescent protein
- E. coli: Escherichia coli
- OD: Optical Density
- MisP: Misfolded Protein
- PBS: Phosphate Buffered Saline
- Kan: Kanamycin
- Amp: Ampicillin
- Cm: Chloramphenicol
- HEPES: 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
- Ni-Nta: Nickel-nitrilotriacetic acid
- SDS-PAGE: sodium dodecyl sulphate–polyacrylamide gel electrophoresis

- APS: Ammonium persulfate
- TEMED: Tetramethylethylenediamine
- TAE: Tris-acetate-EDTA
- EDTA: Ethylenediaminetetraacetic acid
- DTT: Dithiothreitol
- PEG: Polyethylene glycol
- DMSO: Dimethyl sulfoxide
- SEC: Size Exclusion Chromatography
- dH_2O: distilled H_2O
- ddH_2O: double distilled H_2O
- HRP: Horseradish peroxidase
- TBS: Tris-buffered saline
- PMSF: Phenylmethylsulfonyl fluoride
- PBS-T20: Phosphate-buffered saline Tween 20
- Acryl: Acrylamide
- Bis: Bis(dimethylamino)phosphorodiamidate

2. Introduction

The misfolding of protein is causal for a variety of human diseases referred to as protein misfolding diseases (PMD) [1].

One of these protein misfolding diseases is the Amyotrophic Lateral Sclerosis (ALS), the most common motor neuron disease of adult' onset. ALS is characterized by loss of upper and lower motor neurons and leads to death from respiratory failure within 3-5 years of symptoms onset [2]. The average age of onset of the disease is ~60 years [3]. The disorder is generally sporadic (90-95%) and only approximately 5% of cases have a positive family history for ALS [4].

2.1 TDP-43

As it is shown in Neumann et al. 2006, in the frontotemporal cortex and hippocampus of the brain and in the motor neurons of the spinal cord of patients suffering from amyotrophic lateral sclerosis (ALS), there is an accumulation of ubiquitin-positive, tau-negative and a-synuclein-negative protein inclusions [5]. In 2006, the main component of these inclusions was reported to be TDP-43 protein [6]. Moreover, it is shown that the vast majority of ALS cases involve TDP-43's aggregation [7]. Specifically, almost 97% of ALS patients have been found with TDP-43 aggregation.

The immunoreactive inclusions of TDP-43 have been also observed in a broad spectrum of other neurodegenerative disorders including frontotemporal dementia (FTD), ALS/parkinsonism–dementia complex of Guam, Alzheimer's disease (AD), dementia with Lewy bodies (DLB), Pick's disease, argyrophilic grain disease and corticobasal degeneration. These conditions are collectively termed as 'TDP-43 proteinopathies' [8]. Thus, there is a more wide-spread and vital role of TDP-43 in the general process of neuro-degeneration. This could lead to the hypothesis that by targeting the TDP-43 dyshomeostasis, common therapeutics, applicable to a multitude of neurodegenerative diseases could be found [9].

TDP-43 was first identified as a protein that bound to the transactivation response (TAR) element of HIV human immunodeficiency virus and was named TAR DNA-binding protein-43 kDa [10]. It is a highly conserved and an essential DNA and RNA binding protein [11] with 414 amino acids encoded by the TAR DNA binding protein gene on chromosome 1p36 [12] .

TDP-43 is a multifunctional nuclear protein which shuttles between nucleus and cytoplasm and is a part of the heterogeneous nuclear ribonucleoprotein (hnRNP) family which also includes some of the widely known splicing modulators, such as hnRNP I, hnRNP A/B, and hnRNP H [13],[14]. Its shuttling between these two compartments is explained by the fact that it is a protein that controls gene expression through RNA processing. In particular, it is involved in regulation of RNA splicing, microRNA processing, mRNA transport, stability and translation [15]. There are indications that TDP-43 controls expression of various genes and is dynamically distributed in cells.

Dynamic light scattering and SEC analysis have suggested that TDP-43 is a dimer [16]. It is comprised of the N-ubiquitin-like domain, the nuclear localization signal (L), two RNA recognition motifs (RRM1 and RRM2) including a nuclear export signal (E), and a C-terminal Q/N/S/G-rich domain (Prion-like domain) (Fig.1). Nuclear localization and export signal are responsible for TDP-43 shuttling between the nucleus and the cytoplasm. RNA recognition motifs bind to single- or double-stranded RNA/DNA [13] at UG-/TG- repeat. The RRMs are crucial for RNA metabolism as they are present in numerous proteins and affect RNA processing and transport [17][18]. Interestingly, the C-terminal domain over residues 274-414 has a sequence low in complexity and abundant in Gln, Asn, Ser and Gly residues, which is 24.2% identical with the N-terminal yeast prion domain of Sup35 [13]. The C-terminal domain of TDP-43 is crucial for spontaneous aggregation. Mutations within this domain (ex. Q331K, M337V) are linked with ALS, accelerate aggregation of pure TDP-*43 in vitro and* increase the number of TDP-43 aggregates and promote toxicity *in vivo* [19].

Figure 1. TDP-43 Domains. N-ubiquitin-like domain, nuclear localization signal (L), two RNA recognition motifs (RRM1 and RRM2) hosting a nuclear export signal (E), an C-terminal prion like domain abundant in Gln/Asn/Ser/Gly residues [13].

The mutations studied in our experiments are the following:

- G294A, where glycine at position 294 is replaced by alanine. Its expression in SH-SY5Y cells promoted aggregation. A synthesized peptide with this mutation in a phosphate buffer also forms fibrils which are similar to neuronal cytoplasmic inclusions [20][21].
- Q331K, where glutamine at position 331 is replaced by lysine. It was shown in a mouse transgenic model, that this ALS-linked mutant produces loss and gain of splicing function of selected RNA targets compared with WT protein at the early stages of disease [22]. In mature mouse cortical neurons, this mutation led to increased mislocalization to the cytoplasm and compared to wild type impaired neurite formation [23]. Mice expressing this mutant have an age-dependent mild motor pathological phenotype increased compared to WT [24].
- M337V, where methionine at position 337 is replaced by valine. Transgenic mice which express this mutant develop progressive motor dysfunction and early death [25]. In a mouse transgenic model, it was shown that ALS-linked M337V mutant also produces loss and gain of splicing function of selected RNA targets compared with WT protein [22]. In mature mouse cortical neurons, this mutation milsocalized to the cytoplasm and impaired neurite formation compared with wild type [23].

Wild type (WT) TDP-43 has been shown to be intrinsically aggregation-prone [19]. Experiments in cervical carcinoma cells (HeLa) showed that under specific environmental conditions such as heat, viral infection, oxidative stress, or

hypoxia, TDP-43 tends to be redistributed from the nucleus to the cytoplasm [6] in cytoplasmic stress granules [26] that may lead to pathological aggregate formation [27]. Experiments in post-mortem frontal cortical tissue revealed that in these cytoplasmic inclusions TDP-43 is hyperphosphorylated, ubiquitinated and cleaved to form C-terminal fragments [6],[28]. However, in spinal cord neurons the inclusions consist of full length hyperphosphorylated and ubiquitinated TDP-43 [29].

In mouse embryos, homozygous loss of TDP-43 protein was found to be lethal, while heterozygous loss of TDP-43 protein leads to deficits in motor function. These findings demonstrate an absolute requirement for TDP-43 in mice embryogenesis, which is not surprising as TDP-43 protein participates in numerous fundamental cellular functions [30]. Complete TDP-43 loss of function is probably not tolerated in humans either. However, ALS linked mutations are not lethal in humans for embryos but become dysfunctional with ageing [31].

Although the loss of normal nuclear localization and cytoplasmic aggregation correlate with neurodegeneration, the exact mechanisms of neurotoxicity still remain unknown.

2.2 Drug discovery approaches for protein misfolding diseases

As the most of the protein misfolded diseases remain untreatable, new molecules with therapeutic potential of protein misfolding are in demand. Chemical rescuers of protein misfolding are a very promising approach for the misfolding protein diseases treatment. These are typically identified via:

- structure-based drug design. For this procedure the first step is the choice of a target, the determination of its structure, the selection of the appropriate method for the drug lead discovery that have the ability to reach the target and the evaluation of the drug leads [32].
- synthetic or natural chemical library screening using biochemical or biophysical assays to detect protein misfolding (such as Thioflavin T staining, microscopy, spectroscopy techniques).
- high-throughput screening of DNA-encoded chemical libraries via display technologies. High throughput screening is the rapid test of thousands to millions of samples biological activity with the use of automated equipment [33]. The main display technology used is virus/phage display but there are others too, such as cell display, ribosomal display, mRNA display and covalent DNA display [34].

Nevertheless, structural information is available for only a few pathogenic misfolded proteins, and screens (biochemical and biophysical) are very expensive, time consuming and generally with a low throughput. Furthermore, display technologies require available purified misfolded proteins in specific forms and do not provide information for misfolding-rescuing activity of tested molecules.

Microbial genetic screens that have the ability to perform rapidly a high-throughput screening have emerged as useful tools for the discovery of possible

anti-protein misfolding disease's therapeutics. Moreover, due to advances in the field of synthetic biology the microbial production of libraries of different types of peptides recently became possible. Thus, via the combination of these microbial genetic screens with the libraries produced, there is the ability to overcome the difficulties of the methods referred to above in order to find therapeutic molecules against misfolded protein diseases [1].

2.3 Integrated bacterial system for selection of chemical rescuers for protein misfolding diseases

In this intergrated bacterial system manufactured from our laboratory, the bacterial cells are employed to coexpress the protein of interest fused with GFP together with a library of cyclic peptides. GFP fluorescence is indicative of the conformational status of the protein, thus peptides that increase fluorescence might rescue aggregation.

Peptides, in comparison with other small synthetic molecules, are considered less toxic and are useful and common drugs available in the market.

The cyclic peptides are polypeptides whose one end is linked with the other usually with amide bonds. Their cyclic conformation results in some benefits compared with their linear form. Firstly, their rigidity decreases the entropy term of the Gibbs free energy, therefore allowing the enhanced binding toward target molecules, or receptor selectivity. Secondly, thanks to their lack of termini they are resistant to hydrolysis by exopeptidases and they are also resistant to endopeptidases as their structure is rigid. Furthermore, they penetrate easier through the membrane in comparison to the linear conformations [35].

The cyclic peptide library used in our laboratory was formed by split-intein-cyclization with the 'split-intein-mediated circular ligation of peptides and proteins' (SICLOPPS) technique and had a general structure $NuX_1X_2...X_n$, where Nu is Cys, Thr or Ser while X is any amino acid. Peptides expressed by the pSICLOPPS vector acquire a cyclic conformation [36]. In this way a library that consisted of macrocyclic pentapeptides and was characterized by high levels of chemical and structural diversity was biosynthesized in *E. coli* cells [1].

In these systems the ability of these peptides to rescue pathological misfolding was simultaneously screened with an ultrahighthroughput fluorescence-based genetic assay. Effectiveness was tested in amyloid β peptide which is linked with Alzheimer's disease and a mutant of the human Cu/Zn superoxide dismutase (SOD1) which is linked with a familial form of amyotrophic lateral sclerosis (fALS) [37].

In particular, the protein fused with GFP and the cyclic pentapeptides were co-expressed in *E. coli* cells (Fig.2). It has been shown that due to the fact that Aβ is aggregation prone, overexpression of Aβ42-GFP produces a misfolded fusion that is accumulated in inclusion bodies that lack fluorescence. In conditions that inhibit Aβ aggregation, Aβ42-GFP is soluble and fluorescent. Thus, macrocyclic peptides that have the ability to decrease protein misfolding lead to a fluorescent phenotype [1].

Indeed, macrocyclic peptides that managed to decrease protein misfolding for SOD1-GFP and Aβ42-GFP [1], [37] were found by this method.

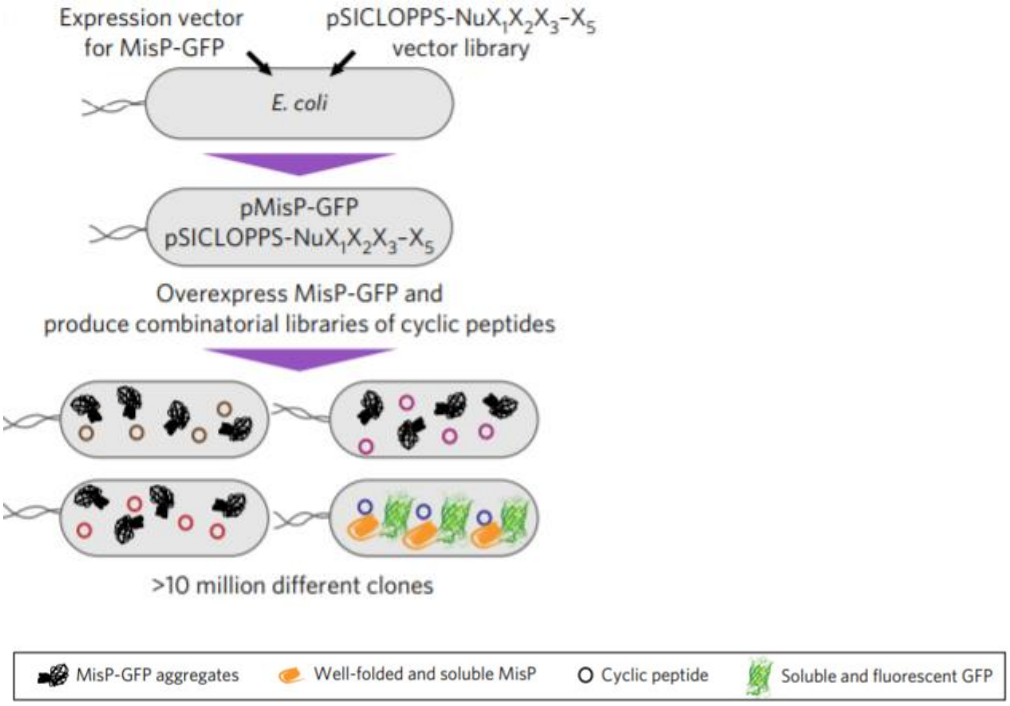

Figure 2: Escherichia coli cells co-expressing the misfolded protein fused with GFP and the combinatorial macrocyclic oligopeptide library cyclo-NuX1X2X3–X5, where Nu is Cys, Ser or Thr and X is any of the 20 natural amino acids. Macrocyclic peptides that inhibit aggregation lead to fluorescent proteins and thus bacteria [1].

2.4 Objective of the thesis:

Amyotrophic Lateral sclerosis is related with TDP-43 protein aggregation. TDP-43 aggregates are also present in numerous other neurodegeneration diseases, thus, the inhibition of TDP-43's aggregation would be advantageous in all these cases. The main objective of this thesis was to investigate the ability of selected macrocyclic pentapeptides to inhibit TDP-43 aggregation *in vitro*. To fulfill this purpose 4 steps should be achieved:

- Cloning of TDP-43 and TDP-43-GFP protein.
- Purification of TDP-43 and TDP-GFP protein via Ni-affinity and size exclusion chromatography.
- Setting up aggregation assays for TDP-43 and TDP-43-GFP as have been described in the literature.
- Monitor TDP-43 aggregation assays in the presence or absence of two selected cyclic peptides.

3. Materials and Methods

3.1 Materials

3.1.1 Strains and antibiotics

During the experiments of this thesis the bacterial strains used were DH5a for molecular cloning and subcloning and Rosetta 2 and BL21(DE3) for bacterial expression of TDP-43 protein. Rosetta 2 were grown in SOC medium, whereas BL21(DE3) and DH5a in LB. The antibiotics used were ampicillin (100 μg/mL), chloramphenicol (40m μg/μL) and kanamycin (50 μg/mL).

3.1.2 Ladders

The ladder used for agarose gel electrophoresis was DNA ladder 1kB NEB while the ladder used for polyacrylamide gel electrophoresis was Novex, Life technologies 10748-010 Benchmark prestained protein ladder.

3.1.3 Proteins and Enzymes

All enzymes used and their buffers -T7 RNA polymerase, Taq polymerase, T4 Ligase, T4 ligase buffer, BamHI, XbaI, Hind III- were from New England Biolabs (NEB).

3.1.4 Antibodies

Anti-His and anti-mouse antibodies HRP conjugated were supplied by Sigma. Anti-Flag antibody was supplied by Cell signalling and Anti-GFP by Takara.

3.1.5 Kits

Plasmid extraction Monarch® Plasmid Miniprep Kit and gel extraction Monarch® DNA Gel Extraction Kit were purchased from NEB.

