EXPRESSION AND PURIFICATION OF THE RECOMBINANT HUMAN TORSIN A

by

YAN WU

B.S., China Pharmaceutical University, China, 2010

A THESIS

submitted in partial fulfillment of the requirements for the degree

MASTER OF SCIENCE

Graduate Biochemistry Group
College of Arts and Sciences

KANSAS STATE UNIVERSITY
Manhattan, Kansas

2012

Approved by:

Major Professor
Michal Zolkiewski

Abstract

Early-onset dystonia (EOTD, also known as DYT1 or Oppenheim's dystonia is the most severe and common form of hereditary dystonia, a neurological disorder characterized by abnormalities in the control of movement. It is linked to the deletion of a single GAG codon in the gene DYT1 that leads to the loss of a single glutamic acid residue in the C-terminal region of the protein torsinA (ΔE-torsinA). It is not known how the GAG deletion alters the torsinA structure and function.

In this thesis, the expression and purification of recombinant torsinA variants from *E. coli* is reported. Wild type torsinA is not soluble after its expression in *E. coli*, possibly due to misfolding caused by cysteine. We produced Cys-less torsinA, and established a purification procedure to produce this mutant torsinA. Furthermore, because of the critical role likely to be played by the C-terminal domain of torsinA that contains the glutamate deletion, we produced fragments encoding the C-terminal domain of torsinA, and attempted to purify it. However, we failed to obtain appreciable amount of active proteins by both of the strategies. A novel SUMO fusion technology was also used in this study. We demonstrated that SUMO, when fused with torsinA variants, was able to enhance its expression and solubility in *E. coli*. A satisfactory yield of the fusion protein was successfully purified. Once we get appreciable quantities of folded torsinA variants, it is our future goal to study their function by using biochemical and high-resolution structural approaches.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

Ala	alanine
Arg	arginine
Asn	asparagine
Asp	aspartic acid
Cys	cysteine
Gln	glutamine
Glu	glutamic acid
Gly	glycine
h	hour
His	histidine
kDa	kilodalton
IIe	isoleucine
Leu	leucine
Lys	lysine
MS	mass spectrometry
Met	methionine
Min	minutes
PAGE	polyacrylamide gel electrophoresis
PCR	polymerase chain reaction
Phe	phenylalanine
Pro	proline
RT	room temperature
S	seconds
Ser	serine
SDS	sodium dodecyl sulfate
Thr	threonine
Trp	tryptophan
Tyr	tyrosine
Val	valine

Acknowledgements

First and foremost, I would like to express my sincere gratitude to my advisor, Dr. Michal Zolkiewski, for his profound insights, constructive advice and impeccable patience. I am greatly indebted for all the time he spent helping me learn how to be a scientist.

I would also like to thank Dr. S.Muthukrishnan and Dr. Jianhan Chen for their willingness to serve on my committee as well as for their valuable research suggestions.

My appreciation is also extended to all the past and present Zolkiewski lab members, Hui-Chuan Wu (Arleen), Ting Zhang, Sam Molina and Fabrice Ngansop. The group has been a source of friendships as well as good advice and collaboration. Appreciation is especially expressed to Arleen and Ting for being a patient teacher during my first year in the lab, and for inspirational discussions with me regarding my experiments. From both of them, I have learned so much and appreciate their willingness to help.

I owe a debt of gratitude to my friends, both inside and outside of Department of Biochemistry. I am grateful for time spent with them. A special thanks to Zeyu Peng (Sara), Yue Qi, Shruti Shrestha, Pallabi Gupta, Zhen Li, who are always here for me whenever I need support.

Finally, I would like to thank my family for all their love and encouragement. Thank you!

Dedication

This thesis is dedicated to my dear parents, Xiaoping Wu and Juan Yuan, who raised me with a love of science and always helped me to pursue my dreams.

Chapter 1 – Introduction

Dystonia

Dystonia has been defined by the Scientific Advisory Board of Dystonia Medical Research Foundation as a syndrome of sustained involuntary muscle contractions, frequently causing twisting and repetitive movements or abnormal postures (1988). Despite early descriptions of dystonia by Oppenheim in 1911; it took more than half a century before physicians accepted that this syndrome was due to brain disease (Marsden, 1976).

Dystonia does not discriminate, and it affects men, women, children of all ages and racial backgrounds. It is the third most common movement disorder behind essential tremor and Parkinson's disease, affecting more than 300,000 people in North America (according to the Dystonia Medical Research Foundation).

Despite the paucity of knowledge about a specific cause of dystonic disorders, the symptomatic treatment strategies have been substantially improved. These treatment options include physical and occupational therapy, oral medications, intramuscular injection of botulinum toxins, and neurosurgical interventions (Jankovic, 2006).

Classification

Clinically, dystonia can be classified in three different ways: (1) by etiology; (2) by distribution in the body; (3) by age of onset (de Carvalho Aguiar and Ozelius, 2002).

From the etiologic standpoint, dystonia is classified as "primary" and " secondary". Primary dystonia develops in the absence of any associated diseases and shows no other neurological symptoms, except tremor and myoclonus (Breakefield et al., 2008). Primary dystonia can be further divided into two groups: Early-onset and late onset. Early-onset torsion dystonia (EOTD, also known as DYT1 or Oppenheim's dystonia) is most common and severe form of primary dystonia, which frequently has a genetic basis. The symptoms of EOTD usually occur first between the age of 5 and 28. Onset before 4 years or after 28 years is rare. It has onset in a limb and can progress to affect multiple body parts (Tarsy and Simon, 2006; Zolkiewski and Wu, 2011). Late-onset primary dystonia develops after the age of 26, and usually is focal (Breakefield et al., 2008).

Secondary dystonia is a large group of disorders in which dystonic symptoms rise from other causes, such as injury, toxins, strokes or trauma. When dystonia is secondary to a hereditary neurological disorder, additional neurological abnormalities are likely to be present (Geyer and Bressman, 2006; Muller, 2006). More than 42 hereditary neurological disorders are in this case, including Huntington's disease, juvenile Parkinson's disease, etc (Breakefield et al., 2008)

Dystonia can also be classified by distribution of the affected region. In focal dystonia, only one single body part is affected (e.g. "blepharospasm" for dystonia affecting eyes; "cervical" for dystonia affecting neck). Cervical dystonia is the most common form of the focal dystonia. When two or more adjacent body areas are affected, we use the term "segmental dystonia". In Multifocal disease, two or more non-adjacent body parts are involved. Hemidystonia dystonia affects arm or leg on same side of the body, whereas generalised dystonia is associated with abnormal movements at least one other side of the body (Aguiar and Ozelius, 2002; Chen et al., 2008; Geyer and Bressman, 2006)

Early-onset dystonia

Early-onset torsion dystonia (or DYT1 dystonia) is the most common and severe form of hereditary dystonia. This disorder is reported to have estimated prevalence of 3.4:100,000 among general population worldwide. However, it occurs most frequently in people of Ashkenazi Jewish heritage, affecting 1 in 3,000 to 9,000 people in this population (Kamm, 2006). The high frequency of this disease is linked to a founder mutation that introduced into the AJ population around 350 years ago, and probably originated in Lithuania or Byelorussia (Bressman et al., 1995; Risch et al., 1995).

The EOTD-linked gene TOR1A (also known as DYT1), which is responsible for both Ashkenazic Jewish (AJ) and non-Jewish family, has been identified by Ozelius and colleagues (1997). TOR1A is expressed ubiquitously in human tissues, and encodes a protein of 332 amino acids, which is named torsinA. Most cases of EOTD are associated with a single GAG deletion in the coding sequence of TOR1A gene, which leads to the loss of a glutamate from a pair of glutamates at position 302/303 (ΔE302/303) (Ozelius et al., 1997). This mutation in the TOR1A gene is inherited in an autosomal dominant pattern. Only 30 to 40 percent people who inherit this mutation will develop symptoms of early-onset primary dystonia. It has been suggested that this

disease was a potential consequence of additional factors, such as environmental factors (Zolkiewski, 2011).