3.1.6 Solutions for bacterial growth

LB: 10 g/L Tryptone (Bioline), 5 g/L Yeast Extract (Applichem), 10 g/L NaCl (Applichem)

LB agar plates: 10 g/L Tryptone (Bioline), 5 g/L Yeast Extract (Applichem), 10 g/L NaCl (Applichem), 15 g/L agar

SOC: 20 g/L Tryptone (Bioline), 5 g/L Yeast extract, 4.8 g/L $MgSO_4$ (Merck), 3.6 g/L Dextrose (Fluka)

IPTG: 1mM IPTG (Thermofischer) diluted in ddH20

ATc: 20 mg/L anhydrous tetracycline (Sigma) diluted in pure Ethanol

3.1.7 Solutions for PCR

dNTPs: 50 µL of each dNTP (10mM, Bioline) and 300 µL ddH_2O

3.1.8 Solutions for protein quantification

Bradford reagent: Coomassie buffer (Thermo)

3.1.9 Solutions for electrophoresis

Running buffer (10x): 250 mM Tris (Techline), 1.92 M Glycine (Applichem), 1% w/v SDS (Applichem)

Transfer buffer: 25 mM Tris (Techline), 192 mM Glycine, 0.01% w/v SDS, 20% v/v Methanol (Merck)

10x PBS: 1.4 M NaCl, 26 mM KCl (Sigma), 100mM Na_2HPO_4 (Sigma), 18 mM KH_2PO_4 (Sigma)

PBS-Tween-20: PBS with 0.1% Tween-20 (Merck)

TAE (50x): 2 M Tris base (Techline), 1 M Glacial acetic acid (Merck), 0.05 M EDTA (Applichem), pH=8.0 diluted in 1 L dH_2O

Laemmli buffer: 62.5 mM Tris-HCl (Applichem), pH=6.8, 2% SDS (Applichem), 25% glycerol (Merck), 0.01% bromophenol blue (Sigma), 5% β-mercaptoethanol (Roche) added fresh

3.1.10 Solutions for chemically competent bacterial cells

TSS: LB, 10% PEG 8000 (Sigma), 5% DMSO (Applichem), 50 mM $MgCl_2$ (Merck)

3.1.11 Primers

-GS088: TAATACGACTCACTATAGGG

-DD035: GAATGGCCAGATGATTAATTC

-DD037: CACACCCGCCGCGCTTAATGC

3.2 Methods

3.2.1 Bacterial culture conditions

For every type of cell a quantity of 5 mL of LB, supplemented with the appropriate amount of antibiotic -ampicillin (100 µg/mL), chloramphenicol (40 µg/mL) or kanamycin (50 µg/mL)- was inoculated with 10 µL of glycerol bacterial stock or transformation mix. Ampicillin was used for cells containing pASK75 plasmid,

chloramphenicol for Rosetta 2 cells and kanamycin for cells with pET28a plasmid.Liquid cultures were incubated at 37 °C with agitation overnight. Cells were plated in petri dishes containing LB supplemented with antibiotic and incubated at 37° C overnight.

3.2.2 Glycerol stock

1.5 mL of overnight bacterial liquid culture, was pelleted at 11000 rpm for 5 min in Eppendorf tubes. The cell pellet was resuspended in 800 µL LB and 200 µL glycerol 80%. Bacterial cells were stored for short-term in -20°C or for long-term in -80 °C [38].

3.2.3 Chemically competent bacterial cells

Bacterial cells from a glycerol stock were streaked on LB-agar plates and incubated overnight at 37 °C. A single colony was inoculated in 5 mL LB overnight at 37 °C with agitation. 1 mL of overnight culture was added at 100 mL LB (1:100 dilution) and incubated, at 37° C with agitation until optical density (OD) reaches 0.4-0.5. The culture was cooled in an ice bath with shaking. Cells were centrifuged at 4°C, 3500xg for 10 minutes and the cell pellet was resuspended in 5 mL ice cold TSS. Cells were distributed in ice cold Eppendorfs and stored at -80°C. The procedure was conducted under flame sterilizing conditions.

3.2.4 Transformation

1 µL of plasmid vector was added to 50 µL of chemically competent cells and subsequently incubated on ice for 30 min. Cells were subjected to heat shock at 42 °C for 50 s and then, incubated on ice 2-3 min. Transformed cells were recovered at 37 °C for 1 h for transformation with kanamycin or chloramphenicol resistant vectors or for 15 min for ampicillin resistance. Cells were plated on a Petri dish containing LB-agar supplemented with the suitable antibiotic and incubated overnight at 37 °C. Single colonies were used for subsequent analysis or assays (e.g. vector propagation, PCR, etc).

3.2.5 Gel DNA extraction and plasmid extraction

According to manufacturer protocol (NEB).

3.2.6 Molecular cloning

To clone TDP-43 into the pET28a vector, the forward primer created contains a BamHI restriction site and adds a 6xHis tag while the reverse primer adds a Flag-tag and creates a HindIII restriction site.

Phusion PCR

1 µL of plasmid DNA was mixed with the master mix in an Eppendorf tube. Master mix contained 10 µL of 5x polymerase buffer, 1 µL dNTPs, 1 µL of forward primer (MT011), 1 µL of reverse primer (MT012), 0.3 µL of Phusion polymerase and double distilled water to a final volume of 50 µL.

MT011 forward TDP-43 primer:
BamHI – RBS – start - 6xHis – cDNA (TDP-43):
aaaaa ggatcc aggaggaaacg atg caccatcaccatcaccat tctgaatatattcgggtaacc

MT012 reverse TDP-43 primer:
HindIII – Stop – Flag – cDNA (TDP-43):
aaaaa aagctt cta cttgtcatcgtcgtccttgtagtc cattccccagccagaagactta

The program of phusion PCR with the primers mentioned had the following steps:

1.	95°C	3min	
2.	5°C	30sec	
3.	T_m(gene)	30sec	
4.	72°C	1min	Step 2 to step 4 for 5 cycles
5.	95°C	30sec	
6.	T_m(whole)**+3°	30sec	
7.	72°C	1min	Step 5 to step 7 for 30 cycles
8.	72°C	5min	
9.	4°C(or 10°C)	for ever	

*T_m(gene): T_m of cDNA part of the primer with the lower T_m. In our case 52°C.

**T_m(whole): T_m of the primer with the lower T_m. In our case 72°C.

The PCR product (insert) and the vector are digested with the same restriction enzymes (BamHI and HindIII).

Digestion

5 μL of plasmid DNA were mixed with 4 μL of Enzyme buffer 10x, 1 μL of restriction enzyme 1, 1 μL of restriction enzyme 2 and 29 μL ddH_20. PCR products digestion: 10 μL of DNA were mixed with 4 μL of enzyme buffer, 1 μL of restriction enzyme 1, 1 μL of restriction enzyme 2 and 24 μL ddH_20.

Samples were vortexed and incubated at 37 °C without agitation. Successful digestion was verified via agarose gel electrophoresis and the DNA is purified via gel extraction.

Purified DNA was quantified in order to have the right ratio of insert:vector for ligation (3:1).

Ligation

The insert and the recipient vector were merged via ligase enzyme in a procedure called ligation. Prerequisite was that they should have been cut with the same restriction enzymes and thus they should have corresponding sticky edges. Vector and insert were mixed in a molecular ratio 1:3. Also, 1 μL of T4 ligase (Bioline), 1 μL of T4 ligase buffer (Bioline) and ddH_20 to 10 μL were added.

Samples were incubated at room temperature for 30 min. After ligation the new plasmids were transformed in 100 μL of chemically competent DH5a cells, cultivated at plates and were incubated overnight at 37 °C. The following day, a colony PCR was performed in order to continue with colonies most likely to have the right insert. PCR positive colonies were incubated overnight in a liquid culture at 37 °C with agitation. Three colonies of WT, G294A, Q331K mutants and six colonies of M337V mutant were selected for colony PCR performed with vector-specific (GS088-forward and DD037-reverse) primers. These primers were used in order to verify that *TDP-43* is inserted with the right direction.

Colony PCR

A colony was picked and placed in an Eppendorf with 15 μL ddiH_20 and 3 μL were mixed with 17 μL of master mix containing: 2μL Thermo Pol buffer (NEB), 0.4 μL dNTPs (Bioline) 10 mM, 0.4 μL MT011 forward primer 100 pmol/μL, 0.4 μL MT012 reverse primer pmol/μL (mentioned above), 0.1 μL Taq polymerase (Biolabs) and 13.7 μL ddH_20.

The program of colony PCR had the following steps:

1. 95°C	3min	
2. 5°C	30s	
3. T_m *	30s	
4. 68°C	1min	Step 2 to step 4 for 30 cycles
5. 68°C	5min	
6. 4°C(or 10°C)	for ever	

*T_m: The lower T_m between the T_m of primers.

Plasmids were extracted (Monarch plasmid extraction kit - NEB) according to the manufacturer's protocol, tested again with a second diagnostic digestion with HindIII and XbaI restriction enzymes. Finally, products were evaluated with diagnostic PCR with pET28a forward and reverse primers - more details in results section-. Finally, the right ones were sequenced and were transformed at BL21(DE3) cells for protein expression.

3.2.7 Subcloning

Subcloning is a procedure in molecular biology for transfer of DNA inserts from one vector to another. This procedure contains the following steps: digestion of the recipient plasmid and of donor plasmid, ligation and finally transformation to the appropriate strain. To verify the product of molecular cloning, a diagnostic PCR and a diagnostic digestion are performed. Firstly, donor and recipient plasmid were digested with the enzymes XbaI and BamHI and gel extraction (Monarch DNA gel extraction kit- NEB) was performed in order to separate the digested product. The DNA fragments of interest were quantified in order to set up a ligation reaction with a 3:1 molar ratio of insert:vector. When ligation was completed, the new plasmid was transformed in DH5a cells and cells were cultivated on plates and incubated overnight at 37°C. The following day, a colony PCR was performed with 3 colonies in order to continue the procedure with the

colonies most likely to have the right insert. The primers used for colony PCR were DD035 forward and DD037 reverse primer. Positive colonies of colony PCR were incubated overnight in liquid culture at 37 °C with agitation. Plasmids extracted (Monarch plasmid extraction kit - NEB) were tested again with a second diagnostic digestion with EcoRI double cutter restriction enzyme. Next verification step was with diagnostic PCR with forward pASK75 primer MT013 and the reverse primer DD037 -more details in result section-. Finally, the right ones were transformed at BL21(DE3) cells for protein expression.

MT013 pASK75 forward primer:
XbaI – RBS – start - 6xHis – cDNA (TDP-43):

aaaaa tctaga aggaggaaacg atg caccatcaccatcaccat tctgaatatattcgggtaacc

*The procedures of subcloning are presented analytically in the molecular cloning section.

3.2.8 Agarose gel

1.2 % agarose diluted in 100 mLTAE buffer with 8 μL of ethidium bromide.

3.2.9 Induction of protein production

E. coli BL21(DE3) cells were transformed with the appropriate vector that enabled the production of the protein of interest. Transformed *E. coli* was inoculated in liquid LB cultures supplemented with the appropriate antibiotic and cells were grown for ~ 16h at *37 °C* with agitation. The following day a subculture with OD_{600}=0.1 was prepared in fresh LB supplemented with the appropriate antibiotic and was incubated at 37 °C with shaking until OD_{600}=0.6-0.8.Then protein production was induced by adding 1mM IPTG or 0.2 μg/mL ATc (IPTG was used for pET28a plasmids whereas ATc for pASK75 plasmids) and cells were incubated for 20 h at 20 °C with constant shaking. OD_{600} was measured and cells were centrifuged at 6000 g at 4°C for 10 min. Finally, cells were resuspended in PBS to a final OD_{600}= 5 for subsequent analysis.

3.2.10 Co-induction

The pASK75 plasmid with the TDP-43 protein fused with GFP and the pARCBD plasmid with SICLOPPS technology containing the selected pentapeptides were transformed at E. coli cells. Liquid culture of cells containing the appropriate antibiotics (ampicillin for pASK75 and chloramphenicol for pARCBD) were shaken for ~16 h. The following day in a subculture with OD=0.1 (fresh LB supplement with the appropriate antibiotics) 50 μL 0.02% Ara were added for the induction of pARCBD vector. The culture was incubated at 37 °C with shaking until OD_{600}=0.3-0.5. Misfolded protein fused with GFP was induced with 0.2 μg/mL ATc and cultures were incubated overnight at 18 °C.

3.2.11 Cell lysis via sonication

1 mL cell culture with OD_{600} =5 was harvested by centrifugation at 11000 rpm for 2 min. Cell pellets were stored at – 20 °C or -80 °C until protein extraction. Cell pellets were resuspended in 100 µL PBS and sonicated in an ice bath for 5 cycles (30 s on/30s off). Total protein extract was centrifuged at 200xg for 5 min to remove cellular debris. An aliquot of 40 µL of supernatant was transferred at a new eppendorf (Total) and 20 µL of 6x Laemmli buffer were added. The remaining supernatant was further fractionated by centrifugation at 21000xg, 4 °C for 30 min. The supernatant was transferred to a new tube (Soluble, ~40 µL) and 20 µL of 6x Laemmli buffer were added. The remaining pellet (insoluble fraction) was resuspended with 30 µL 6x Laemmli buffer and 30 µL ddH_20 were added.

3.2.12 Solubilization of inclusion bodies

BL21(DE3) cells were transformed with a vector containing the gene coding the TDP-43 protein and were incubated overnight at 37 °C with shaking. The following day a subculture was prepared with 2.5 mL of liquid culture, 250 mL LB and 250 µL antibiotic at 37°C with agitation until OD_{600}=0.6. Cells were induced with 1mM IPTG or 0,2 µg/mL ATc (IPTG was used for pET28a plasmids whereas ATc for pASK75 plasmids) and were incubated at 20°C for 20 h. Cultures were centrifuged at 4000xg at 4°C for 10 min. Cell pellets were resuspended with 10 mL of 50 mM Tris, pH=8.0, 0.5M NaCl, 250 µL of lysozyme, 25 µL of DNAse and 100 µL PMSF protease inhibitor and were incubated on ice with constant stirring for 30 min. Cells were sonicated for 6 cycles of 20 s each. Lysates were centrifuged at 24000xg for 30 min to separate insoluble inclusion bodies from soluble protein fraction.

Isolation of inclusion bodies from the insoluble fraction: The pellet was subjected to a series of washes: (1) with 50mM HEPES, pH=8.0 and 1%TritonX-100 (2) with 50mM HEPES, pH=8.0 and 2 M NaCl and (3) 50mM HEPES, pH=8.0. The pellet was centrifugated at 24000xg for 30 min at 4 °C between washes and after the final wash.

Solubilization of inclusion bodies: The resulting IBs were dissolved into 25 mL of denaturing buffer containing: 50 mM HEPES, pH-8.0, 0.3 M NaCl, 10 mM imidazole and 8 M urea with constant stirring at room temperature for at least 15 h. The next day, the solubilized IBs were centrifuged at 24000g for 30 min at 4 °C to remove any debris and were stored at -80°C. Samples from total, soluble and insoluble fraction were mixed with 6x Laemmli buffer and were tested with SDS-PAGE.

3.2.13 Protein quantification

5 different known concentrations of bovine serum albumin (BSA) protein (standards) were diluted 1:50 in Coomassie buffer (Bradford reagent). Water was used as a blank for standards. The final concentrations of standards were:

Standards	**1:50 dilution**	**final standards**
mg/mL		**µg/mL**
0.1		**2**
0.25		**5**
0.5		**10**
0.75		**15**
1		**20**

Samples were diluted from 1:50-1:500 depending on the experiment. As a blank for samples, the buffer they were diluted in was used. 200 µL were loaded in the microplate wells and absorbance was detected at 600 nm.

3.2.14 Polyacrylamide gel

10% resolving gel:	stacking gel:
3.35 mL 30% Acryl:Bis (29:1)	1.35 mL 30% Acryl:Bis (29:1)
2.5 mL 1.5 mM Tris, pH=8.8	2. 5 mL 0.5 mM Tris, pH=6.8
4.15 mL H_20	6.15 mL H_20
100 µL 10% SDS	100 µL 10% SDS
100 µL 10% APS	100 µL 10% APS
5 µL TEMED	10 µL TEMED

3.2.15 Coomassie Brilliant Blue stain

SDS-page gels were stained with Coomassie Brilliant Blue buffer for 45 minutes. Then they were rinsed with water, washed for 45 minutes with Coomassie destaining buffer and overnight with water.

3.2.16 Ponceau S stain

After western blot, membranes were stained with ponceau S for 2 minutes subsequently, rinsed with water to remove unspecifically bound stain and an

image was captured. Membranes were destained with PBS-Tween-20 before subsequent blocking.

3.2.17 Concentration with Amicon cells

Samples were loaded in Amicon cells and centrifuged at 3000 g, 4 ºC until 500 μL or 1 mL final volume.

3.2.18 Ni-NTA gravity flow affinity chromatography

Ni-NTA chromatography is an immobilized metal-affinity chromatography used for purification of recombinant proteins containing a polyhistidine tag. The column is covered with an immobilized metal ion Ni^{2+} and the recombinant protein with the affinity tag of polyhistidine sequence forms strong interactions with Ni^{2+}. By addition of buffer with free imidazole the protein with polyhistidine residues is eluted [39].

6 mL of Ni slurry (equivalent with 3 mL Ni beads) were loaded to a column with 25 mL volume. Column was washed with 30 mL H_2O and 30 mL Binding buffer (50 mM HEPES, pH=8, 0.5 M NaCl, 8 M urea and 10 mM imidazole for inclusion bodies and 50 mM Tris, pH=8, 0.5 M NaCl, and 10 mM imidazole for soluble fraction). The sample was loaded on the column and the beads with the sample were rotated for 1 hour at 4 °C and transferred back to the column. Flow through was collected. Two washes followed with Wash buffer (50 mM HEPES, pH=8, 0.5 M NaCl, 8 M urea and 25 mM imidazole for inclusion bodies and 50 mM Tris, pH=8, 0.5 M NaCl, and 25 mM imidazole for soluble fraction) with rotation for 30 min at 4 °C each. Wash solution was collected. Protein was eluted by adding 5-10 mL Elution Buffer (50 mM HEPES, pH=8, 0.5 M NaCl, 8 M urea and 300 mM imidazole for inclusion bodies and 50 mM Tris, pH=8, 0.5 M NaCl, and 300 mM imidazole for soluble fraction). The column was washed with ~20 mL H_2O and Ni-beads in 1:1 20% Ethanol, in order to be stored.

3.2.19 Ni-Nta His-Trap AKTA

Nickel-nitrilotriacetic acid (Ni-Nta) AKTA column preparation for samples started with a wash of 25 mL of ddH2O, with 25 mL of 300 mM imidazole and with 50 mL of equilibration buffer containing 50 mM HEPES, pH=8, 0.5 M NaCl, 8 M urea and 10 mM imidazole. In order to purify the protein, the sample was loaded, and washed with 50 mM HEPES, pH=8, 0.5 M NaCl, 8 M urea and 25 mM imidazole. The protein was eluted with 50 mM HEPES, pH=8, 0.5 M NaCl, 8 M urea and 300 mM imidazole. Between different samples the column was washed with 25 mL 500 mM imidazole and subsequently equilibrated with 50 mL equilibration buffer. Finally, a final wash with 25 mL 500 mM imidazole and 50 mL equilibration buffer followed and column was stored with 20% EtOH.

3.2.20 Size exclusion chromatography

Size-exclusion chromatography (SEC) is a chromatographic technique used for separating substances according to their molecular size as they pass through a bed of porous particles [40]. At SEC the first eluates are the ones with the bigger volume.

Protein samples eluted from Ni-Nta chromatography were concentrated to 600 μL with Amicon cells and filtered through 20 μm syringe filter. All the buffers used were degassed. 500 μL (2.2 mg/mL) of soluble WT protein was loaded. 500 μL (2.8 mg/mL) of solubilized WT IBs were loaded.

Samples were loaded onto a HiLoad 16/600 Superdex 200 column driven by an Akta Pure system at a flow rate of 0.8 mL/min. After the sample was eluted with the elution buffer: 50 mM HEPES and 0.2 M KCl. The first samples eluted were those with the larger molecular size as they are not trapped into the porous particles of the column.

The HiLoad 16/600 Superdex 200 prep grade manual refers to a number of proteins with known molecular weight and volume of elution, which were used as standards. These data may give an insight at the oligomerization state of the eluates according to their volume of elution. In the experiments presented in HiLoad 16/600 Superdex 200 prep grade manual the time of elution of each molecular weight was known as mentioned, as well as the flow rate which was 1.5 mL/ min. Thus, the volume of the elution as well as the log (MW) were calculated. The diagram of log(MW)=f (volume) is also presented (Diagram 1).

time (min)	volume	MW	log(MW)
50	75	440	2.6434527
65	97.5	158	2.1986571
75	112.5	67	1.8260748
82	123	43	1.6334685
95	142.5	17	1.2304489

Table 1: Calculation of volume of elution at HiLoad 16/600 Superdex 200 prep grade column and of log(MW) when time of elution, molecular weight and flow rate=1.5 mL/min is known.

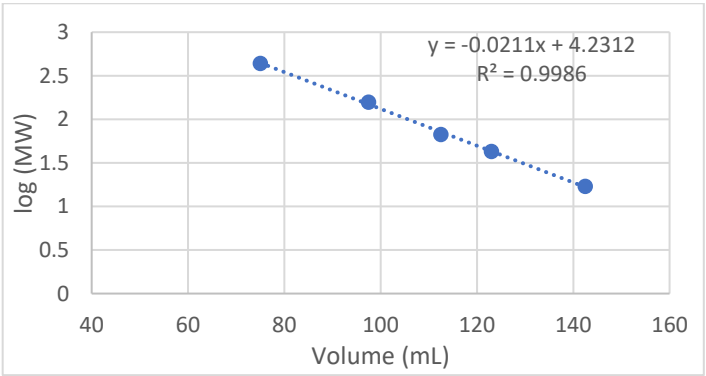

Diagram 1: The log(MW)=f (volume) diagram of HiLoad 16/600 Superdex 200 prep grade column.

3.2.21 Turbidity assay

The insoluble components in a liquid environment absorb and therefore increase the overall absorbance of a solution. By using this property the formation of insoluble aggregates in the course of time can be monitored. In our case the measurement of optical density of aggregates within the aggregation buffer selected is measured. Specifically, samples were diluted in dilution buffer with a final concentration range between 3-20 μM. Samples were shaken at 14000 rpm at room temperature and absorbance was measured at 340 nm and 390nm.

3.2.22 Sedimentation assay

Sedimentation assay is the measurement of the amount of aggregated protein which is aggitated during a specific amount of time, after centrifugation. Protein was diluted in a dilution buffer dropwise and was shaken at 1400 rpm at room temperature for 2 hours. Every 30 min samples were centrifuged at 16100 g and the supernatant was separated from the pellet. Samples were analyzed by Western blot.

4. Results

4.1 Determining the suitable bacterial strain for protein expression

Two different strains of *E. coli* were tested and in particular Rosetta 2 and BL21(DE3), in order to determine the proper strain for TDP-43 expression.

BL21(DE3) strains contain the lambda DE3 prophage that carries the gene for T7 RNA polymerase under control of a lacUV5 promoter. Thus, the expression of T7 RNA polymerase can be induced with IPTG. Additionally, this strain is deficient in two key proteases Lon and OmpT that reduce the degradation of heterologous proteins expressed in cells [41], [42].

Rosetta 2 strains are derivatives of BL21 that enhance the codons that are rarely used in *E. coli*. These strains supply tRNAs for AGG, AGA, AUA, CUA, CCC, GGA codons on a compatible chloramphenicol-resistant plasmid [43]. In this way, they can enhance the expression of proteins, whose genes contain these rare codons. Such a protein is TDP-43 which contains some codons rarely used in *E. coli*, as it results from the codon usage frequency distribution in figure 3. Specifically, in figure 3 the percentage of codons for DNA polymerase I and RNA polymerase sigma E factor which are enzymes usual in *E. coli* cells and TDP-43 is compared. As it is observed, ~60% of DNA polymerase I codons and ~50% of RNA polymerase sigma E factor codons are used with a frequency of >90% while ~40% TDP-43 codons are used in the same frequency.

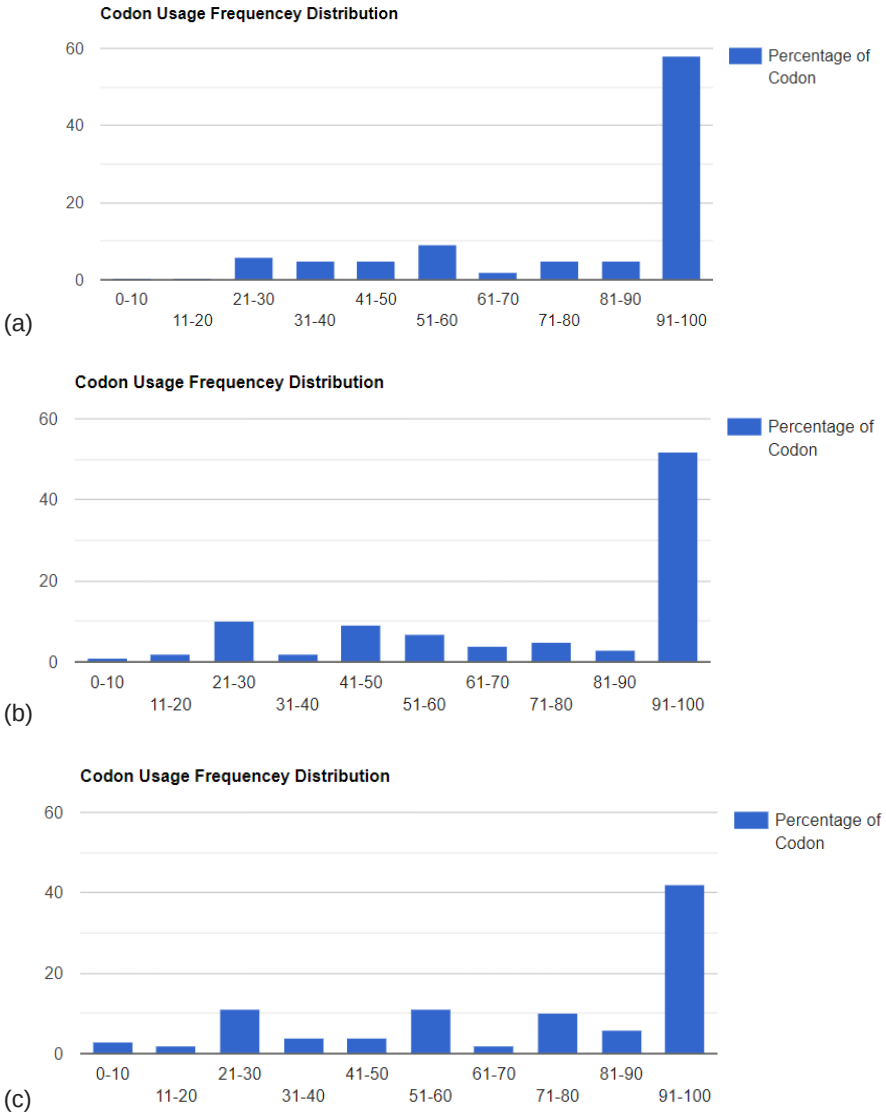

Figure 3: The percentage of codon (a) for DNA polymerase I (b) RNA polymerase sigma E factor and (c) TDP-43 is presented with information available at https://www.biologicscorp.com/tools/RareCodonAnalyzer#.YbCWOdBBxPb.

Rosetta 2 and BL21(DE3) cells containing pET28a-TDP-43-GFP plasmid that was already available in the laboratory, were induced at 30 °C for 3 hours. Then viability and *in vivo* fluorescence were measured with UV-Vis spectrophotometer and Magellan respectively and were found to be almost identical in all cases (Fig. 4).

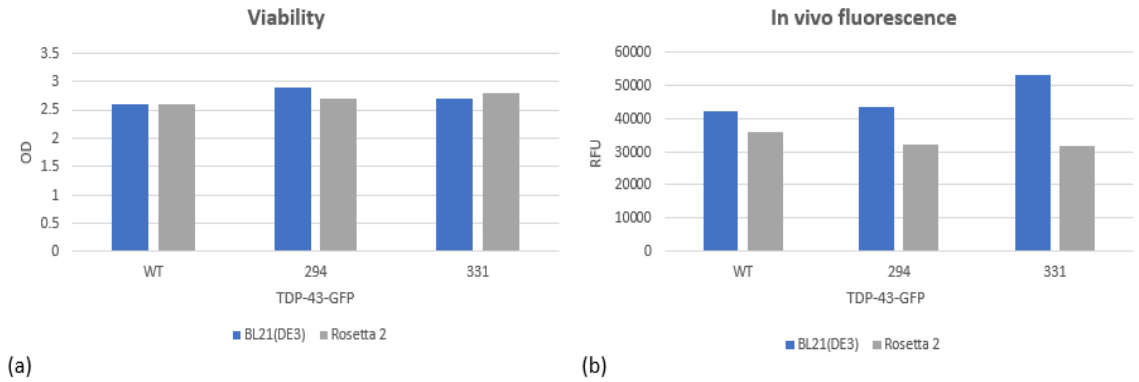

Figure 4: (a) Viability of BL21(DE3) and Rosetta 2 cells when expressing TDP-43-GFP protein in a pET28a plasmid. (b) Relative fluorescence units (RFU) of BL21(DE3) and Rosetta 2 cells when expressing TDP-43-GFP protein in a pET28a plasmid.

Next, TDP-43-GFP (70 kDa) expressed in BL21(DE3) or Rosetta 2 cells was extracted via sonication and fractionated in soluble and insoluble fraction via centrifugation (Fig.5).

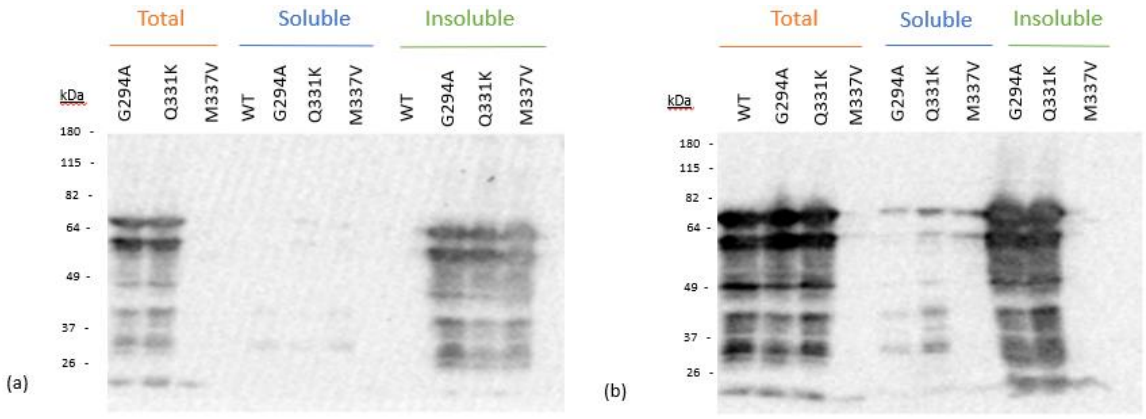

Figure 5: Western blot with anti-GFP antibody of total soluble and insoluble fractions of TDP-43-GFP expressed in (a) Rosetta 2 and (b) BL21(DE3) cells.

It was observed that the yield of expressed TDP-43-GFP protein (70kDa) was higher in BL21(DE3) cells. This could be possibly attributed to the fact that Rosetta 2 cells contained two plasmids, while BL21(DE3) contained one. Specifically, Rosetta 2 cells contained one plasmid that produced the rare tRNAs (chloramphenicol resistant) and another with TDP-43 protein in a pET28a vector (kanamycin resistant). On the other hand, BL21(DE3) contained only the plasmid with the TDP-43 protein. There are indications that two plasmids with two different antibiotic resistances may reduce the productivity of cells and in our case Rosetta 2 cells [44].

Furthermore, it was observed that in both cases, a higher amount of protein was in the insoluble fraction. In previous experiments of our laboratory it has been again reported that pET28a-TDP-43-GFP plasmid leads to a protein with higher yield of insoluble fraction.

All things considered, BL21(DE3) cells successfully expressed TDP-43 protein, as they performed similar viability and fluorescence with Rosetta 2 cells as well as even higher yield of produced protein. Thus, BL21(DE3) cells were the ones selected for the rest of the experiments.

4.2 Production of TDP-43 fused with GFP and a His tag

TDP-43 protein was transferred from a pET28a vector where it was already fused with a GFP to a pASK75 vector containing a GFP but also a 6xHis tag which was necessary for protein's purification (Fig.6).

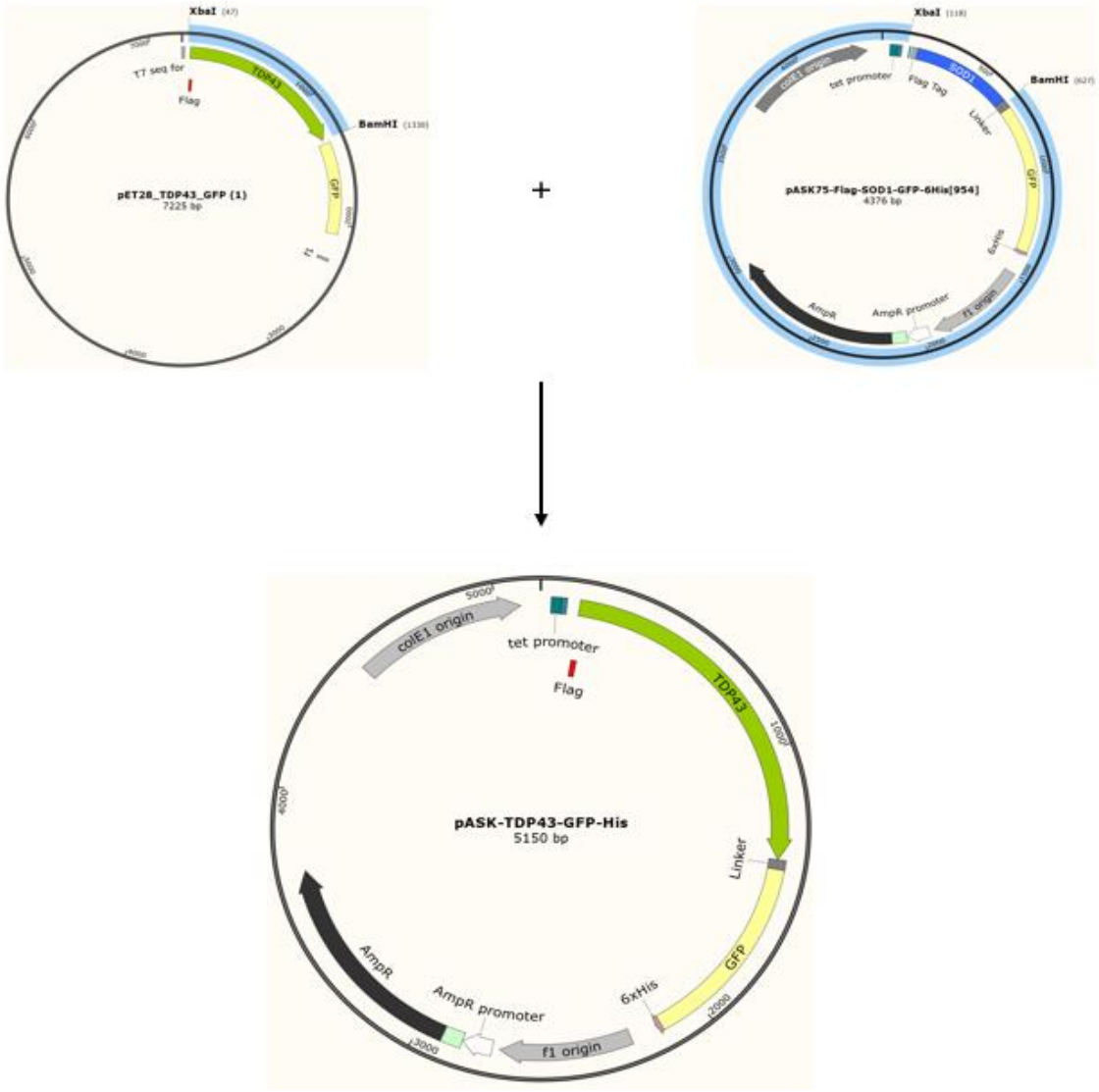

Figure 6: TDP-43 insert is transferred from pET28a plasmid to a pASK75 that has a GFP and a 6xHis tag.