AAA+ protein superfamily

AAA proteins (ATPases Associated with diverse cellular Activities) constitute a superfamily, which can be found in all kingdoms of living organisms. They participate in diverse cellular processes, such as protein degradation and DNA replication in all organisms, membrane fusion the movement of microtubule motors in eukaryotes (Hanson and Whiteheart, 2005). The most characteristic feature of AAA+ superfamily is 200–250 amino-acid ATP-binding domain (AAA domain), which contains four conserved motifs: Walker A is involved in nucleotide binding; Walker B is important for ATPase activity; sensor 1 is essential for nucleotide hydrolysis and substrate binding; sensor 2 is crucial for nucleotide binding and hydrolysis (Hanson and Whiteheart, 2005; Neuwald et al., 1999). AAA domain consists of two subdomains: an N-terminal nucleotide-binding subdomain and a smaller C-terminal helical subdomain. The larger N-terminal subdomain is αβα fold with five parallel β sheets in the core. The C-terminal subdomain is mostly α helix that is one of main features of AAA proteins different from other nucleotide-binding proteins (Hanson and Whiteheart, 2005). AAA proteins function as oligomers, in most cases by forming hexametric rings, which links to the nucleotide-driving conformational changes. However, the extent to how the nucleotide binding and hydrolysis are coordinated within hexametric ring structures for their biological functions is not well defined yet (Joly et al., 2006).

ClpB

ClpB belongs to the Clp ATPases family, which also includes ClpA,ClpE, ClpL and ClpX. ClpB is found not only in prokaryotes and in eukaryotic organelles, but also in the cytosol of plants and yeast. ClpB has a remarkable ability to disaggregate and refold the aggregated proteins, with the cooperation of the DnaK chaperone system (DnaK, DnaJ, and GrpE; termed KJE) (Zolkiewski, 2006).

TorsinA and other torsin proteins in the AAA+ superfamily

Sequence analysis showed that TorsinA belongs to the AAA+ superfamily of ATPases (Giles et al., 2008). Based on the properties of AAA proteins, TorsinA is likely to function as

oligomers, and may use the nucleotide-driven conformational changes to promote movement within the subunit (Breakefield et al., 2001). However, currently torsinA oligomerization is still a controversial topic.

Four homologous mammalian torsins have been found: tosinA, torsinB (84% similar with torsinA), torsin2, and torsin3 (62% and 50%, respectively) (Jungwirth et al., 2010). All four torsins are found in mouse and chicken, and torsin2 and torsin3 are also present in zebrafish (Breakefield et al., 2001). Related coding sequence have also been found in Drosophila (torp4a), nematodes (OOC-5, Y37A1B.12, and Y37A1B.13), but not found in prokaryotes, fungi, yeast or plants, which suggests that torsins are specifically related to animal physiology (Breakefield et al., 2001; Zolkiewski, 2011).

The four mammalian torsin proteins are homologous in the AAA domain which contains: Walker A, Walker B, sensor1 and sensor 2 motif. Both of torsinA and torsinB have hydrophobic segment following the signal sequence. Torsin2 lacks this distinguishing segment, whereas torsin3 possess the hydrophilic segment. These features may imply the difference in protein solubility. The hydrophobic domains of torsinA and torsinB anchor them to the ER membrane, whereas torsin2 and torsin3 may be soluble, due to the lack of hydrophobic segment (Zolkiewski, 2011).

Structural model of torsinA

The 20 N-terminal amino acids in torsinA is a signal sequence, which targets the nascent protein to the translocation machinery in the ER membrane, and then is removed by signal peptidases. The amino acids 21-40 form a hydrophobic segment, which is proposed to anchor torsinA in ER membrane. Previous study showed that torsinA behaved as a type II membrane protein, and this orientation within the translocon stabilized the final structure of torsinA (Liu et al., 2003). However, Callan et al. (2007) proposed that torsinA was a peripherally associated with ER membrane, most likely via binding to an integral membrane protein, instead of being an integral membrane protein as we thought before. However, these findings failed to show a specific ER subdomain where torsinA located to (Granata et al., 2009).

Unfortunately, there is no crystal structure of torsinA available to date. Based on a multiple alignment of torsinA in a family with the closest homologue of C-terminal AAA+ domains of ClpA from Escherichia coli, and T. thermophilus ClpB, an initial prediction of

human torsinA structure was proposed (Callan et al., 2007; Zhu et al., 2008). Both of two proposed models suggested that torsinA possessed two subdomains, a N-terminal α/β subdomain from residue 41-271, and a C-terminal α-helical subdomain from residue 272-332, which contained the dystonia-linked mutation (Figure 1) (Zolkiewski, 2011). Importantly, a unique feature of torsinA is six conserved cysteine residues (Cys44, Cys49, Cys50, Cys162, Cys280, Cys319), which are not found in ClpA, ClpB, or other AAA+ ATPases (Zhu et al., 2008). The structural model also predicted that there are at least two disulfide bonds, one was located in the large N-terminal domain, and another one was in the small C-terminal domain. Zhu et al. (2008) proposed that the Cys 319 within the sensor II motif was linked to Cys 280 in the C-terminal domain , and it might function as a redox sensor to regulate nucleotide binding.

Expression pattern of torsinA

In human brains, in-situ hybridization and immunohistochemistry studies demonstrated that torsinA protein was highly expressed in several parts of brain region, including hippocampus, basal ganglia, substantia nigra and cerebellum (Siegert et al., 2005). No discernible differences in torsinA staining and no neuropathological abnormalities were found in the neostriatum or midbrain of dystonia brains, which suggested that dysfunction rather than neurodegeneration was the cause of DYT1 dystonia (Rostasy et al., 2003). High expression of torsinA was also found in peripheral tissues such as kidney, liver and muscle, but not found in spleen (Shashidharan et al., 2000).

Xiao et al. (2004) examined torsinA expression in rats from the embryonic period through adulthood. At postnatal day 14, the highest expression of torsinA was noted in both striatal cholinergic interneurons and cerebellar Purkinje cells. This result might suggest that torsinA plays an important role in postnatal maturation in the brain.

Subcellular localization of torsinA and dystonia-linked mutant form (torsinAΔE)

Several early studies showed that torsinA was a membrane-associated enzyme that resided in the endoplasmic reticulum (ER) lumen, by using a variety of biochemical techniques, such as co-localization with ER proteins (Hewett et al., 2003b), subcellular fractionation (Kustedjo et al., 2003). How does torsinA retain in ER lumen and escape fluxing out of the ER into the secretory pathway? Vander Heyden et al. (2011) proposed that the hydrophobic N-terminal domain (residue 21-43) was a monotopic membrane-associating domain that was

directly responsible for static retention in the ER. Upon over-expression, torsinA was also found in nuclear envelope (NE), where it appeared to interact with substrates. This distribution of torsinA is unique in two ways: 1) torsinA is the only known non-transmembrane protein that resides in the NE. 2) significant amounts of protein resides in the ER and the NE at the same time (Goodchild and Dauer, 2004).

Dystonia-linked mutant form (torsinAΔE) exhibited a distinct subcellular location to a large spheroid structures surrounded with membrane whorls, which probably were derived from the NE (Hewett et al., 2000; Kustedjo et al., 2003). In contrast to these data, Walker et al. (2002) showed that there was no evidence of different localizations of torsinA by using brains from normal and DYT1 patients. Furthermore, in Caenorhabditis elegans, there was no apparent change in nematode TOR-2 mutant protein localization compared with the wild type protein (Caldwell et al., 2003). So the relationship between dystonia pathology and intracelluar mislocalization of the mutant protein is still unknown.

TorsinAΔ323 – 8 indicates a 18 bp deletion in the torsinA gene and it is found in one family associated with DYT1 dystonia and myoclonus (Giles et al., 2008). O'farrell et al. (2002) reported that this mutant protein showed similar subcellular localization with wild-type torsinA, which indicated that intracellular inclusions formation is relevant only to torsinAΔE, and torsinAΔ323 – 8 mutation associated disease had different mechanism.