Data of digestion and of colony PCR are presented in Figure 7.

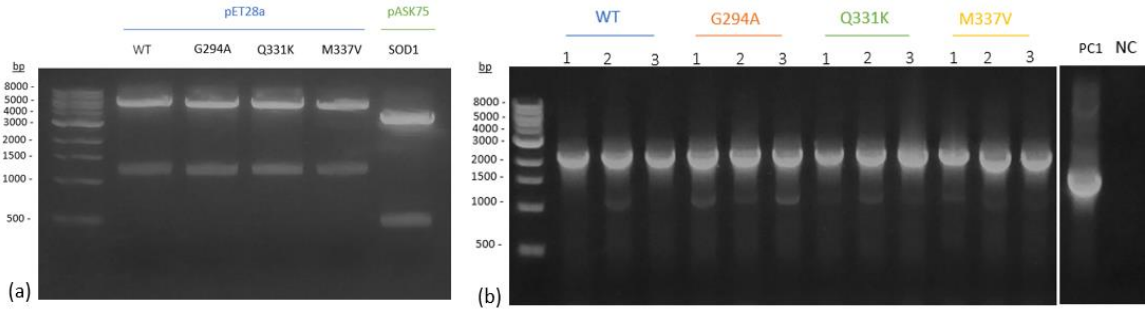

Figure 7: (a) Digestion of pET28a donor and pASK75 recipient plasmid with restriction enzymes XbaI and BamI. After pET28a-TDP-43-GFP digestion, 2 fragments were created. The first one with TDP-43 protein (1283 bp) and the second with pET28a donor plasmid with the GFP (5942 bp). pASK75 recipient plasmid that contained SOD1 protein was cut and 2 parts were created too. The first part was the Superoxide dismutase (SOD1) gene (509 bp) and the second part the pASK75 recipient vector with its GFP and 6xHIs tag (3867 bp). (b) Colony PCR with DD035 and DD037 primers at TDP-43 mutants. As positive control pASK75-SOD1 plasmid was used that it resulted at a 1519 bp fragment. As negative control ddH2O was used. All the colonies had the expected fragment (2293 bp).

The diagnostic digestion and diagnostic PCR performed for the evaluation of the created products are presented in Figure 8. Due to time limitations TDP-43-337-GFP-His mutant could not be constructed. The colonies selected for the rest of the experiments were the WT (2,) G294A (2), Q331K (2).

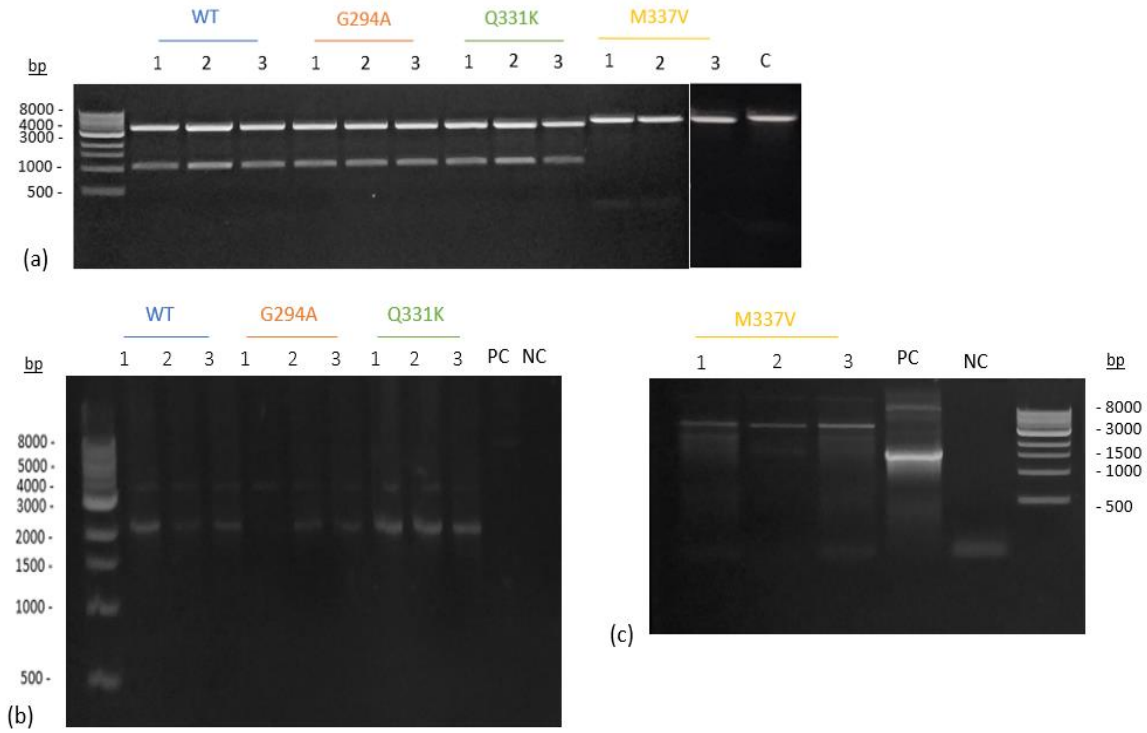

Figure 8: (a) Diagnostic digestion with double cutter restriction enzyme EcoRI leads to a 4053 bp fragment and a 1097 fragment. (b) Diagnostic PCR with pASK75a forward primer and DD037 reverse primer at WT, G294A, Q331K and M337V mutants.

Finally, in order to check subcloning further, BL21(DE3) cells transformed with pASK75-TDP-43-GFP-His plasmid were induced with 0.2 μg/mL ATc for 20 hours at 20 °C. Protein was extracted via sonication and was fractionated. SDS-page and Western blot with anti-His antibody followed.

Ponceau S revealed that equal amount of proteins were loaded at WT and G294A while less at Q331K mutant. Western blot revealed that WT, G294A, Q331K mutants had indeed the 70 kDa protein (Fig. 9). Satisfying amount of protein was expressed both in soluble and insoluble fractions.

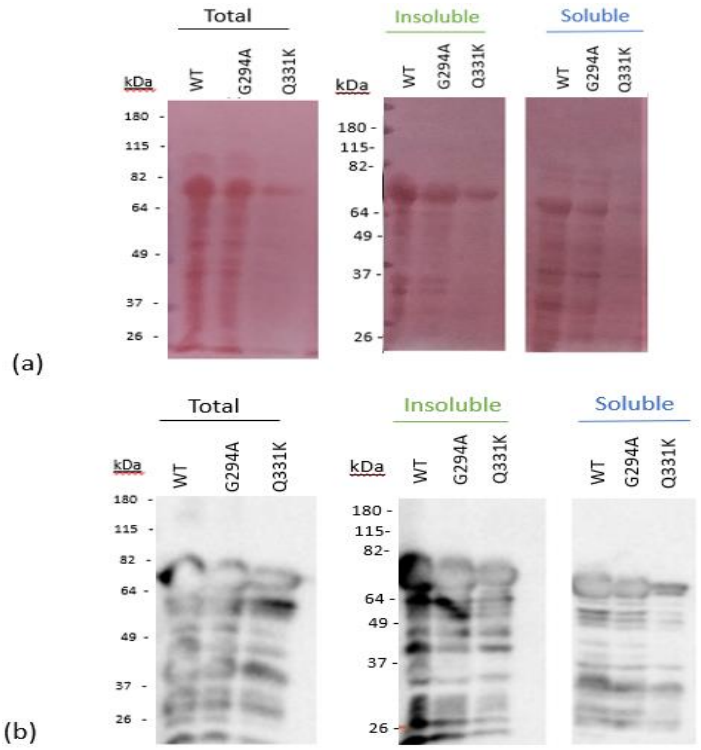

Figure 9: Denaturing gels after induction of pASK75-TDP-43-GFP-His plasmid with anti-His-HRP conjugated antibody. WT, G294A, Q331K mutants that were manufactured and were induced with 0.2 μg/mL ATc.

4.3 Purification of TDP-43-GFP-His with Ni-NTA

For TDP-43-His protein purification the WT and Q331K mutant (that as mentioned in the introduction is more aggregation prone from G294A mutant) were chosen. Both TDP-43 soluble and insoluble fractions were purified as described below.

The purification of protein's soluble fraction was performed with AKTA Ni-NTa column and SDS-page denaturing gels followed. Coomassie staining and western-blot with anti-His-HRP conjugated antibody evaluated the efficiency of the purification (Fig. 10a, 10b). The numerous bands in Coomassie stained gel

possibly indicate that protein was moderately purified. The numerous bands with antibody could be due to antibody's low specificity and the fragments of protein created during sonication [45].

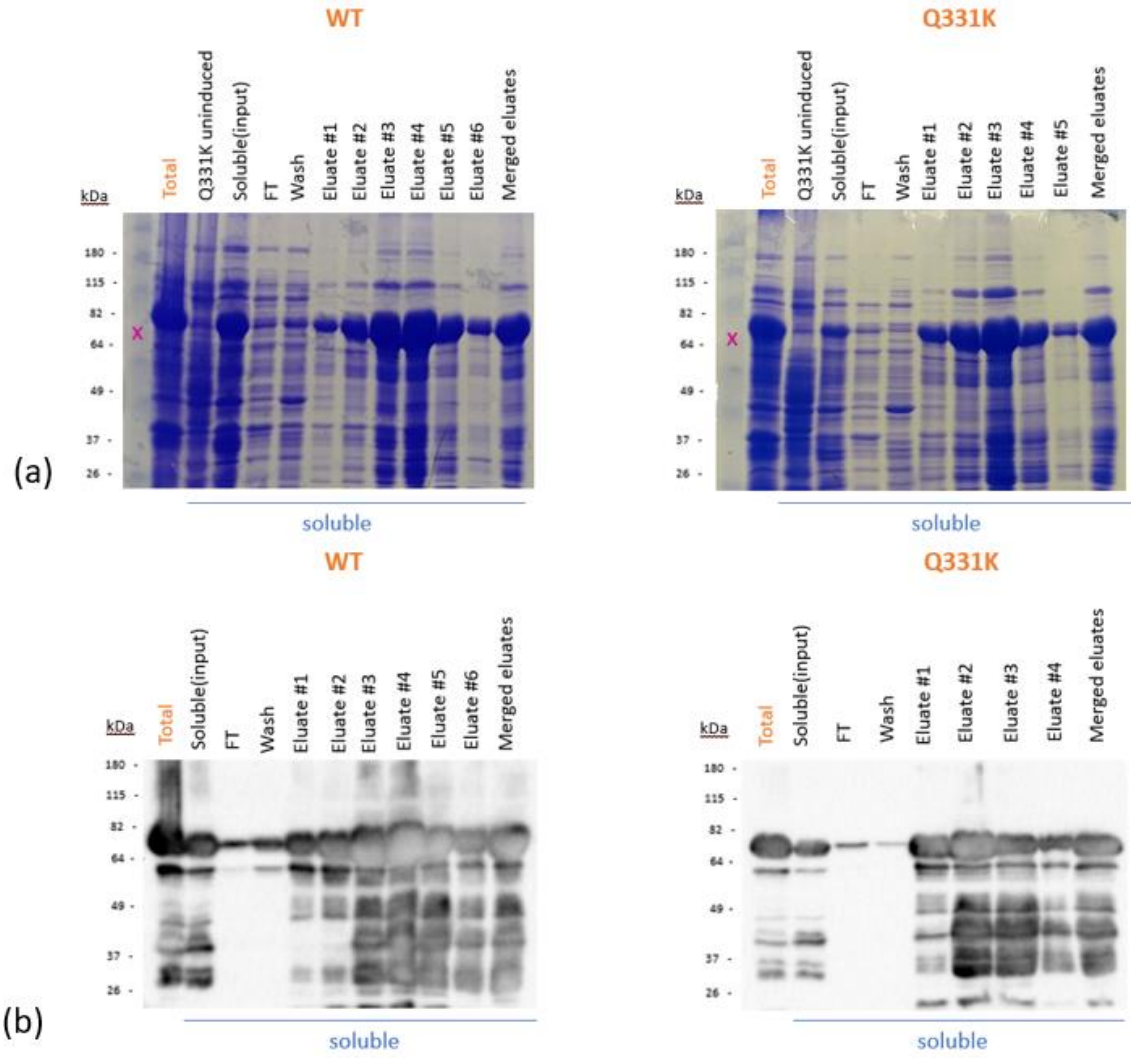

Figure 10: Denaturing SDS-page gels of soluble fraction eluates of TDP-43-GFP-His purified with AKTA Ni-Nta column: (a) Stained with Commassie (b) Western blot with anti-His anti-body.

The purification of protein's soluble fraction was also performed with Ni-NTa gravity flow column. Via native PAGE of protein's soluble fraction, information about the protein eluates would be revealed. The bands were presented via Coomassie staining (Fig. 11a,11b) and western-blot with anti-His-HRP conjugated antibody (Fig. 11c,11d). Although there is a prevalent band in Coomassie staining, there are also traces beyond and beneath, suggesting the co-existence of other protein species such as TDP-43 oligomers, other proteins etc. Furthermore, due to the lack of a proper molecular weight standard it is not possible to decide the kind of prevalent species, i.e. monomer, dimer, tetramer etc.

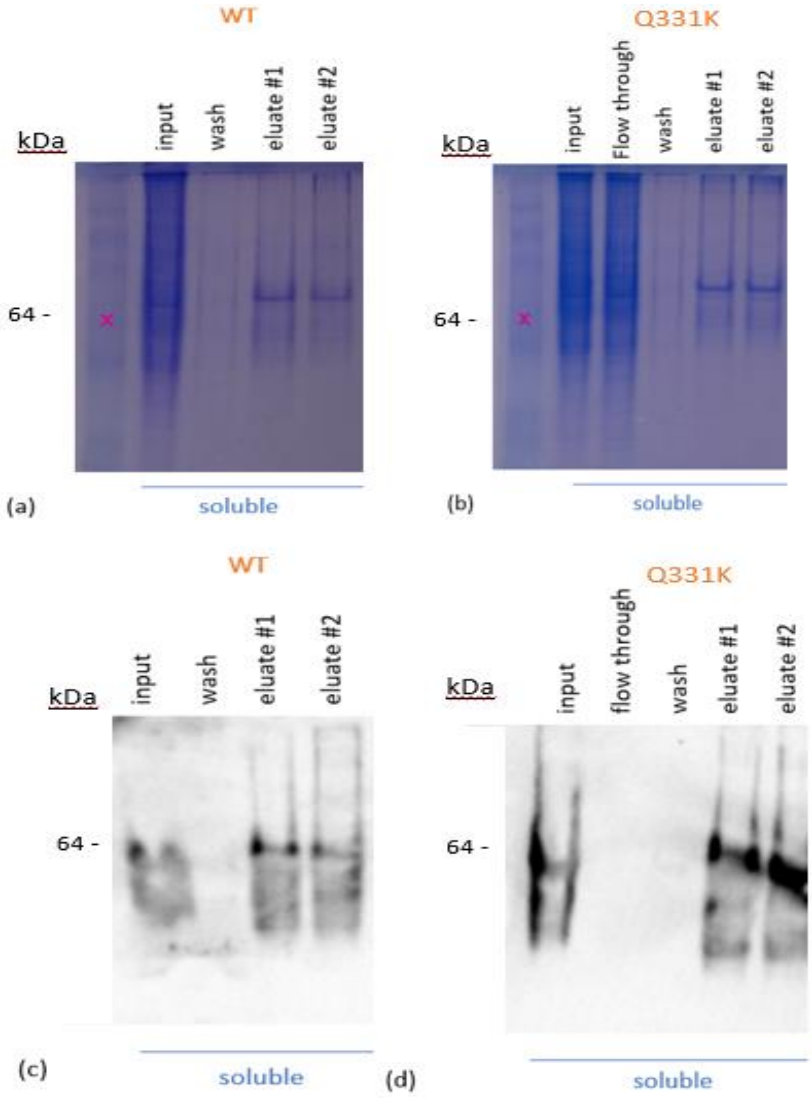

Figure 11: Native gels of soluble fraction eluates of TDP-43-GFP-His purified with Ni-Nta gravity flow column: (a),(b) stained with Coomassie. (c),(d) western blot with anti-His- HRP conjugated antibody.

The purification of protein's insoluble fraction was performed with Ni-NTa gravity flow column, too. SDS-page denaturing gels, Coomassie staining and western-blot with anti-His-HRP conjugated antibody followed (Fig. 12).

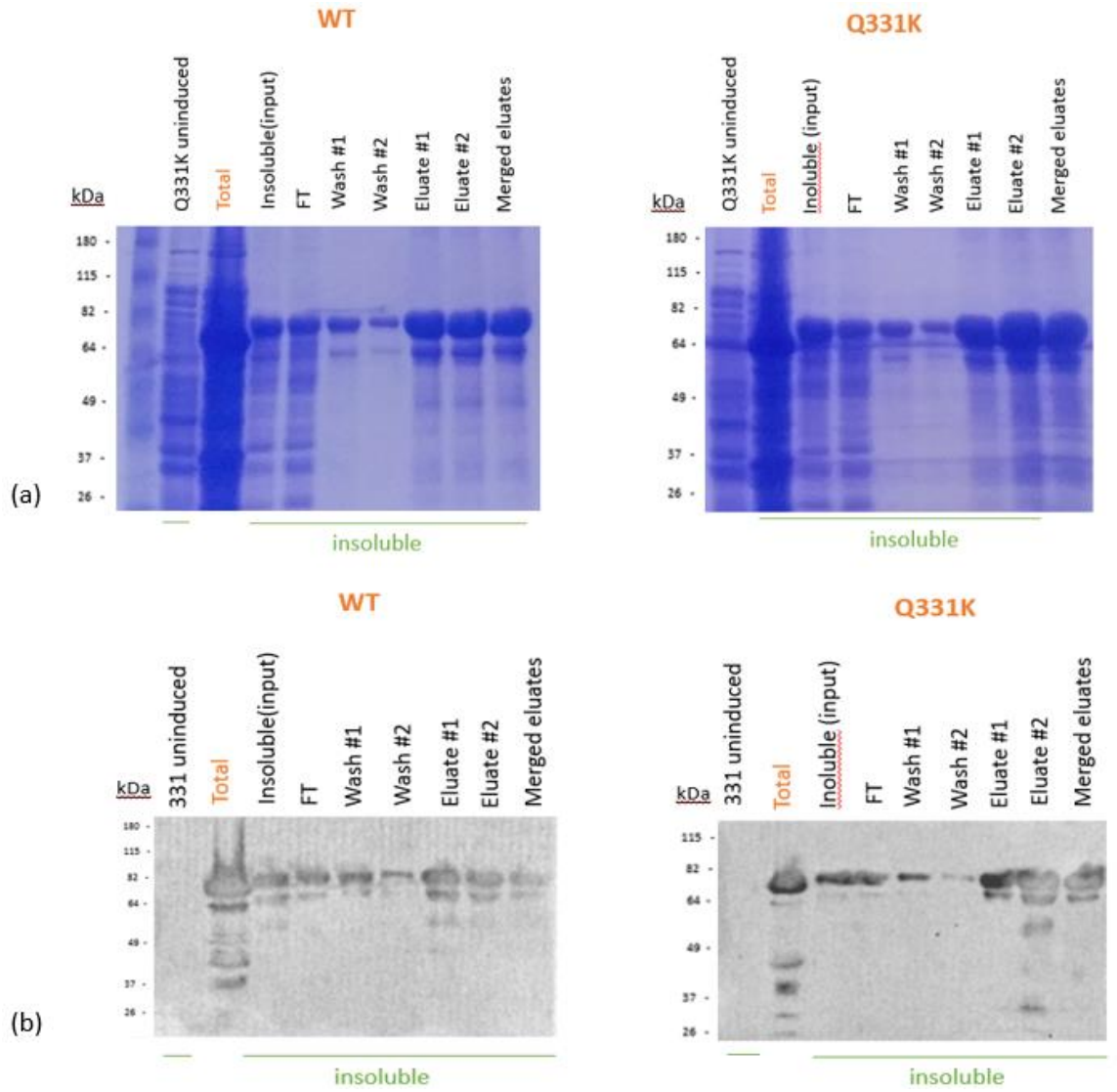

Figure 12: Denaturing SDS-page gels of insoluble fraction eluates of TDP-43-GFP-His purified with Ni-Nta gravity flow: (a) Stained with Coomassie (b) Western blot with anti-His anti-body.

Insoluble fraction of TDP-43-GFP was sufficiently purified. Inclusion bodies fraction had almost pure TDP-43-GFP as they usually contain the overexpressed protein almost exclusively [46].

4.4 Purification of TDP-43-GFP-His with SEC

After Ni-Nta chromatography the purification of TDP-43-GFP protein from soluble fraction and from inclusion bodies continued with size exclusion chromatography (SEC) and fluorescence of each eluate was measured (Fig.13a,13b). Further purification of TDP-43-GFP protein with SEC chromatography was essential in order to specify the state of oligomerization of the protein (oligomer, dimer, monomer etc) and select ideally the dimer which is its physiological oligomerization state for the aggregation assays [16]. By achieving this, the following aggregation experiments would begin with a non oligomeric protein or a zero aggregation.

Due to time limitations, only WT mutants were loaded. In both cases, almost 90% of TDP-43-GFP protein was retained within the column thus there was a really small yield of eluates of protein.

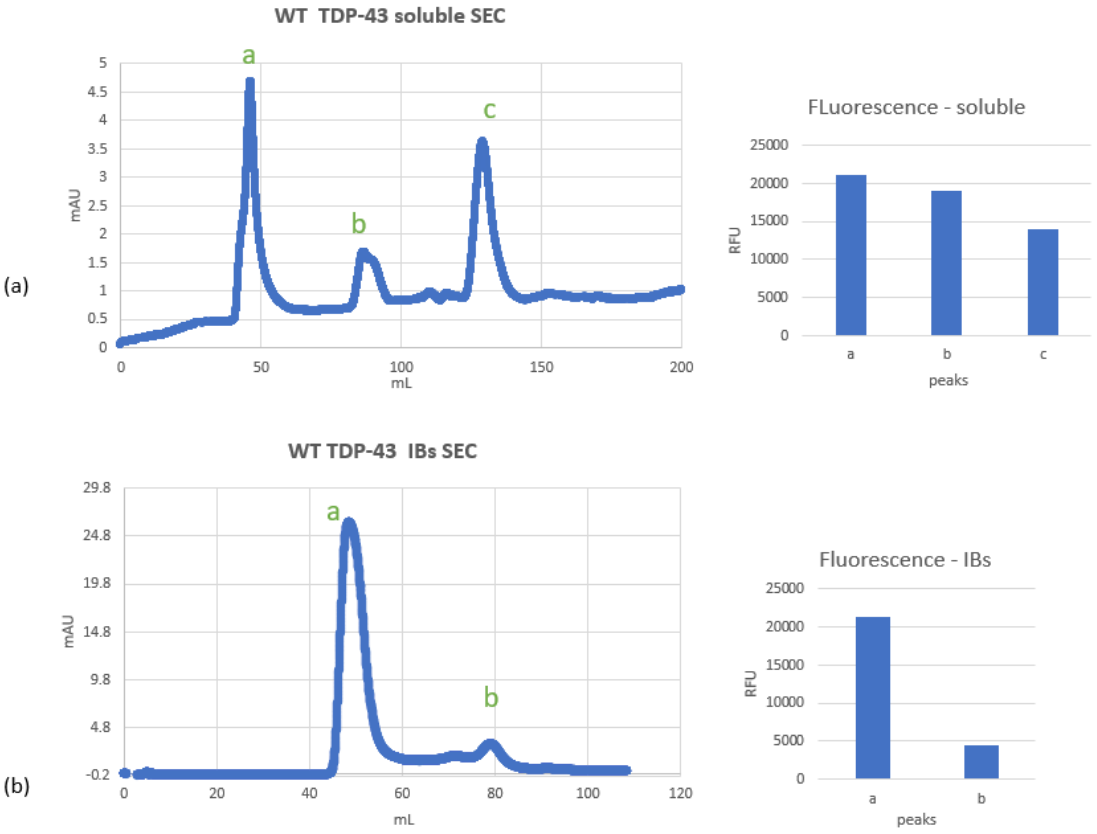

Figure 13: SEC chromatogram of (a) WT TDP-43-GFP soluble and (b) WT TDP-43-GFP inclusion bodies with the relative fluorescence units of each peak respectively.

The majority of WT TDP-43-GFP protein eluate from IBs was concentrated almost only in one peak, probably because as mentioned before IBs usually contain the overexpressed protein almost exclusively [46]. Moreover, the IBs eluates performed five times higher absorbance values than the eluates of soluble fraction. Interestingly, at soluble protein eluates while peak b had lower absorbance than peak a, it had a higher specific fluorescence. This could be

attributed to the fact that the more the protein is aggregated the less fluorescent is.

In order to calculate the theoretical state of oligomerization, the equation of HiLoad 16/600 Superdex 200 prep grade from diagram 1 at SEC - methods section was used. For the elution volume of each peak, the log (MW) was calculated and MW was found (Table 2). Thus, by dividing the predicted molecular weight of the species present at each elution peak with the molecular weight of TDP-43-GFP (70 kDa), the oligomerization state was found. So, the oligomerization state for soluble eluates was 26 for peak a, 4 for peak b, 1 for peak a, and for insoluble eluates, 21 for peak a and 5 for peak b. It is observed that both soluble and insoluble fractions contain very similar species, and soluble fractions contains one more with oligomerization state 1 that it may be truncated.

Soluble

Peak	Elution volume	log(MW)	MW	Oligomerization state
a	46.2	3.260	1804.600	25.779
b	87	2.400	248.600	3.551
c	129	1.509	32.307	0.462

IBs

Peak	Elution volume	log(MW)	MW	Oligomerization state
a	50	3.176	1500.376	21.434
b	80	2.543	349.301	4.990

Table 2: Calculation of theoritical oligomerization state of each peak's eluates of HiLoad 16/600 Superdex 200 prep grade.

Taking into account the above considerations, in soluble fraction peak b with oligomerization state 4 which had the highest specific, would be the best eluate to be selected for the aggregation assays fluorescence (while the physiological oligomerization state of TDP-43 is a dimer). This is attributed to the fact that the less the protein is oligomerized the better the results about its aggregation propensity would be.

This theoretically calculated oligomerization state could be proved with SDS-page. However, the protein eluted in both cases was so diluted that SDS-page of eluates was not possible.

4.5 TDP-43-GFP Aggregation assays

TDP-43 aggregates did not react with amyloid-specific dyes Thioflavin-T and Congo red [19]. Thus, turbidity and sedimentation assays were preferred for monitoring the aggregation propensity of TDP-43.

4.5.1 Sedimentation assay

WT and Q331K TDP-43-GFP-His solubilized inclusion bodies passed through a sedimentation assay before their purification with Ni-Nta gravity.

Solubilized inclusion bodies were diluted in buffer A and in buffer B (Fig. 14). Buffer A was consisting of 50 mM HEPES, 150mM NaCl. Buffer B was consisting of 50 mM HEPES, 50 mM KCl, 75 mM NaCl, 0.05% PEG3350, 0.05% sucrose, 0.0125% tritonX-100. In buffer B, PEG reagent as well as the nonionic surfactant tritonX-100 reduce aggregation [47] and sucrose stabilizes the protein [48]. All these reagents aid protein stay in its solubilized form and thus, delay aggregation.

When protein is diluted at buffer A the quantity of protein both at WT and Q331K mutants' pellets was elevated during 4 hours of shaking. The quantity of protein at supernatants was gradually decreasing. At Q331K mutant after 3 hours there was a band of oligomers that over time was denser. At WT mutant after 3.5 h oligomers started to be created but there was not an obvious band. This could possibly be attributed to the fact that Q331K mutant is more aggregation prone. However, no conclusion can be drawn as there were aggregates from the beginning of the experiment and it was only conducted once.

So, a problem created both in sedimentation and in turbidity assays mentioned below was that at the starting point of aggregation measurements an important amount of protein had already been aggregated.

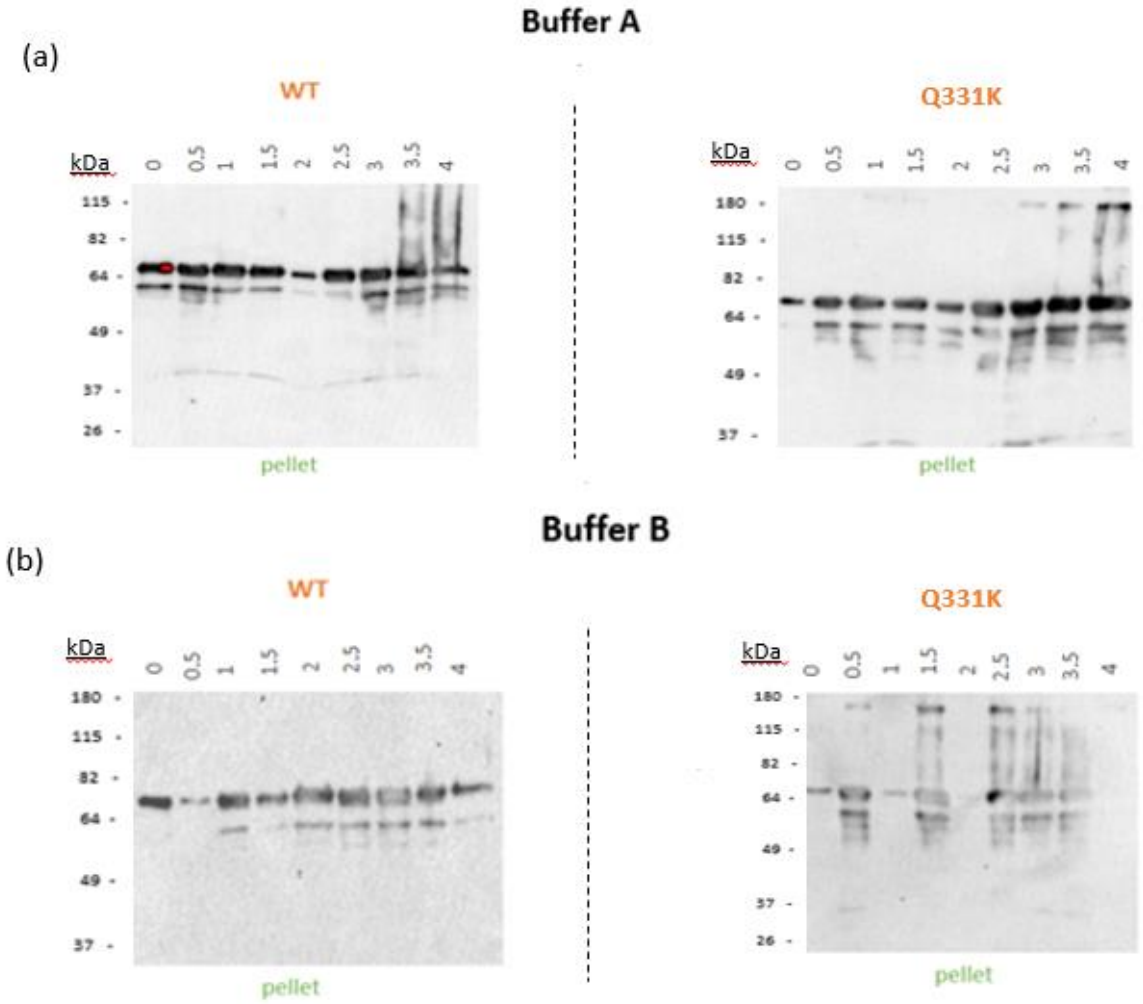

Figure 14: Sedimentation assays of WT, Q331K solubilized IBs before purification with Ni-Nta in denaturing gels (a) diluted at buffer A. (b) diluted at buffer B.

4.5.2 Turbidity assay

The solubilized inclusion bodies of WT and Q331K mutant of TDP-43-GFP protein before their purification with Ni-Nta column, passed through the turbidity assay. In turbidity assays the first parameter to determine was the nm of absorbance. Two wave lengths were measured: 340 nm and 390 nm (Fig.15). Both nanometers of absorbance led to similar behavior in turbidity assays. 340 nm were selected for the rest of the experiments as it was more often used in bibliography.

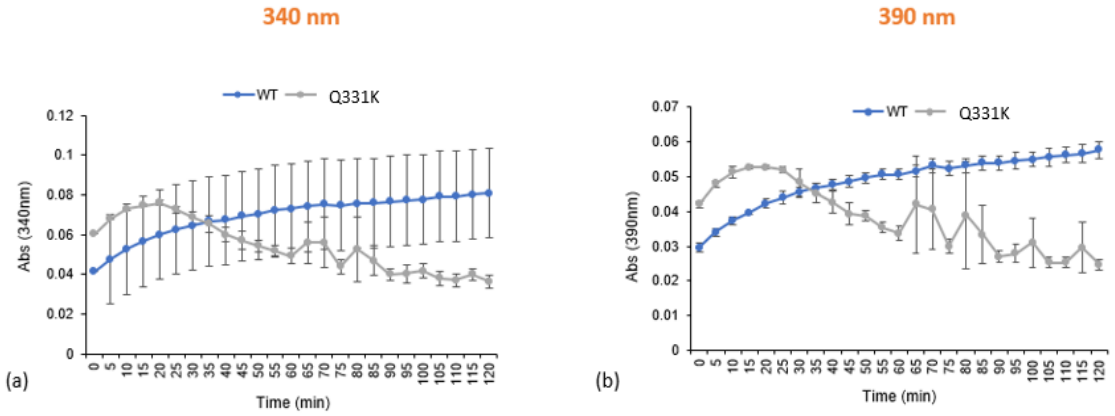

Figure 15: Turbidity assays of solubilized IBs of TDP-43-GFP before purification with Ni-Nta at: (a) 340 nm and (b) 390 nm.