TorsinA localization in neuronal cells is different from non-neurons; It shows neuronal cell-type-specific enrichment in the NE. Giles et al. (2008) also indicated that dystonia-linked mutation induced translocation of torsinA from the ER to the NE in neuronal cells but not in non- neuronal cells. Based on these results, it may explain the neuronal cell-specific phenotype of EOTD.

Biochemical properties of torsinA

The membership of the torsins in the AAA family offers us some clues to the possible functions of torsinA. Based on the fact that members of the AAA+ family are typically chaperone proteins, one can hypothesize that torsinA is a chaperone with specific targets in the secretory pathway. McLean et al. (2002) demonstrated that torsinA and HSP molecular chaperones colocalized with α-synuclein in Lewy bodies, and overexpression of torsinA and HSPs could suppress α-synuclein aggregation in cellular model. It suggestd that torsinA might

exhibit chaperone-like activity and that dystonia-linked mutant torsinA lost this function. In support of these data, another study showed that C. elegans torsin-related protein TOR-2 could suppress polyglutamine-induced protein aggregation (Caldwell et al., 2003). However, torsinA did not appear to behave as a typical ER chaperone protein, and it was not found to be up-regulated in response to a variety of stresses, including many ER stresses(Hewett et al., 2003a).

Previous studies regarding torsinA oligomerization have been controversial. An early study showed that torsinA forms monomers in solution (Kustedjo et al., 2003). In contrast, Giles et al. (2008) indicated that wild type torsinA was capable of self-associating into an oligomeric complex, and this process did not require carboxyl- terminal region, which was necessary for other AAA+ protein oligomerization. They also reported that the dystonia-linked mutation did not disrupt the ability of self-association. This result was consistent with data from Torres et al. (2004), which indicated the property of oligomerization was enhanced in torsinAΔE.

Konakova and Pulst (2005a) provided the first evidence for a role of torsinA as an active ATPase and suggested that the dystonia-linked mutation might impair its ATPase activity. However, the significance of this study was unclear, as in this case torsinA is fused to other large domains, and the folding state after purification from E. coli was not determined (Zolkiewski, 2011).

Recent studies have addressed the question regarding how torsinA and torsinAΔE are degraded in cells. These findings suggested that torsinA was a very stable protein with a half-time of ≈3.5 days, whose turnover was mediated by basal autophagy. TorsinAΔE reduced half-time to ≈18 h, and was degraded by both proteasome-mediated ER-associated degradation (ERAD) and macroautophagy–lysosome pathways (Giles et al., 2008; Gordon and Gonzalez-Alegre, 2008). Furthermore, results from Giles et al. (2009b) raised the possibility that monomeric form of torsinA mutant proteins was cleared by ERAD, whereas the oligomeric and aggregated forms of torsinA mutant proteins were cleared by ER stress-induced autophagy. In in vitro studies, the dystonia-linked mutation did not cause proteins misfolded (Kustedjo et al., 2003). However, it was still unclear why torsinAΔE was specifically recognized by ERAD, a more selective pathway than autophagy, as an abnormal protein.

Substrate Interactions with TorsinA

Since the distinct localizations of torsinA and torsinAΔE have been established, efforts have been made to discover torsinA substrates and characterize the interactions, which may provide the insight into DYT1 dystonia disease pathogenesis. TorsinA have been shown to bind the inner nuclear membrane protein lamina-associated polypeptide 1 (LAP1), and lumenal domain-like LAP1 (LULL1) in the ER (Goodchild and Dauer, 2005; Hewett et al., 2006). A recent study indicated that LULL1 dynamically regulated the distribution of torsinA between the ER and the NE, and dystonia-linked mutation was less efficient in enacting this process(Vander Heyden et al., 2009).

A novel protein named printor (protein interactor of torsinA) was reported to selectively interact with wild type torsinA instead of torsinAΔE. Interestingly, printor selectively bound to the ATP-free form but not to the ATP-bound form of torsinA, which might support a role of printor as a cofactor rather than a substrate of torsinA (Giles et al., 2009a).

Although torsinA was generally thought to be an ER-lumen protein, a significant pool of torsinA has been shown to exhibit the AAA+ domain facing the cytoplasm (Kamm et al., 2004). TorsinA has been reported to bind several cytosolic proteins, such as snapin and kinesin light chain (Granata et al., 2008; Kamm et al., 2004). A recent study reported an association between torsinA and nesprin-3, a resident of the outer nuclear membrane implicated in connecting the nucleus to the cytoskeleton (Nery et al., 2008). This torsinA-nesprin-cytoskeletal association was supported by the interaction between the C-terminal region of torsinA and KASH domain of nesprin-3, which was the only portion of nesprin-3 located in the lumen of the NE. Importantly, compared with torsinA, torsinAΔE was found to bind more nesprin-3 tightly, which might be responsible for apparent trapping of torsinAΔE in the NE (Nery et al., 2008; Zolkiewski, 2011). Furthermore, in the absence of torsinA, a wound-healing assay shows that nuclear polarization and cell migration were delayed in embryonic fibroblasts (Nery et al., 2008). This finding supported a role for torsinA in regulating dynamic interactions between the KASH domains of nesprins and their protein partners in the lumen of ER, and this interaction contributed to intracellular movement of the nucleus during cell migration (Gomes et al., 2005; Nery et al., 2008).

Biological models for DYT1 dystonia

Several mouse lines have been published to model DYT1 dystonia: Transgenic mice overexpressing either human ΔE or wt torsinA (Sharma et al., 2005; Shashidharan et al., 2005); Dyt1ΔGAG knock-in heterozygous mice (Dang et al., 2005); Dyt1ΔGAG knock- down mice (Dang et al., 2006; Goodchild et al., 2005). In the DYT1 ΔGAG knock-in mice model, a GAA trinucleotide deletion was made to mimic the GAG deletion in human DYT1 mutation, which was supposed to encode one of the glutamic acids at the position corresponding to human E302/303 (Dang et al., 2005). Notably, none of these mouse models showed a behavioral phenotype typical of dystonia, although mouse models generated by Shashidharan et al. (2005) exhibited self-clasping behavior suggestive of dystonia (Granata et al., 2009). Such dystonia-like postures, detected in both transgenic and control mice, have to be regarded as a clasping reflex (Lange et al., 2011). In addition, the cortex-specific and striatum-specific Dyt1 conditional knockout mouse have been generated separately (Yokoi et al., 2011; Yokoi et al., 2008). A recent study indicated that the loss of striatal torsinA affected signal transduction pathway through dopamine receptor 2 in the basal ganglia circuit and exhibit motor deficits.(Yokoi et al., 2011).

Non-mammalian models of early-onset dystonia have also been reported. The expression of ΔE HtorA in Drosophila model showed behavioral and ultrastructural defects that involved impaired TGF-β signaling. This raised the possibility that a defect in TGF-β signaling might also underlie early onset torsion dystonia in humans (Koh et al., 2004). Caldwell et al. (2003) represented the first evidence of a role of torsinA in reducing the formation of protein aggregates, and defects in this mechanism might have significant consequences on neuronal dysfunction associated with dystonia.

Therapy for early-onset dystonia

There are few treatments with broad efficacy in the dystonias. The botulinum toxins are the most broadly effective (Jinnah and Hess, 2008). They must be injected directly into the affected body areas, and they are the first choice for the treatment of focal and segmental dystonia. However, with regard to generalized dystonic patients, clinical study showed that botulinum toxin injections might not be very helpful (Kamm, 2006; Roubertie et al., 2000). Several oral medications, such as anticholinergics, are used to treat dystonia mostly by affecting

the neurotransmitter chemicals in the nervous system that regulate the control of movement (Albanese et al., 2006). A recent study showed that ampicillin, which enhanced wild type torsinA activity in Caenorhabditis elegans, might represent a promising new approach toward therapeutic development for EOTD (Cao et al., 2010).

Gonzalez-Alegre et al. (2005) have generated short hairpin RNAs (shRNAs) that mediated allele-specific suppression of torsinA(ΔE) and rescued cells from its dominant-negative effect, and demonstrated significant suppression of endogenous torsinA in mammalian neurons without triggering an interferon response. Based on this result, viral-mediated RNAi might support a novel therapeutic strategy for DYT1 dystonia.