The next parameters tested were the dilution buffer of protein for turbidity assays, as well as protein's concentration (Fig. 16). Buffer A contained 50mM Tris, 0.5M and NaCl. Buffer B contained 50mM Tris, 0.25M KCl, 0.25M NaCl, 50mM imidazole, 0.05% PEG 3350, 0.0125% triton X-100, 50mM sucrose. Buffer B because of the reagents that it was consisted of, it enhanced protein's well folding and delayed protein's aggregation. As the experiment revealed, the best conditions for the turbidity assays was at 3 μM and with buffer A because there was a gradient rise in aggregation which agrees with the assays of Johnson 2019 [19].

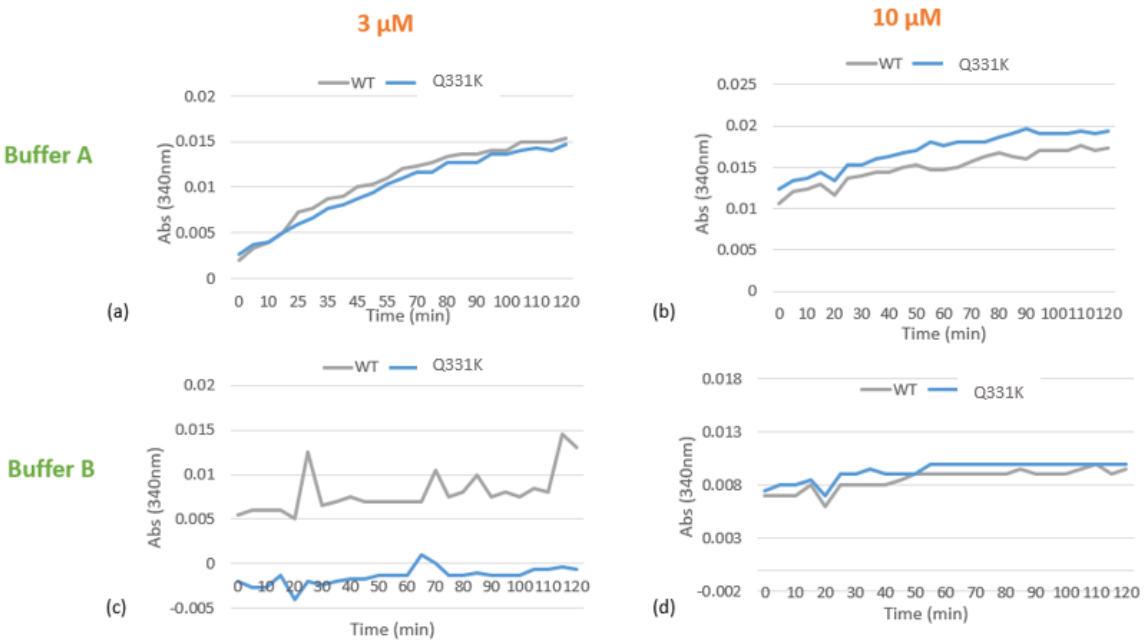

Figure 16: Turbidity assays of solubilized IBs of TDP-43-GFP before purification with Ni-Nta at 340 nm diluted in: (a) bufferA and 3 μM protein (b) buffer A and 10 μM protein (c) buffer B with 3 μM protein and (d) buffer B with 10 μM protein.

Finally, the extend of aggregation of soluble TDP-43-GFP after purification with Ni-Nta column was determined by turbidity. For the first 15 minutes there was an

augmentation in WT and for the first 40 minutes an augmentation in Q331K mutant (Fig. 17). After reaching a peak OD decreased. This could be attributed to the fact that aggregates are settled at the periphery of microplates.

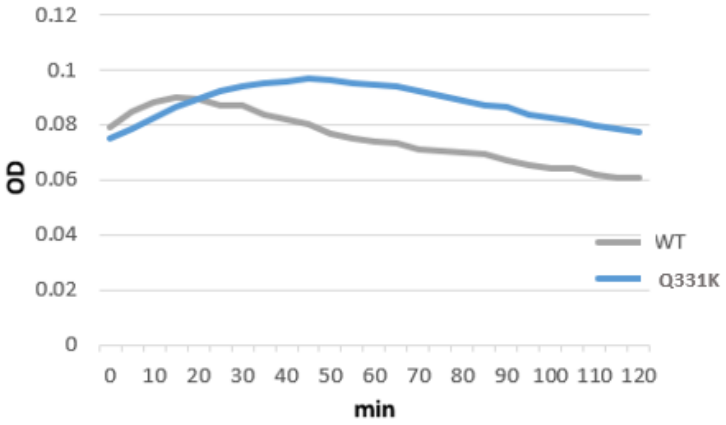

Figure 17: Turbidity assays of 3 μM soluble TDP-43-GFP after purification with Ni-Nta at 340 nm diluted in buffer A.

4.6 Determining the proper growth phase for induction

Growth curve of BL21(DE3) cells in the absence and presence of pASK75-TDP-43-GFP plasmid was determined in order to define the proper OD for protein induction. OD was measured for 12.5 hours in BMG Fluostar's microplate reader with constant agitation. Both plain BL21(DE3) and BL21(DE3) cells transformed with the pASK75 plasmid containing WT TDP-43-GFP were in a stationary phase after reaching OD=1 (Fig.5).

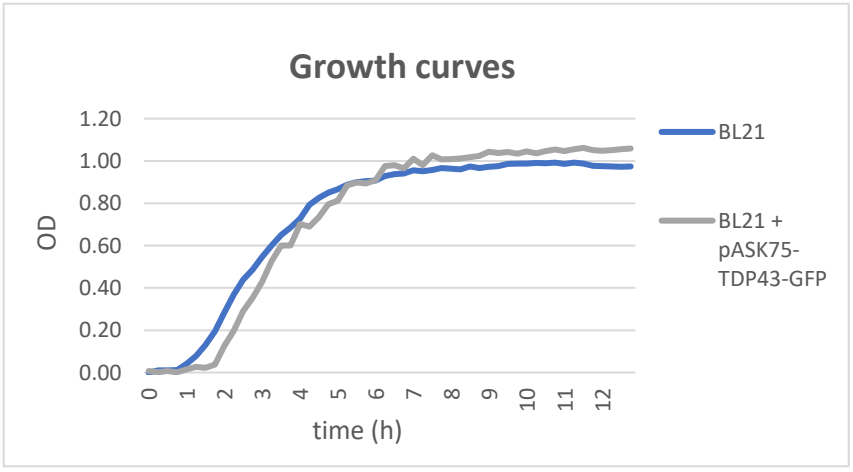

Figure 5: Growth curve of plain BL21(DE3) strain and of BL21(DE3) transformed with WT-pASK75-TDP-43-GFP during 12.5 hours.

Most protocols for expressing recombinant proteins recommend the induction of protein expression during the mid-log phase. This is presumably because the cultures are growing rapidly and protein translation is maximal. However, the

expression of soluble recombinant proteins as TDP-43, in late log phase *of E. coli* cultures yields higher recovery than that obtained from mid log phase cultures [49].

According to these results, in our experiments BL21 (DE3) cells were induced in an OD between 0.6 and 0.8.

4.7 Co-induction of TDP-43-GFP with macrocyclic pentapeptides

Co-induction of TDP-43-GFP protein with macrocyclic pentapeptides was a step needed in order to test if any of the four pentapeptides available had the ability to inhibit aggregation.

WT and Q331K mutants of TDP-43-GFP protein were co-induced with the following pentapeptides: TWSVW, CTWMR, TAFDR, TTYAR. R2 random pentapeptide was used as a control and PBS as blank (Fig. 18). These pentapeptides in previous experiments of our laboratory with other misfolded proteins had managed to increase the fluorescence while R2 was a random pentapeptide that did not enhance it. Specifically, TWSVW enhanced fluorescence of SOD1 protein fused with GFP and TAFDR, TTYAR and CTWMR of Aβ also, fused with GFP. As mentioned in the introduction, usually when there is a decrease in protein misfolding there is an augmentation in its fluorescence. Thus, pentapeptides that lead to a fluorescent phenotype should decrease protein misfolding.

CTWMR, TAFDR and TTYAR led to higher fluorescence than R2 in WT TDP-43-GFP and TAFDR led to equivalent fluorescence with R2 -and not less- in Q331K mutant of TDP-43-GFP. So TAFDR and TTYAR macrocyclic pentapeptides were used to test the fluorescence.

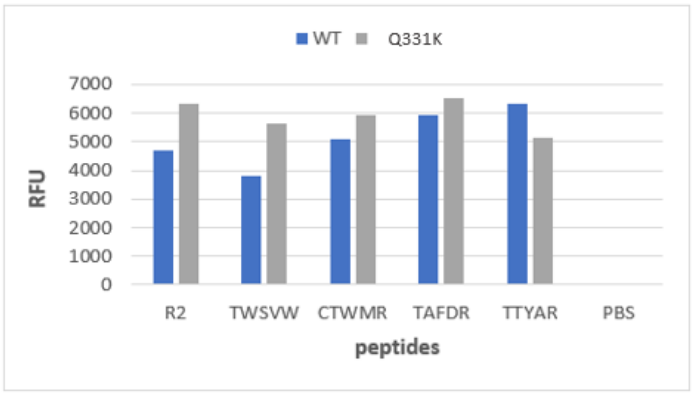

Figure 18: Relative fluorescence units of TDP-43-GFP co-expressed with TWSVW, CTWMR, TAFDR, TTYAR, a random peptide and PBS as blank.

4.8 Production of TDP-43 with a His and a Flag tag

In this molecular cloning the target was to transfer TDP-43 gene from a pET28a containing GFP to a recipient pET28a plasmid without GFP (Fig.19). So, the resulting protein is the TDP-43 tagged with 6xHis-tag and a Flag-tag. His-tag is necessary for protein's purification and Flag-tag is important in order to monitor the production of the full-length protein. The reason why GFP is absent is because it may alter the aggregation properties of the protein.

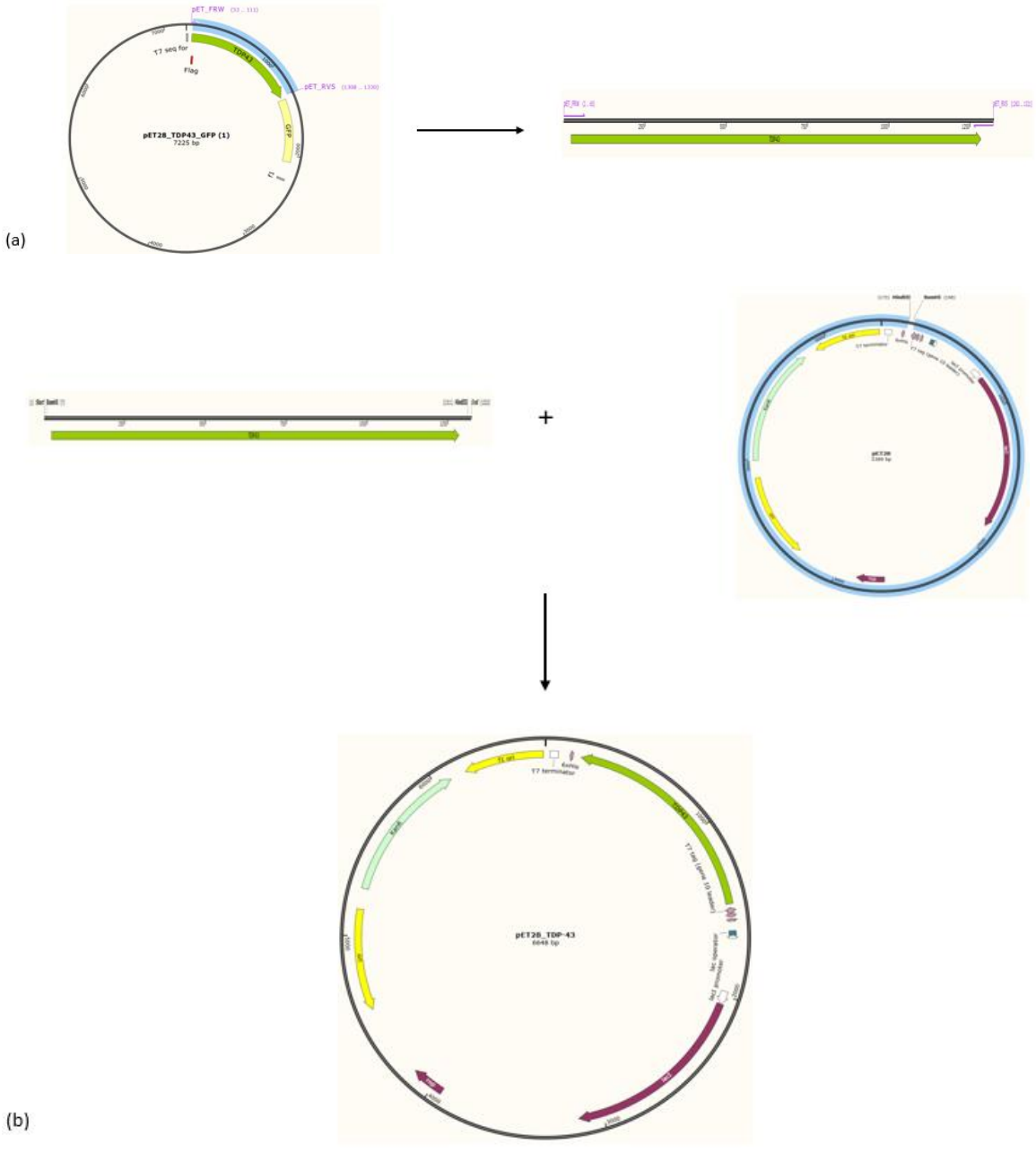

Figure 19: (a)Phusion PCR where the insert is amplified (b) The insertion of phusion PCR product in a plain pET28a vector.

Phusion and colony PCR results are presented at Figure 20. Diagnostic digestion and PCR for evaluation of the plasmid manufactured are presented in Figure 21. The experiments continued with WT (2), G294A (2), Q331K (1), M337V (1) TDP-43 mutants.

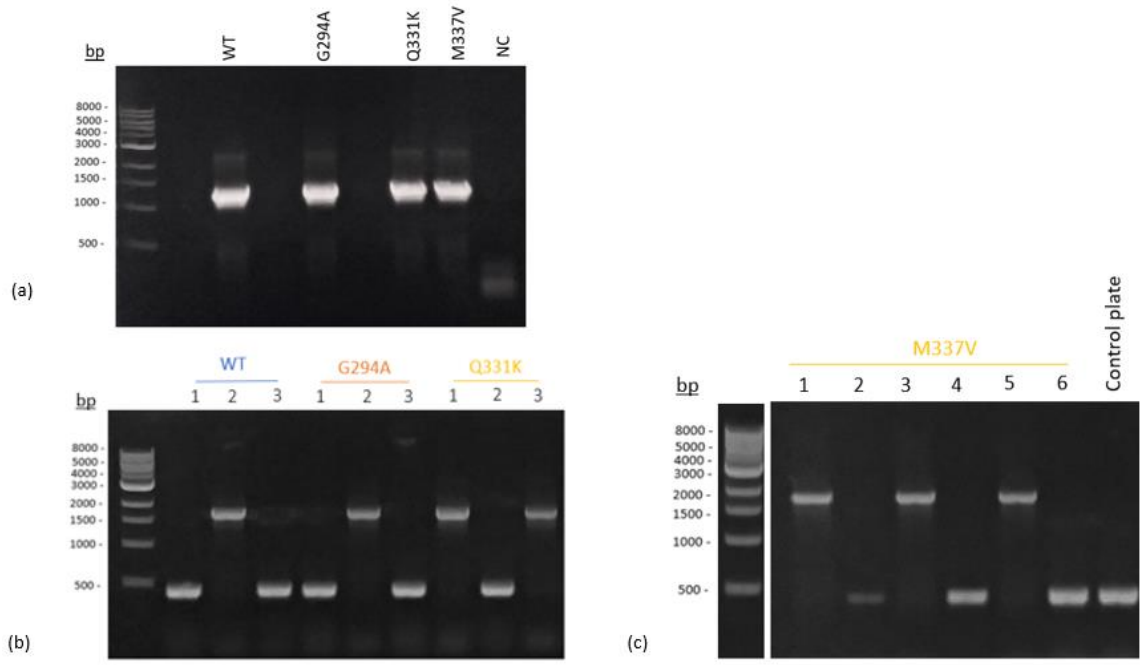

Figure 20: (a) Phusion PCR in pET28a-TDP-43-GFP plasmid. The forward pET28a primer contains a BamHI restriction site and a 6xHis tag and the reverse pET28a primer contains a HindIII restriction site and a Flag-tag.The DNA amplified is the TDP-43 protein with BamHI, HindIII restriction sites and 6xHis and Flag tag (1321bp). (b), (c) Colony PCR at three colonies of WT,G294A,Q331K mutants and 6 colonies of M337V mutants with a control colony of the control plate that was not containing the T4 ligase. GS088 and DD037 primers. Expected product at 1710 bp.