Objectives

Previous studies

The characterization of biochemical and structural properties of torsinA and its dystonia-linked mutant form has been a challenge because of numerous difficulties in producing appreciable amount of active proteins. Initial attempts in *E. coli* were unsuccessful. Kustedjo et al. (2003) made an expression construct in which the first 20 amino acids of torsinA were replaced with a cleavable export sequence followed by a 6×His tag. ~1 mg of purified torsinA protein per liter of cell culture volume was obtained by using Sf9-baculovirus expression system. Konakova and Pulst (2005a) demonstrated that purification of His-NusA-His-S-tagged torsinA was also successful by using a bacterial T7 promoter-driven pET expression system. However, untagged torsinA proteins precipitated after the cleavage of Nus epitope.

Goal of this study

In this dissertation study, the major goal is to produce sufficient amount of native, soluble and fully active protein in *E. coli* for extensive biochemical characterization. For this purpose, we considered two hypotheses: 1). Wild type torsinA is not soluble after its expression in *E. coli*, possibly due to misfolding supported by cysteine; 2) the critical role likely to be played by the C-terminal domain of torsinA that contains the glutamate deletion. To address these issues, we established the purification procedures to produce these mutant torsinA varients. Furthermore, A novel SUMO fusion technology was also used in this study.

Chapter 2 - Methods

Cloning and Site-directed Mutagenesis

DNA fragment encoding the N-terminal domain (residues 41-271) and the C-terminal domain (residues 272-332) of human torsinA with the membrane-anchoring hydrophobic domain deleted (Δ40 torsinA) were produced separately by PCR. Briefly, the plasmid containing the coding regions of human Δ40 torsinA in the pcDNA3 vector was used as the PCR template. For one-tube PCR, 1 ng template was used with 0.2 mM dNTP mix (Stratagene) ,0.5 μM of each forward and reverse primer (IDT Inc.Coralvile, IA), 2.5 U Pfu turbo (Stratagene) and the volume was brought to 50 μl. The cycling parameters were as followings: 95 °C 2 mins, followed by 30 cycles: 30 seconds at 95 °C, 30 seconds at 55 °C, and 1 mins at 72 °C in each cycle. Then further elongated for 5 mins at 72 °C. Reaction products were gel purified by using QIAEX II Gel Extraction Kit (Qiagen). DNA concentration was determined by UV absorbance at 260 nm. PCR products and vectors pProExHTb (Invitrogen) in an N-terminal histidine tagged form (0.5 μg each) were digested with restriction enzymes *BamHI* (Biolabs) and *XbaI* (Biolabs) at 37 °C for 2 h, and purified from agarose gels using QIAEX II Gel Extraction Kit (Qiagen). Ligation was performed with Rapid DNA Ligation Kit (Roche) at 16 °C for overnight. The ligated vectors with inserts were used to transform to DH5α competent cells (Invitrogen). The transformation colonies were screened by GoTaq® Green Master Mix kit (Promega), and then verified by DNA sequencing at the Plant Pathology sequencing facility.

Primers used for amplification of the N-terminal domain of Δ40 torsinA:

Forward:

5’ – CGCGGATCCGCGCGTCTCTACTGCCTC -3’

Reverse:

5’ – TGCTCTAGAGCAGGGGAGGAAGGGAACAA -3’

Primers used for amplification of the C-terminal domain of Δ40 torsinA:

Forward:

5’ – CGCGGATCCGCGCTGGAATACAAACAC – 3’

Reverse:

5’ – TGCTCTAGAGCATCAATCATCGTAGTA – 3’

In the case of the mutant constructs (C280S, C319S and ΔE302/303) produced in this study, site-directed mutagenesis was carried out according to the QuickChange Site-directed Mutagenesis Kit (Stratagene). After digestion of the using QIAprep Spin Miniprep Kit (Qiagen) and confirmed by the DNA sequencing facility at Department of Plant Pathology, Kansas State University.

Primers used for each constructs for mutagenesis are listed in Appendix A. Each primer pair (complementary to each other) for mutagenesis listed contained nucleotide changes (bold) in one codon (underlined). For each mutation, only the primer on the (+) strand is listed.

Recombinant His-tagged torsinA variants expression, solubility test and purification

Plasmids kept in either DH5α or XL-1 blue were isolated and used to transform *Escherichia coli* strain BL21 (DE3) cells (Invitrogen) for protein production (except for C-terminal of Δ40TorsinA (C44S, C49S, C50S, C162S, C280S) and C-terminal of Δ40TorsinA (C44S, C49S, C50S, C162S, C280S, C319S), which Rosetta strains (Novagen) were used.

Protein expression

For expression of recombinant torsinA and torsinA mutants, a single colony harboring the recombinant plasmid was inoculated into 50 mL LB containing 0.1mg/mL ampicillin and incubated overnight at 37 °C with shaking at 225 rpm. Overnight cultures were diluted 1:20 into 1000 mL fresh LB with the same concentration of ampicillin and grown at the same conditions until reaching mid-log phase (OD600 ~0.6). At this time IPTG was added to a final concentration of 1mM. Expression was induced at 37°C for 2 h or 20 °C for 16 h.

Solubility test

Cell pellets were isolated by centrifugation at 4000 × g for 20 min at 4 °C, and the mass of the cells was weighted. Ice-cold lysis buffer (300 mM NaCl, 10 mM imidazole in 50 mM sodium phosphate buffer, pH 8.0) at about 4 ml/g was added to the pellets. The cells were disrupted by sonication on ice for 25 ×15 seconds with a 10 sec cooling between each burst. After sonication, 100 µL lysates was taken out and centrifuged at 14,000 × g for 5 min at 4 °C. The pellets were solubilized in 100 µL cold lysis buffer (300 mM NaCl, 10 mM imidazole in 50 mM sodium phosphate buffer, pH 8.0), and this is the insoluble fraction. The rest of lysates were

subjected to centrifugation at 4000 × g for 20 min at 4 °C. Take out 100 µL supernatant as soluble fraction in solubility test. The remaining soluble fractions were saved for the following purification. Aliquots (15 µL) of each cellular fraction were run on SDS- PAGE followed by Coomassie blue staining.

Purification of His-tagged torsinA variants under native state

The soluble fraction of recombinant C-terminal of Δ40TorsinA and Δ40TorsinA (C44S, C49S, C50S, C162S, C280S, C319S) were purified under native conditions by nickel- affinity chromatography (Ni2+-NTA) or Talon cobalt affinity resin separately. The supernatant was mixed overnight at 4 °C with 50% pre-equilibrated Ni-NTA agarose (Qiagen), or Talon (Clontech). The mixture was then loaded into an empty 1.5 cm diameter column, followed by washing. Impurities were eluted with the lysis buffer that contain imidazole gradient: 10 mM, 50 mM, 100 mM, 250 mM, and 4 mL for each gradient. Fractions (1 mL each) were collected and subjected to SDS-PAGE followed by Coomassie blue staining.

Purification of His-tagged torsinA variants under denaturing state

Purification of denatured recombinant Δ40TorsinA (C44S, C49S, C50S, C162S, C280S, C319S) was done in a similar manner except that the cell extract and buffers loaded on the column contain a denaturant (8 M urea). Purified proteins were analyzed with SDS-PAGE.

Attempted renaturation of purified His-tagged torsinA variants. The fractions containing relatively pure recombinant protein were combined and refolded by step-wise dialyzing against dialysis buffer 1(100 mM NaH_2PO_4, 10 mM Tris-HCl, PH 7.5) with a gradient of decreasing concentration of urea from 7M to 1M. The combined fractions (~3mL) was added to a 0.5-3 mL Slide-A-Lyzer 10,000 MW Dialysis Cassette (Thermo Scientific) and dialyzed in 500 mL dialysis buffer (each dialysis step was for overnight) at 4°C. After final dialysis against 1L dialysis buffer 2 (50 mM Tris-HCl, 1mM EDTA, 20 mM $MgCl_2$, 0.2 M KCl, 10% Glycerol) for overnight at 4°C, the dialyzed sample was centrifuged at 13,000 g for 5min, and the supernatant was collected. The concentration of purified torsinA proteins variants was calculated by measuring absorption at 280 nm and extinction coefficients for all constructs were calculated with the ExPASy ProtParam tool.

Gel Filtration

For further purification, Gel Filtration was carried out by using Superdex 75 column (GE Healthcare). The column was equilibrated with 50 mM Tris-HCl, 1mM EDTA, 20 mM $MgCl_2$, 0.2 M KCl, at a flow rate of 1mL/min. Partially purified C-terminal of Δ40TorsinA (1mL) was loaded at 0.1 mL/min into the system. Because UV-vis absorption detector did not work, so eluted fractions were collected at 30 min interval and time length was 900 min. All samples were analyzed by 16% SDS-PAGE to determine which fractions contained the most C-terminal of Δ40TorsinA and the least contaminating proteins.

The protein reactivation assay

The recombinant Δ40TorsinA (C44S, C49S, C50S, C162S, C280S, C319S) was purified by Ni column under denaturing conditions as described before. In this assay, we took the first fraction (1.53 μM) eluted by 250 mM imidazole buffer as protein sample, and diluted 10-fold in buffer A (50 mM Tris-HCl, pH 7.5, 20 mM $Mg(OAc)_2$,150 mM KCl, 1mM β-mercaptoethanol, 1mM EDTA) containing either no chaperones; 1 μM DnaK, 0.5 μM DnaJ and 0.5 μM GrpE (KJE) with 1.5 μM ClpB; KJE with 1.5 μM ClpB 95 and 0.2 M ATP; or 0.2 M ATP. The control experiment was the protein diluted the same folds in 8 M urea. After incubation at 30 °C for 1 h, aliquots were withdrawn and subjected to Western blot analysis.

The same protein sample diluted 10-fold in buffer A was treated with a series of different concentration of BSA (3.6 μM, 18 μM, 36 μM, 72 μM). The assay was carried as described above.

Expression and purification of torsinA variants by SUMO fusion in *E. coli*

The SUMO fusion protein expression and purification were performed according to Champion™ pET SUMO Protein Expression System kit (Invitrogen) with a few modifications.

Construction of 6xHis-SUMO fusion- Δ40torsinA and 6xHis-SUMO fusion- C-terminal of Δ40 torsinA (C44S, C49S, C50S, C162S, C280S, C319S): for PCR amplification, 20 ng template, 0.5 mM each dNTP mix, 1 μM each primer, and 1 unit Taq Polymerase which adds a single deoxyadenosine (A) to the 3′ ends of PCR products were mixed in reaction buffer to bring the volume to 50 μL. The mixture were heated to 94°C for 2 min, followed by 25 cycles of 94°C for 1 min, 55°C for 1 min, and 72°C for 1 min in each cycle, then further extended for 10 min at

72 °C. Next, the resulting PCR products were ligated into pET SUMO expression vector via T/A ligase technology. The ligation reaction was performed by using a 1:3 molar ratio of vector: insert. 6xHis-SUMO fusion-Δ40torsinA and 6xHis-SUMO fusion-C-terminal of Δ40 torsinA (C44S, C49S, C50S, C162S, C280S, C319S) expression constructs were created. Primers used for amplification of Δ40 torsinA:

Forward: 5' –CGTCTCTACTGCCTCTTCGCC– 3'

Reverse: 5' –TCAATCATCGTAGTAATAATC– 3'

Primers used for amplification of C-terminal of Δ40 torsinA (C44S, C49S, C50S, C162S, C280S, C319S):

Forward: 5'–CTGGAATACAAACACCTAAAAATG– 3'

Reverse: 5' –TCAATCATCGTAGTAATAATC– 3'

Primers were synthesized at IDT Inc. (Coralvile, IA)

To express the target protein with the SUMO moiety as an N-terminal fusion partner, we utilized the same experimental scheme as C-terminal of Δ40TorsinA variants. To further purify the protein, the fractions containing SUMO-fusion torsinA variants were pooled and dialyzed against dialysis buffer (20 mM Tris-HCl, pH 8.0, 150 mM NaCl, 10% (v/v) glycerol. The SUMO fusion protein (~20 mL) was added to a Slide-A-Lyzer 3.5K Dialysis Cassette (Thermo Scientific) and dialyzed in 1liter dialysis buffer (twice, overnight) at 4°C. the dialyzed sample was centrifuged at 13,000 g for 5min, and the supernatant was collected. The concentration of purified torsinA proteins variants was calculated by measuring absorption at 280 nm and extinction coefficients for all constructs were calculated with the ExPASy ProtParam tool.

To investigate the optimal cleavage reaction conditions, a small-scale experiment was set in 200 µl volume: 20 µg of fusion protein was mixed with 20 µl 10x SUMO protease buffer without salt. Hydrolysis was initiated by addition of 10 unit or 20 unit of SUMO protease and was performed at 30 °C for a total of 3 h. During the 3 h 20 µl aliquots were withdrawn at different time points: 0 h, 1 h, 2 h, 3 h and added with 20 µl 2× SDS sample buffer. Samples were boiled for 5 min and analyzed by SDS-PAGE followed by Coomassie blue staining and silver staining.

SDS-PAGE and Western blot analysis

Proteins were treated with 2 × SDS loading buffer containing SDS and β-mercaptoethanol, and heated at 95 °C for 5min. samples were loaded on 4-15% TGX™ precast polyacrylamide gels (Bio-Rad) or 16% SDS-PAGE, and run on constant voltage of 200 V for approximately 30 min. Electrophoresis buffer was Tris-glycine-SDS buffer. Gels were stained for 20-60 min in 0.2% (w/v) Coomassie Blue R- 250 in 20% (v/v) methanol and 20% (v/v) acetic acid, followed by destaining with 40% (v/v) methanol and 10% (v/v) acetic acid. When the concentration of expected protein was low and could not be detected by Coomassie blue, gels were silver stained using the SilverXpress® Silver Staining Kit (Invitrogen).

For western blot analysis, proteins were transferred from SDS-PAGE gels to nitrocellulose membranes using western transfer buffer (0.25 M Tris, 1.92 M glycine, PH 8.3) and run on constant voltage of 100 V for 1 h. Membranes were blocked for overnight at 4°C with 10 ml Sea Block Blocking Buffer (Pierce) diluted with 40 ml PBS, in order to block non-specific protein binding site. The membrane was then incubated at RT for 1 h with an appropriate primary anti-torsinA antibody KSU 1451 (provided by Dr. Zhonghua Liu) diluted 1:2500 in 20% (v/v) Sea Block Blocking Buffer in PBS (12 µl anti-torsinA: 30 ml 20% (v/v) Sea Block Blocking Buffer in PBS). The membrane was washed with PBST 3 times, at 10 min per wash, followed by incubation with a secondary antibody (Goat Anti-Rabbit IgG –HRP; Southern Biotechnology Associates) diluted 1:4000 in Blocker Casein in PBS (Thermo Scientific). Membranes were again washed three times with PBST, and then developed with SuperSignal West Pico Chemiluminescent Substrate Kit (Pierce).

Chapter 3 - Results

Previously in our laboratory, Dr. Zhonghua Liu purified a variant of torsinA without the membrane-anchoring hydrophobic domain, torsinAΔ40. This mutant is used in this thesis (Figure 2).

Expression and purification of Cysteine-less torsinA in *E. coli*

TorsinA protein contains six conserved cysteine residues (Cys44, Cys49, Cys50, Cys162, Cys280, and Cys319) and could possess up to three disulfide bonds: one disulfide is located in the large N-terminal domain and the other in the C-terminal small domain (Konakova and Pulst, 2005b; Zhu et al., 2008). Upon overexpression in *E. coli*, wild type torsinA proteins were accumulated as insoluble inclusion bodies that required refolding to become soluble and functionally active (Konakova and Pulst, 2005b). We assume that wild type torsinA is not soluble after its expression in *E. coli*, possibly due to misfolding due to incorrect disulfide bond formation. To explore whether the cysteine residue may affect the solubility of protein, we changed all the cysteine to serine in torsinA sequence. The torsinA mutant C44S, C49S C50S, C162S generated by Hui-Chuan Wu in our laboratory was used as the template. The mutation C ➔S was introduced at position 280 and 319. Therefore two new torsinA variants 6×HisΔ40TorsinA (C44S, C49S, C50S, C162S, C280S) and 6×HisΔ40TorsinA (C44S, C49S, C50S, C162S, C280S, C319S) (Figure 3) were produced.