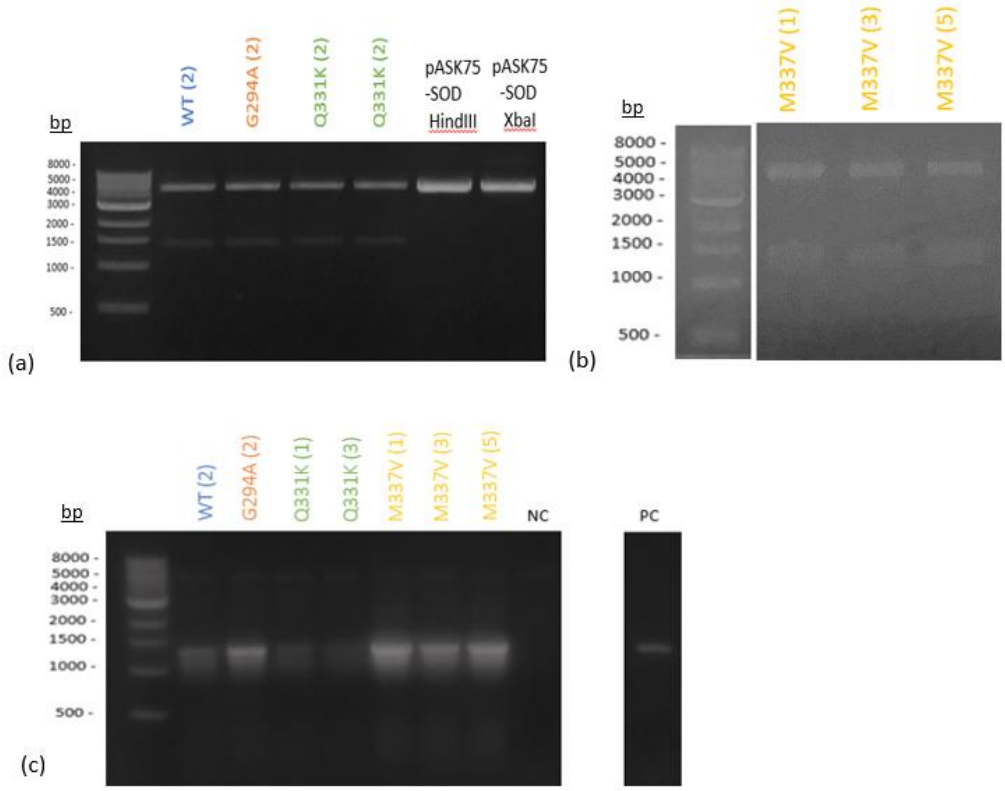

Figure 21: (a) Diagnostic digestion with XbaI and HindIII at WT and G294A, Q331K mutants. (b) Diagnostic digestion with XbaI and HindIII at M337V muatnt. Expected products: 1441 bp and 5207 bp. (c) Diagnostic PCR with pET28a forward and reverse primer. Expected products 1310 bp.

In order to do a final check in molecular cloning, BL21(DE3) cells transformed with pET28a-Flag-TDP-43-His plasmid were induced with 1 mM IPTG for 20 hours at 20 °C. Protein was extracted via sonication and it was fractionated. SDS-page and Western blot with anti-His and anti-Flag antibodies followed.

Ponceau S staining revealed that all TDP-43 mutants were expressing proteins -the same amount of proteins was loaded- (Fig. 22a). Western blot revealed that only G294A and M337V mutants had the expected 43 kDa protein (Fig. 22b, 22c). The same black spots on anti-His and anti-Flag western blot mean that the protein contains both Flag and His -tag and thus the full-length protein is produced. Thus, only G294A and M337V mutants were well manufactured and able to be correctly expressed. Satisfying amount of protein was expressed both in soluble and insoluble fractions.

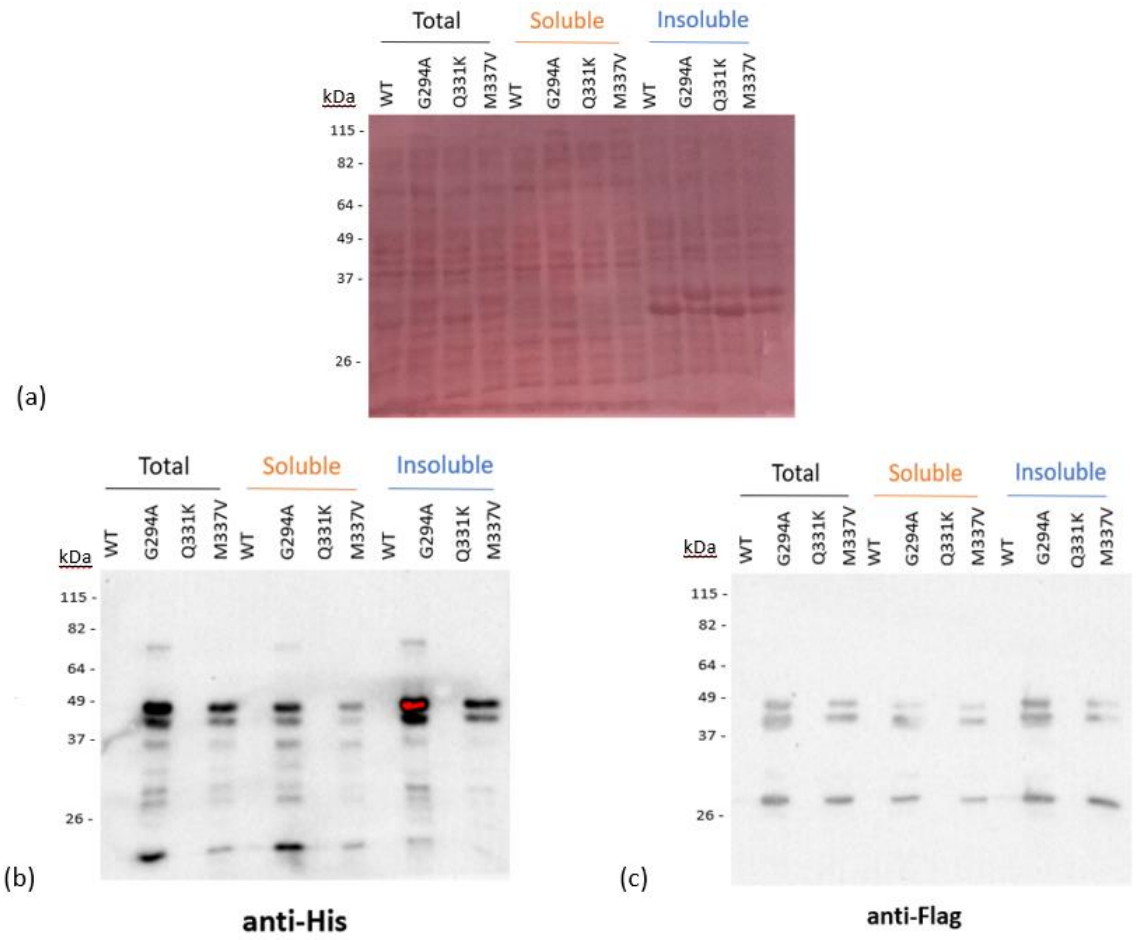

Figure 22: (a) Ponceau S stained denaturing gels after induction of pET-Flag-TDP-43-His plasmid (b) Denaturing gels after induction of pET-Flag-TDP-43-His plasmid with anti-His-HRP conjugated antibody (c) Denaturing gels after induction of pET-Flag-TDP-43-His plasmid with anti-Flag-HRP conjugated antibody.

4.9 Purification of TDP-43-His with Ni-Nta

As the protein mutations successfully manufactured without GFP were G294A and M337V, the experiments continued only with these two mutations. pET28α-Flag-TDP-43-His mutations were transformed into BL21(DE3) cells. The cells were induced with 1mM IPTG and there was a control sample, too, that was not induced. The protein was extracted via sonication and was fractionated via centrifugation. Their soluble as well as their insoluble fraction were purified via Ni-Nta gravity chromatography (Fig. 23).

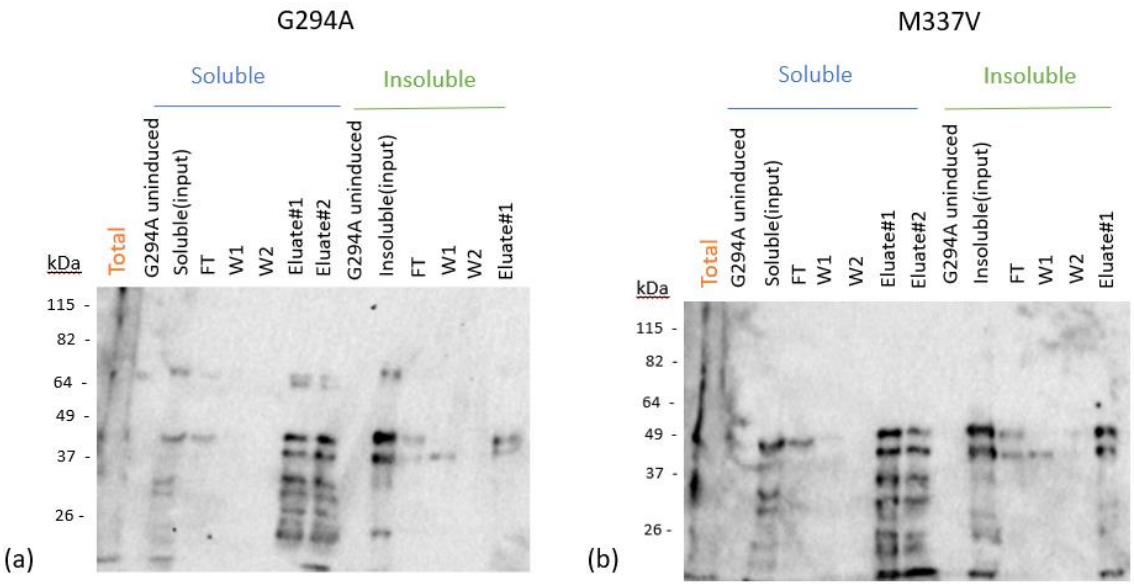

Figure 23: Denaturing total fraction and Ni-Nta gravity eluates of soluble and insoluble fraction of TDP-43 in mutants: (a) G294A and (b) M337V.

4.10 TDP-43 Aggregation assays

For turbidity assay soluble G294A and M337V mutants after Ni-Nta purification, were diluted with Tris 50 mM buffer to a final concentration of 20 μM. They were shaken at the microplate reader for 2 and a half hours and absorbance was measured every 10 minutes (Fig. 25). After one hour the increase in aggregation started to be apparent. At a concentration of 20 μM the aggregation was better. In order to reach a plateau more time was needed.

As described in Johnson 2009 [19] and as these experiments revealed, TDP-43 mutants are intrinsically aggregation prone.

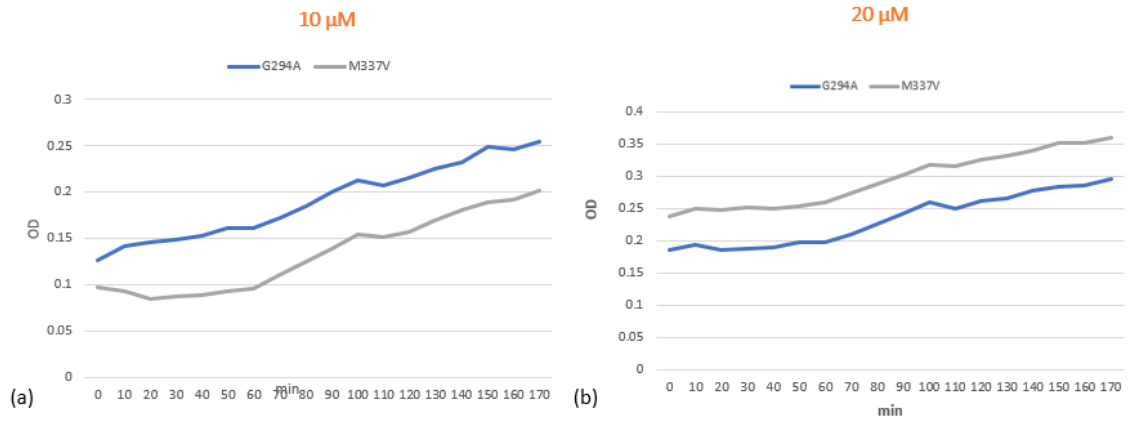

Figure 25: Turbidity assays in soluble fraction of G294A and M337V mutants of TDP-43 after Ni-Nta gravity flow column: (a) at 10 μM concentration (b) at 20 μM concenttration.

Finally, the effect of TAFDR and TTYAR pentapeptides, that were selected in previous experiments, in decreasing aggregation was tested.

TDP-43 solubilized inclusion bodies (insoluble fraction) before purification with Ni-Nta column were diluted in a buffer with Tris 50 mM, 0.5 mM NaCl, so that the final concentration was 20 μM. Samples were tested in a sedimentation assay for 2 hours alone and with 20 μM TAFDR or TTYAR pentapeptides (Fig. 24). At the start point of the measurements (t=0h) there was already aggregation, which remained stable for the whole experiment. In none of the pellets there was a rise in aggregation and in none of the supernatants there was a decrease in protein quantity. Possibly, supernatants had stable quantity of protein because the amount of protein that aggregated was small. Nevertheless, TAFDR and TTYAR did not affect the already aggregated protein.

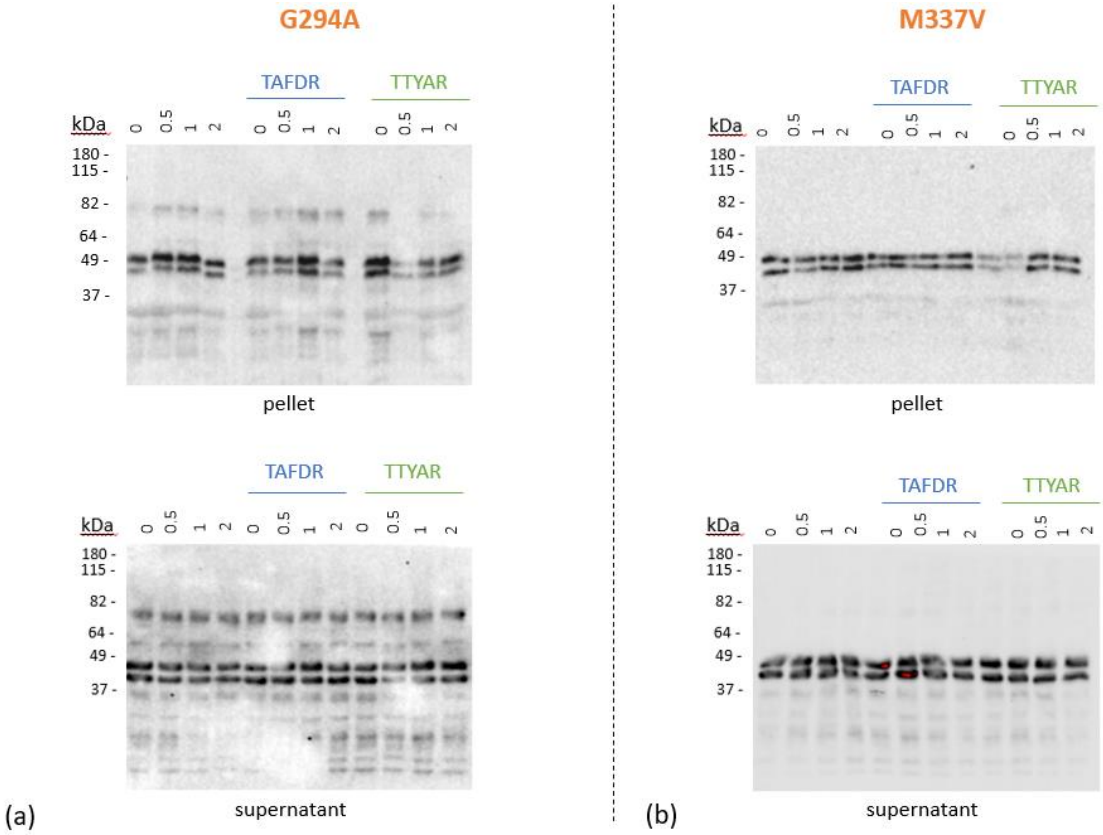

Figure 24: Sedimentation assays of plain WT, Q331K solubilized IBs before purification with Ni-Nta and with TAFDR and TTYAR pentapeptides. (a) G294A mutant (b) M337V mutant.

Finally, turbidity assays with plain TDP-43 and with TDP-43 with TAFDR and TTYAR macrocyclic pentapeptides evaluated the fact that the selected peptides did not affect TDP-43's aggregation (Fig. 26). The results are not clear and possibly more time was needed to see a difference in aggregation.