For these two torsinA variants, the solubility tests were then performed by measuring relative amount of protein in the cytoplasmatic fraction and in inclusion bodies by SDS–PAGE. The N-terminal His-tagged fusion proteins were expressed in *E. coli* BL21 (DE3). High expression levels of recombinant proteins were obtained after 1 mM IPTG induction at 37 °C (Figure 4,5). The strong Coomassie-stained bands corresponding to the predicted size of these two variants (~35kDa) were detected. After sonication, cell lysates from *E. coli* were subsequently separated by centrifugation into supernatant fractions, which contained soluble proteins, and pellet fractions, which contained transmembrane proteins. For both of Δ40TorsinA (C44S, C49S, C50S, C162S, C280S) and Δ40TorsinA (C44S, C49S, C50S, C162S, C280S, C319S), the bands (~35kDa) consistent with the results of induced *E. coli* cells were found

predominantly accumulated in the pellet fractions (Figure 4,5). To determine the efficiency of the solubility test, we monitored the AAA+ chaperone ClpB that was thought to be a soluble, cytoplasmically localized protein. The major band with molecular weight of 95 kDa was observed in the soluble fraction (data not shown). The results shown in Figure 4,5 indicate that Δ40TorsinA (C44S, C49S, C50S, C162S, C280S) and Δ40TorsinA (C44S, C49S, C50S, C162S, C280S, C319S) are not soluble after expression in *E. coli.*

Having identified these two torsinA variants as insoluble proteins, we attempted to purify 6×HisΔ40TorsinA (C44S, C49S, C50S, C162S, C280S, C319S) by nickel- affinity chromatography under denaturing conditions. Insoluble inclusion bodies containing 6×HisΔ40TorsinA (C44S, C49S, C50S, C162S, C280S, C319S) were isolated from the rest of the lysed host cell by centrifugation. The pellet containing the protein was subsequently solubilized and completely unfolded using 8 M urea. The mixtures were then applied directly to Ni^{2+}-affinity column. Proteins were eluted with an imidazole gradient containing 8 M urea. 4 mL elution buffer for each gradient, and fractions (1 mL each) were collected and subjected to SDS-PAGE. The purity was checked by SDS- PAGE (Figure 6). At the end of elution steps, the torsinA variant protein became increased upon the addition of 250 mM imidazole in the elution buffer. This suggests that 6×HisΔ40TorsinA (C44S, C49S, C50S, C162S, C280S, C319S) was probably isolated most effectively in the condition of elution buffer with 250 mM imidazole. These fractions were collected and subjected to refolding carried out by step-wise dilution thatwas described in Method. The sample was not stable over a long time and precipitation was observed during dialysis. The calculated final concentration of purified proteins was 0.021 mg/ml. Then purified protein was concentrated approximately 10 times to allow detection with SDS-PAGE. However, there is no visible band observed on SDS-PAGE (Figure 7), indicating that the amount of purified 6×HisΔ40TorsinA (C44S, C49S, C50S, C162S, C280S, C319S) was so low may due to precipitation during refolding by dialysis. In conclusion, the purification result was not successful, and we could not get sufficient amount of folded proteins.

The protein reactivation assay

Zolkiewski (1999) reported a multi-chaperone system from *E. coli*: ClpB together with DnaK, DnaJ, GrpE (KJE) can suppress and reverse protein aggregation. As we found that purified 6×HisΔ40TorsinA (C44S, C49S, C50S, C162S, C280S, C319S) precipitated during

refolding process by dialysis, we then asked whether ClpB with KJE might refold cysteine -less torsinA and improve recovery of recombinant proteins. However, the molecular function of torsinA is still largely unknown, and the ATPase activity of torsinA proteins was slightly lower (2.5-10-fold) than those displayed by ClpB and DnaK (Kustedjo et al., 2003). In this case, we could not detect the activity of torsinA. As described in Methods, we evaluated the relative amounts of soluble and insoluble Cys -less torsinA in the reactions with and without chaperone system.

It has been observed before that Cys -less torsinA eluted by 250 mM imidazole buffer containing 8 M urea precipitated into insoluble aggregates upon decreasing the urea concentration from 7 M to 0 M. We assumed that aggregates, which could be removed by centrifugation, were able to form when the elutant fraction diluted 10 folds into Tris-HCl based buffer. The supernatants were pooled and subjected to western blot analysis after incubating with and without chaperone system, as described in Methods. When ClpB, KJE and ATP were included in the refolding assay, one band with approximately 35 kDa were present, but not in the system without chaperone system (lane 1 compared to lane 4, Figure 8). A control experiment containing 8M urea was also performed (lane 5, Figure 8). The results above indicated that ClpB/DnaK/DnaJ/GrpE were able to refold Cys -less torsinA and improve recovery of recombinant proteins. Interestingly, the assay containing ClpB, KJE but in the absence of ATP revealed similar soluble amount of Cys-less torsinA compared with the ClpB, KJE and ATP system (lane 3, Figure 8). Existing data have showed that the coordinated function of ClpB and DnaK/DnaJ/GrpE required ATP hydrolysis (Zolkiewski, 1999). Next, we asked whether this refolding assay was chaperone-specific. We performed the assay with BSA, which had a strong ability to reversibly bind and transport an enormous variety of substances. As described in Methods, four different amounts of BSA (3.6 μM, 18 μM, 36 μM, 72 μM), which were enough to bind with Cys-less torsinA (1.53 μM), were used. However, no detectable bands could be observed in Figure 9. Thus, BSA had no effect on refolding Cys-less torsinA, and the refolding assay was chaperone-specific. The reason why the apparent activity of ClpB with KJE does not require ATP is currently unknown, and it will need to be further verified by more direct assays.

Expression and purification of the C-terminal fragment of wild type-torsinA and torsinA mutants in *E. coli*

The dystonia-linked mutation ΔE302/303 lies within the human torsinA C-terminal 62 residues region, which may play a critical role. To characterize this region and its potential function, a recombinant protein containing the C-terminal 62 residues of torsinA (Figure 2) named C-torsinA, and three mutants, C-torsinA (C280S), C-torsinA (C280S, C319S), C-torsinA ΔE were produced and expressed in *E. coli* BL21 DE3 cells.

In order to increase the solubility of the recombinant protein C-torsinA, the two different IPTG induction conditions: 37 °C 2 h and 20 °C 16 h were tested. As shown in Figure 10A, C-torsinA was predominantly present in the pellet upon 37 °C with 2 h of IPTG induction conditions, while the result obtained at 20 °C with 16 h of induction shows increased protein solubility (Figure 10B). This was consistent with the expectation that lowing the temperature increased protein solubility since proteins had more time to properly fold. Next, expressed proteins at 20 °C with 16 h induction were further purified using nickel- affinity chromatography using the 6×his rich region provided by the *pPROEX* vector. The purity was checked by SDS-PAGE (Figure 11). Unfortunately, it showed that proteins were bound poorly to the column. One possible explanation was that the protein could not fold correctly and his-tag might be buried within the protein structure. However, in the first two fractions eluted by 50 mM imidazole, a specific faint band of expected size for C-torsinA (~10 kDa) in each fraction was still detected, and verified by MS analysis (data not shown). With the partially purified C-torsinA sample, a gel filtration experiment was performed in order to further purified proteins. C-torsinA level in fractions was so low that it could not be detected by SDS PAGE. In this case, efforts to purify sufficient amount of C-torsinA in native folded form have failed.

The recombinant C-torsinA (C280S, C319S) was expressed in the Rosetta (DE3) E. coli strain, which has been reported an improved expression yield compared with *E. coli* BL21 (DE3) (Tegel et al., 2010). This protein was subsequently expressed at two temperatures 37 °C for 2 h and 20 °C for 16 h. The results showed the latter induction condition could further optimize C-torsinA (C280S, C319S) expression and enhance insolubility (Figure 12). To optimize the purity and the homogeneity of protein, metal-affinity chromatograph was performed with another metal ion, Cobalt (Co^{2+}-NTA), using the same conditions as for the Ni^{2+}-affinity purification. Similar to the Nickel results, proteins were bound poorly to the column, and mostly lost in the flow through

(data not shown). Moreover, purified C-torsinA (C280S, C319S) was unstable and precipitated over one day during storage at 4 °C. Still, purification of C-torsinA (C280S, C319S) by Talon cobalt resin, failed to give any good results.

Construct C-torsinA ΔE were also made and tested for solubility in 20 °C for 16 h induction conditions (Figure 13).

Expression and purification of torsinA variants by SUMO fusion in *E. coli*

The strategy for the construction of SUMO fusion Δ40torsinA and C-torsinA (C280S, C319S) (Figure15) was described in "methods" section. As shown in Figure 17, after induction with 1mM IPTG at 20 °C for 16 h or 37 °C for 2 h, SUMO- C-torsinA (C280S, C319S) was efficiently expressed in *E. coli*, and runs on an SDS-PAGE gel as around 23KD band. Most strikingly, the SUMO had been shown to strongly improve the C-torsinA (C280S, C319S) solubility. In addition, most of the SUMO fusion Δ40torsinA proteins were expressed in the pellet fraction (Figure 16), which indicates that SUMO fusion Δ40torsinA proteins expressed were insoluble and retained within inclusion bodies.

Soluble 6×His-SUMO-C-torsinA (C280S, C319S) protein was purified on Ni-NTA resin. Most of the proteins without 6×His tags were removed from Ni-column using washing buffer containing 10 mM imidazole. There was still some loss of target protein during washing (Figure 18 A, lane 3 and 4). The His-tagged fusion proteins were efficiently eluted with elution buffer containing gradient concentration of imidazole. Although there was still co-purification of less-abundant protein (impurities) appeared, the purified SUMO-C-torsinA (C280S, C319S) protein bands were most highly intense and reasonably purer as compared with the previous purification method (Figure 18). Moreover, the protein was stable at 4°C for one week. But during dialysis process, there was a tiny amount of protein precipitated, which may be due to the high protein concentration. Finally, we obtained ~15 ml of the purified SUMO--C-torsinA (C280S, C319S) protein sample in which the protein concentration was 4.76 mg/ml. It is apparent that SUMO fusion dramatically enhanced the yield of recombinant protein.

To separate the C-torsinA (C280S, C319S) from SUMO-C-torsinA (C280S, C319S) protein, the dialyzed fusion protein sample was added with the SUMO protease at a ratio of 1 or 2 unit of the enzyme to 2 μg of the substrate (One unit of SUMO Protease is defined as the amount of enzyme needed to cleave 85% of 2 μg of substrate protein at 30°C in 1 h). The time

course experiment was carried out. Coomassie blue stained SDS-PAGE gel could not detect the expected band, even fusion protein band (~23 kDa) could not be detected. A few highly intense bands were detected using the high sensitivity of silver staining, of which two bands were close to the expected size of hydrolysis products- the expected molecular weights of SUMO (~13 kDa) and the C-torsinA (C280S, C319S) protein (~10 kDa) (Figure 19). However, because of lack of anti- C-torsinA (C280S, C319S) antibody, we could not confirm that the band (~10 kDa) appeared close to the bottom of the gel in all the samples was our target protein. In sum, the result indicates that the SUMO fusion was not efficiently cleaved by SUMO protease, and modified reaction condition or site-specifically modified fusion proteins will be required for optimal cleavage.

Chapter 4 - Discussion

TorsinA is one of the few members of the AAA+ ATPases family involved in the human neurological disease and the only known AAA+ ATPase to target to the lumen of the secretory pathway. Moreover, torsins are found only in animal proteomes, which suggests that their function is essential for animal physiology. However, the role of torsinA is still unknown. Specific questions need to be addressed such as how to produce and study torsinA using biochemical and high resolution structural analysis, how the GAG deletion alters torsinA structure and function, substrate partner of torsinA.

A sufficient quantity of protein is the first step for initial functional and structural studies. To date, one of the major challenges in the study of torsinA has been the problem of obtaining a large amount of soluble and functional of protein. In the present study, efforts have been made to recombinantly express and purify torsinA and its mutated form from *E. coli*.

Expression and purification of torsinA in *E. coli*

We expressed torsinA in *E. coli* that was the system of first choice for expression of many heterologous proteins. Wild type torsinA was not soluble after its expression in *E. coli*. It was possibly due to misfolding supported by the six conserved cysteines, which were not found in other AAA+ ATPases. The six conserved cysteine indicated that the protein may possess up to three disulfide bonds and were subjected to posttranslational modifications (Kustedjo et al., 2003). So we tried to mutate the six cysteine to serine, and tested the solubility of torsinA. It showed that the majority of the torsinA variant ended up in the insoluble fraction. We attempted to purify His-6-tagged torsinA by using Ni^{2+}-affinity chromatography under denaturing conditions. SDS-PAGE analysis after Ni^{2+}-affinity chromatography confirmed the presence of torsinA variant in the eluted fractions. However, considering the initial expression level observed for torsinA, which was good, the limited yield (0.021 mg/ml) after refolding process indicated that a large amount of torsinA was lost, possibly due to protein precipitation upon step-wise dialysis.

Purification of C-terminal domain of torsinA was also unsuccessful. As predicted by a critical role likely to be played by the C-terminal domain of torsinA that contains the glutamate deletion, we produced fragments encoding the C-terminal of torsinA. A low-temperature, long-

incubation time induction strategy was also used to increase the yield of soluble, functional proteins. Then we tried to purify the protein by Ni^{2+}-affinity chromatography under native condition. The results showed that target protein did not efficiently bind to the column, and there were high non-specific of other proteins. Several possible reasons might contribute to this result: 1). The expression level of C-terminal region was not high. To overcome this problem, we used Rosetta (DE3) *E. coli* strain to optimize the expression level in the following experimental process. 2). The folding of the C-terminal region protein might result in partial hiding instead of exposure of His-tag, which makes torsinA less discriminable from other proteins. This aspect may be corrected by dual tag sequence or increasing the number of histidine residues. 3). The presence of inherent histidine-rich regions in host proteins may result in non-specific binding, leading to yield contamination or decreased purification efficiency. Because we did not find the similar problem in purifying Cys-less torsinA, it could not become a problem. The identity of torsinA in the band was unclear, but it could be confirmed by MS analysis. Then we used gel filtration for further purification. However, we still could not get significant amount of torsinA variant. Talon cobalt resin is an alternative immobilized metal affinity chromatography (IMAC) resin, which resulted in decreased levels of nonspecifically bound proteins and consequent yields consisting of increased purity. Similar to Ni column results, the C-terminal of Cys-less torsinA was bound poorly to the talon column and many contaminant proteins were bound. Another challenge encountered was that C-terminal of Cys-less torsinA tended to form insoluble aggregates during storage. Overall, the above- mentioned results of torsinA variants shows that are inherently hard to purify this protein in its native form.

Enhanced expression and purification of torsinA variants by SUMO fusion in *E. coli*

The small ubiquitin-related modifier (SUMO) proteins are highly conserved from yeast to human, but are absent in prokaryotes. There are three different human SUMO proteins-- SUMO-1, SUMO-2 and SUMO-3 with distinct functions. Yeast contains only one SUMO protein, called Smt3 (Butt et al., 2005). Even though SUMO proteins share only 17% sequence identity with ubiquitin, it resembles ubiquitin in its three-dimensional structure, as well as its ability to modify proteins by covalently attaching to lysine residues of target proteins. However, the ultimate fate for SUMOylated and ubiquitinated proteins is different. Ubiquitinated proteins are generally

targeted to degradation pathway, while sumoylation does not directly target proteins for degradation. Rather, SUMO modification participates in many important cellular processes such as transcriptional regulation, signal transduction pathways. (Butt et al., 2005; Melchior, 2000). Recent studies show that SUMO fusions can markedly enhance the level of expression and solubility of proteins in *E. coli*. Several proteins, such as hirudin variant-1 (rHV1), B lymphocyte stimulator (BAFF), have been successfully recombinant expressed and purified by SUMO fusion technique (Lu et al., 2012; Lu et al., 2009).