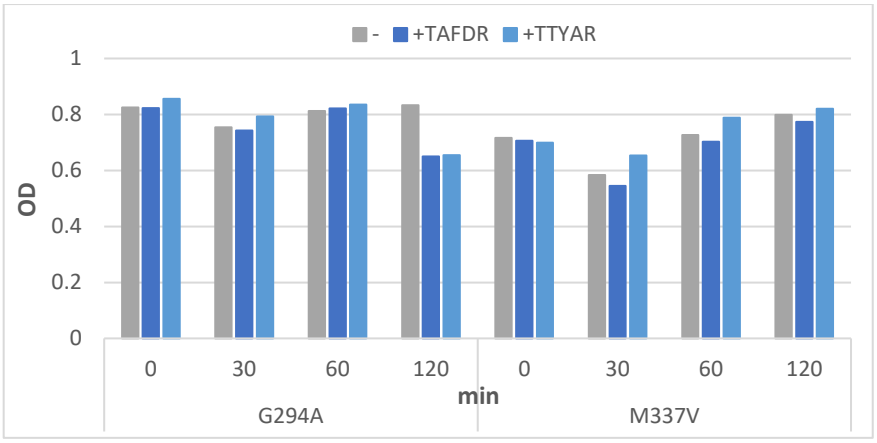

Figure 26: Turbidity assays of soluble G294A and M337V mutants of TDP-43 after purification with Ni-Nta gravity when it is plain and when it is with macrocyclic peptides.

Discussion

At this thesis it was evaluated that TDP-43 protein was indeed aggregation prone by the usage of sedimentation and turbidity assays, as it had already been described at Johnson 2009 [19]. Furthermore, at sedimentation assays the Q331K TDP-43-GFP started to create bands of oligomers more rapidly than the WT TDP-43-GFP did. Although, the sedimentation assay was conducted once and there was already protein aggregated from the beginning of the experiment this result could evaluate the literature describing Q331K variant as more aggregation prone than WT [19]. Moreover, as the experiments revealed the pentapeptides selected (TAFDR and TTYAR) did not inhibit the aggregation.

Due to time limitations in this thesis some of the problems that occurred could not be surpassed. Specifically, molecular cloning was successful only at WT and Q331K variant and subcloning only at G294A and M337V variants. Furthermore, there were difficulties in protein purification that perforce stopped the experiments in crucial time points. Akta Ni-Nta column could not work with a buffer with such a high concentration in urea at 4°C, as crystals were created inside the column. Protein was also stuck in Ni-Nta gravity flow column as well as, at HiLoad 16/600 Superdex 200 prep grade column. Additionally, in aggregation assays the protein was aggregated from the beginning of the experiments so the turbidity curve and the results of the sedimentation were not ideal. Finally, protein aggregates tend to adhere in microplate well's periphery which leads to problems at absorbance meausurements.

In future work, ameliorations of these procedures could occur and experiments could continue in many directions. Firstly, all variants of TDP-43 should be manufactured both with plain, as well as with fused with GFP protein. Secondly, an optimization of the protein purification procedure is necessary so that the protein is purified in high yields without losing great amounts within the column. Thirdly, aggregation procedure should be ameliorated, too. Low-binding microplates that do not absorb the protein at their periphery should be preferred. At the beginning of the aggregation assays there should be zero aggregation and probably more time is needed. Finally, more experiments should be based in the soluble fraction of protein as it is easier to manipulate.

Bibliography

[1] I. Matis *et al.*, “An integrated bacterial system for the discovery of chemical rescuers of disease-associated protein misfolding,” *Nat. Biomed. Eng.*, vol. 1, no. 10, pp. 838–852, Oct. 2017, doi: 10.1038/S41551-017-0144-3.

[2] C. A *et al.*, “Prognostic factors in ALS: A critical review,” *Amyotroph. Lateral Scler.*, vol. 10, no. 5–6, pp. 310–323, 2009, doi: 10.3109/17482960802566824.

[3] E. Broussalis *et al.*, “Late age onset of amyotrophic lateral sclerosis is often not considered in elderly people,” *Acta Neurol. Scand.*, vol. 137, no. 3, pp. 329–334, 2018, doi: 10.1111/ane.12869.

[4] C. A *et al.*, “Prevalence of SOD1 mutations in the Italian ALS population.,” *Neurology*, vol. 70, no. 7, pp. 533–537, Feb. 2008, doi: 10.1212/01.WNL.0000299187.90432.3F.

[5] M. Neumann *et al.*, “Ubiquitinated TDP-43 in frontotemporal lobar degeneration and amyotrophic lateral sclerosis,” *Science (80-.).*, vol. 314, no. 5796, pp. 130–133, Oct. 2006, doi: 10.1126/SCIENCE.1134108.

[6] E. Bentmann, M. Neumann, S. Tahirovic, R. Rodde, D. Dormann, and C. Haass, “Requirements for Stress Granule Recruitment of Fused in Sarcoma (FUS) and TAR DNA-binding Protein of 43 kDa (TDP-43),” *J. Biol. Chem.*, vol. 287, no. 27, p. 23079, Jun. 2012, doi: 10.1074/JBC.M111.328757.

[7] R. H. Tan, Y. D. Ke, L. M. Ittner, and G. M. Halliday, “ALS/FTLD: experimental models and reality,” *Acta Neuropathol.*, vol. 133, no. 2, pp. 177–196, Feb. 2017, doi: 10.1007/S00401-016-1666-6.

[8] H. M *et al.*, “Molecular dissection of TDP-43 proteinopathies,” *J. Mol. Neurosci.*, vol. 45, no. 3, pp. 480–485, Nov. 2011, doi: 10.1007/S12031-011-9571-X.

[9] A. Prasad, V. Bharathi, V. Sivalingam, A. Girdhar, and B. K. Patel, “Molecular Mechanisms of TDP-43 Misfolding and Pathology in Amyotrophic Lateral Sclerosis,” *Front. Mol. Neurosci.*, vol. 0, p. 25, Feb. 2019, doi: 10.3389/FNMOL.2019.00025.

[10] L. SC, P. M, and C. DW, “Converging mechanisms in ALS and FTD: disrupted RNA and protein homeostasis,” *Neuron*, vol. 79,

no. 3, pp. 416–438, Aug. 2013, doi: 10.1016/J.NEURON.2013.07.033.

[11] W. Li *et al.*, “Heat shock-induced phosphorylation of TAR DNA-binding protein 43 (TDP-43) by MAPK/ERK kinase regulates TDP-43 function,” *J. Biol. Chem.*, vol. 292, no. 12, pp. 5089–5100, 2017, doi: 10.1074/jbc.M116.753913.

[12] W. Huang *et al.*, “TDP-43: From Alzheimer’s Disease to Limbic-Predominant Age-Related TDP-43 Encephalopathy,” *Front. Mol. Neurosci.*, vol. 13, no. February, pp. 1–7, 2020, doi: 10.3389/fnmol.2020.00026.

[13] L. Lim, Y. Wei, Y. Lu, and J. Song, “ALS-Causing Mutations Significantly Perturb the Self-Assembly and Interaction with Nucleic Acid of the Intrinsically Disordered Prion-Like Domain of TDP-43,” *PLoS Biol.*, vol. 14, no. 1, pp. 1–35, 2016, doi: 10.1371/journal.pbio.1002338.

[14] A. YM *et al.*, “Structural determinants of the cellular localization and shuttling of TDP-43,” *J. Cell Sci.*, vol. 121, no. Pt 22, pp. 3778–3785, Nov. 2008, doi: 10.1242/JCS.038950.

[15] C. Capitini *et al.*, “TDP-43 inclusion bodies formed in bacteria are structurally amorphous, non-amyloid and inherently toxic to neuroblastoma cells,” *PLoS One*, vol. 9, no. 1, pp. 1–14, 2014, doi: 10.1371/journal.pone.0086720.

[16] M. V. Vega *et al.*, “Isolation and characterization of soluble human full-length TDP-43 associated with neurodegeneration,” *FASEB J.*, vol. 33, no. 10, pp. 10780–10793, 2019, doi: 10.1096/fj.201900474R.

[17] T. Afroz *et al.*, “Functional and dynamic polymerization of the ALS-linked protein TDP-43 antagonizes its pathologic aggregation,” *Nat. Commun.*, vol. 8, no. 1, pp. 1–14, 2017, doi: 10.1038/s41467-017-00062-0.

[18] D. SenGupta, “RNA-Binding Domains in Proteins,” *Brenner’s Encycl. Genet. Second Ed.*, pp. 274–276, Jan. 2013, doi: 10.1016/B978-0-12-374984-0.01356-5.

[19] B. S. Johnson, D. Snead, J. J. Lee, J. M. McCaffery, J. Shorter, and A. D. Gitler, “TDP-43 is intrinsically aggregation-prone, and amyotrophic lateral sclerosis-linked mutations accelerate aggregation and increase toxicity,” *J. Biol. Chem.*, vol. 284, no. 30, pp. 20329–20339, 2009, doi: 10.1074/jbc.M109.010264.

[20] C. S. Sun *et al.*, “The Influence of Pathological Mutations and Proline Substitutions in TDP-43 Glycine-Rich Peptides on Its Amyloid Properties and Cellular Toxicity,” *PLoS One*, vol. 9, no.

8, p. e103644, Aug. 2014, doi: 10.1371/JOURNAL.PONE.0103644.

[21] A. K. H. Chen *et al.*, “Induction of amyloid fibrils by the C-terminal fragments of TDP-43 in amyotrophic lateral sclerosis,” *J. Am. Chem. Soc.*, vol. 132, no. 4, pp. 1186–1187, Feb. 2010, doi: 10.1021/JA9066207.

[22] E. S. Arnold *et al.*, “ALS-linked TDP-43 mutations produce aberrant RNA splicing and adult-onset motor neuron disease without aggregation or loss of nuclear TDP-43,” *Proc. Natl. Acad. Sci. U. S. A.*, vol. 110, no. 8, Feb. 2013, doi: 10.1073/PNAS.1222809110.

[23] J. C. Mitchell *et al.*, “Wild type human TDP-43 potentiates ALS-linked mutant TDP-43 driven progressive motor and cortical neuron degeneration with pathological features of ALS,” *Acta Neuropathol. Commun.*, vol. 3, p. 36, 2015, doi: 10.1186/s40478-015-0212-4.

[24] V. B. Tripathi, P. Baskaran, C. E. Shaw, and S. Guthrie, “Tar DNA-binding protein-43 (TDP-43) regulates axon growth in vitro and in vivo,” *Neurobiol. Dis.*, vol. 65, pp. 25–34, 2014, doi: 10.1016/j.nbd.2014.01.004.

[25] S. D’Alton, M. Altshuler, A. Cannon, D. W. Dickson, L. Petrucelli, and J. Lewis, “Divergent phenotypes in mutant TDP-43 transgenic mice highlight potential confounds in TDP-43 transgenic modeling,” *PLoS One*, vol. 9, no. 1, 2014, doi: 10.1371/journal.pone.0086513.

[26] D. CM, C. B, S. CF, J. BA, H. J, and Y. G, “TDP-43 aggregation in neurodegeneration: are stress granules the key?,” *Brain Res.*, vol. 1462, pp. 16–25, Jun. 2012, doi: 10.1016/J.BRAINRES.2012.02.032.

[27] L. YR, K. OD, S. J, and G. AD, “Stress granules as crucibles of ALS pathogenesis,” *J. Cell Biol.*, vol. 201, no. 3, pp. 361–372, Apr. 2013, doi: 10.1083/JCB.201302044.

[28] N. M, R. R, R. S, B. M, K. HA, and M. IR, “A new subtype of frontotemporal lobar degeneration with FUS pathology,” *Brain*, vol. 132, no. Pt 11, pp. 2922–2931, 2009, doi: 10.1093/BRAIN/AWP214.

[29] I. LM *et al.*, “Enrichment of C-terminal fragments in TAR DNA-binding protein-43 cytoplasmic inclusions in brain but not in spinal cord of frontotemporal lobar degeneration and amyotrophic lateral sclerosis,” *Am. J. Pathol.*, vol. 173, no. 1, pp. 182–194, 2008, doi: 10.2353/AJPATH.2008.080003.

[30] B. C. Kraemer *et al.*, “Loss of murine TDP-43 disrupts motor function and plays an essential role in embryogenesis,” *Acta Neuropathol.*, vol. 119, no. 4, p. 409, Apr. 2010, doi: 10.1007/S00401-010-0659-0.

[31] T. Niccoli, L. Partridge, and A. M. Isaacs, “Ageing as a risk factor for ALS/FTD,” *Hum. Mol. Genet.*, vol. 26, no. R2, pp. R105–R113, Oct. 2017, doi: 10.1093/HMG/DDX247.

[32] A. C. Anderson, “The Process of Structure-Based Drug Design,” *Chem. Biol.*, vol. 10, no. 9, pp. 787–797, Sep. 2003, doi: 10.1016/J.CHEMBIOL.2003.09.002.

[33] M. S. Attene-Ramos, C. P. Austin, and M. Xia, “High Throughput Screening,” *Encycl. Toxicol. Third Ed.*, pp. 916–917, Jan. 2014, doi: 10.1016/B978-0-12-386454-3.00209-8.

[34] A. Sergeeva, M. G. Kolonin, J. J. Molldrem, R. Pasqualini, and W. Arap, “Display technologies: Application for the discovery of drug and gene delivery agents,” *Adv. Drug Deliv. Rev.*, vol. 58, no. 15, p. 1622, Dec. 2006, doi: 10.1016/J.ADDR.2006.09.018.

[35] S. H. Joo, “Cyclic peptides as therapeutic agents and biochemical tools,” *Biomol. Ther.*, vol. 20, no. 1, pp. 19–26, Jan. 2012, doi: 10.4062/BIOMOLTHER.2012.20.1.019.

[36] C. P. Scott, E. Abel-Santos, M. Wall, D. C. Wahnon, and S. J. Benkovic, “Production of cyclic peptides and proteins in vivo,” *Proc. Natl. Acad. Sci. U. S. A.*, vol. 96, no. 24, pp. 13638–13643, Nov. 1999, doi: 10.1073/PNAS.96.24.13638.

[37] D. C. Delivoria *et al.*, “Bacterial production and direct functional screening of expanded molecular libraries for discovering inhibitors of protein aggregation,” *Sci. Adv.*, vol. 5, no. 10, pp. 14–17, 2019, doi: 10.1126/sciadv.aax5108.

[38] “Storing Bacterial Samples for Optimal Viability - GR,” Accessed: Nov. 12, 2021. [Online]. Available: //www.thermofisher.com/tr/en/home/industrial/microbiology/microbiology-learning-center/storing-bacterial-samples-optimal-viability.html.

[39] J. A. Bornhorst and J. J. Falke, “[16] Purification of Proteins Using Polyhistidine Affinity Tags,” *Methods Enzymol.*, vol. 326, p. 245, 2000, doi: 10.1016/s0076-6879(00)26058-8.

[40] C. J. Chirayil, J. Abraham, R. K. Mishra, S. C. George, and S. Thomas, “Instrumental Techniques for the Characterization of Nanoparticles,” *Therm. Rheol. Meas. Tech. Nanomater. Charact.*, vol. 3, pp. 1–36, Jan. 2017, doi: 10.1016/B978-0-323-46139-9.00001-3.

[41] “BL21(DE3) Competent Cells.” https://www.thermofisher.com/order/catalog/product/EC0114?SID=srch-srp-EC0114#/EC0114?SID=srch-srp-EC0114 (accessed Oct. 22, 2021).

[42] “BL21(DE3) Competent E. coli | NEB.” https://www.neb.com/products/c2527-bl21de3-competent-e-coli#Product Information_Properties & Usage (accessed Oct. 22, 2021).

[43] “Rosetta™(DE3) Competent Cells - Novagen | 70954.” https://www.merckmillipore.com/INTL/en/product/RosettaDE3-Competent-Cells-Novagen,EMD_BIO-70954 (accessed Oct. 22, 2021).

[44] N. H. Tolia and L. Joshua-Tor, “Strategies for protein coexpression in Escherichia coli,” *Nat. Methods*, vol. 3, no. 1, pp. 55–64, 2006, doi: 10.1038/nmeth0106-55.

[45] J. Jin, O. D. Okagu, A. E. G. A. Yagoub, and C. C. Udenigwe, “Effects of sonication on the in vitro digestibility and structural properties of buckwheat protein isolates,” *Ultrason. Sonochem.*, vol. 70, no. July 2020, p. 105348, 2021, doi: 10.1016/j.ultsonch.2020.105348.

[46] M. M. Carrió and A. Villaverde, “Protein aggregation as bacterial inclusion bodies is reversible,” *FEBS Lett.*, vol. 489, no. 1, pp. 29–33, Jan. 2001, doi: 10.1016/S0014-5793(01)02073-7.

[47] J. L. Cleland, S. E. Builder, J. R. Swartz, M. Winkler, J. Y. Chang, and D. I. C. Wang, “Polyethylene glycol enhanced protein refolding,” *Biotechnology. (N. Y).*, vol. 10, no. 9, pp. 1013–1019, 1992, doi: 10.1038/NBT0992-1013.

[48] G. Graziano, “How does sucrose stabilize the native state of globular proteins?,” *Int. J. Biol. Macromol.*, vol. 50, no. 1, pp. 230–235, Jan. 2012, doi: 10.1016/J.IJBIOMAC.2011.10.025.

[49] G. CA, S. MP, and S. HC, “Increasing the yield of soluble recombinant protein expressed in E. coli by induction during late log phase,” *Biotechniques*, vol. 34, no. 3, pp. 524–530, Mar. 2003, doi: 10.2144/03343ST04.