In our research, when the C-torsinA (C280S, C319S) was expressed as a yeast SUMO (Smt3) fusion, the solubility of the protein substantially increased in both 37°C for 2 h and 20°C for 16 h induction conditions (Figure 17). The exact mechanism by which SUMO exerts an effect on the protein solubility is not very clear now. It has been hypothesized that the solubility enhancer function is based on the special structure feature of SUMO protein. SUMO has a dense inner hydrophobic core. The surface of protein is hydrophilic and highly water-soluble. This structure is similar to amphipathic detergents and may exert a detergent-like effect on its fusion proteins, which may exhibit robust folding characteristics. Henceforward, folded SUMO act as a general molecular chaperone to help its less soluble fusion protein to quickly fold in its correct conformation (Lu et al., 2012).

The soluble protein lysate of 6xHis-SUMO- C-torsinA (C280S, C319S) was then purified on Ni-NTA resin. They were efficiently eluted with elution buffer containing gradient concentration of imidazole (Figure 18). Although some unwanted impurities were still present in the eluted samples, the most intense band (~23 kDa) detected on SDS-PAGE gel was the SUMO fusion protein. The purity and the amount of protein purified after Ni-NTA affinity chromatography significantly enhanced compared with previous purification methods. The fusion protein was stable when kept in 4°C for one week, and there was no aggregate occurred. In conclusion, the use of fusion partner SUMO was an effective system for heterologous protein expression. It helped to stabilize and increase the yield of recombinant protein.

Then SUMO tag was removed in order to generate free active target protein. In our study, only a minor proportion of the SUMO -C-torsinA (C280S, C319S) protein could be cleaved by SUMO protease. There may be several reasons for the poor SUMO cleavage of the fusion protein: 1). Folding of the fusion protein: unlike other proteases that only recognize a peptide sequence, SUMO protease 1 also recognizes the tertiary structure of the SUMO tag and cleaves

precisely after the final Gly residue of SUMO (Figure 14). Due to this unique feature, the folding of the protein might affect the cleavage process. So one possible reason is that the cleavage site became buried in the fusion protein and unavailable for cleavage by the SUMO protease. 2). N-terminal: The recent study showed that a constrictive hydrophobic tunnel within the active site of SUMO protease likely served to recognize the SUMO Gly-Gly motif, and the carboxyl terminus of bound SUMO must pass through this tunnel that lead towards 5-7Å of the active site (Elmore et al., 2011b). Butt et al. (2005) indicated that SUMO protease could not cleave target proteins which contain an N-terminal proline. It has been suggested that the tight tunnel could not accommodate the structural change induced by proline following the cleavage site, so it was unable to allow the substrate to come within the active site. Moreover, according to the Champion™ pET SUMO kit by Invitrogen, the efficiency of cleavage may be very low if the first amino acid in the target protein is leucine, valine, or lysine. However, the underling mechanism is still unknown. We hypothesize that SUMO protease does not prefer these substrates to come within 5-7Å of the active site, which may be not consistent with the length of leucine, valine large hydrophobic side chain. Following the suggestion by Invitrogen, additional serine spacer could be added between SUMO and the fusion partner. However, because of this additional serine, the cleaved target protein will be native-like instead of authentic native torsinA variant. 3). Cleavage reaction condition: Increase in incubation time and 1-fold excess of the protease did not improve the cleavage. However the longer incubation was not advantageous as it may result in protein aggregation (Johanssen et al., 2012). An active site cysteine residue that locates at the end of the tunnel is necessary for SUMO protease (Elmore et al., 2011a). So oxidization of DTT in SUMO protease buffer will cause the low efficiency of cleavage. 4). Different fusion proteins: Johanssen et al. (2012) suggested that a slightly increase in cleavage efficiency was most likely involved in larger protein. So it is a good hint for us to try the whole length torsinA. 5). Optimization of fusion protein expression and purification: trying to purify the fusion protein in denaturing conditions may help us get more fusion protein and do the following experiment. Unfortunately, there was no time to optimize the cleavage reaction, but addition of serine residue would be the first measure to take.

The relationship between the primary sequence of torsinA and its propensity to be insoluble on overexpression in *E. coli*

Obtaining sufficient amount of fully active protein is the biggest challenge in our study. The lack of solubility of proteins is a major obstacle to obtaining purified protein. Strategies we have tried to improve protein solubility during heterologous expression include fusion with solubility enhancing SUMO tag, introduction of mutations, truncations at the N and C termini as well as different induction conditions. These methods have manifested in improved solubility. However, we still could not purify sufficient amounts of torsinA in native folded form.

Under a given set of experimental conditions, protein solubility is ultimately determined by its primary strcture (Smialowski et al., 2007). To investigate the relationship between the protein sequence features of torsinA and its propensity to be insoluble on overexpression in *E. coli*, we compared the amino acid composition of torsinA and its closest homologue of C-terminal AAA+ domain of *T. thermophilus* ClpB which was soluble, cytoplasmically localized proteins. We calculated the parameters of sequence features for these two given proteins according to Idicula-Thomas (Idicula-Thomas and Balaji, 2005) and used the ProtParam web server (http://web.expasy.org/protparam/). Five protein characteristics exhibited significant correlations with protein solubility: average net charge, aliphatic index, instability index, hydropathicity and residue composition (Table 3).

Average net charge

Davis et al. (1999) suggested that a critical parameter for solubility was the approximate-charge average, determined by differences in the numbers of Asp plus Glu vs. Lys plus Arg residues, and there was a inverse correlation between absolute net charge and inclusion formation. Lack of a net charge fosters interactions between protein molecules rather than between protein and water molecules, making protein aggregation or protein precipitation more likely. The average charge of C-terminal AAA+ domain of thermophilus ClpB is higher compared to that of torsinA, which implies that torsinA may be more prone to form inclusion bodies.

Aliphatic index

The aliphatic index is directly related to the relative volume occupied by aliphatic residues and has usually been associated with thermostability (Espargaro et al., 2008). Idicula-Thomas and Balaji (2005) indicated that aliphatic index was another crucial determinant of solubility, and thermostability and solubility on overexpression had a positive correlation. The aliphatic index of C-terminal AAA+ domain of thermophilus ClpB was found to be significantly higher than that of torsinA and hence, an increase in the thermostability might favor a higher solubility.

Instability index

Instability index (II_P) provides an estimate of the stability of proteins by considering all the residues. Thus it can be viewed as a measure of in vivo half-life of a protein. We observed that II_P of torsinA was lower than C-terminal AAA+ domain of thermophilus ClpB. It was consistent with the result from Idicula-Thomas and Balaji (2005) that II_P of soluble proteins were higher compared with insoluble ones. It was reported that II_P might influence the solubility of proteins for two reasons: 1). The long-lived partially folded intermediates had more chance to interact with other intermediates. 2). The longer-lived partially folded intermediates would exhaust the available molecular chaperones that prevented them from aggregating (Idicula-Thomas and Balaji, 2005).

Hydropathicity and residue composition

A negative correlation between protein hydrophobicity and solubility has previously been noted for proteins in general. The grand average of hydropathicity index (GRAVY) indicates the solubility of the proteins: with a positive GRAVY a protein can be characterized as hydrophobic and with a negative GRAVY a protein is hydrophilic (Kyte and Doolittle, 1982). TorsinA and C-terminal AAA+ domain of thermophilus ClpB have respectively a GRAVY -0.025 and -0.214. From these results we can conclude that torsinA has a higher hydrophobicity than C-terminal AAA+ domain of thermophilus ClpB, and therefore lower solubility.

A number of previous studies addressed the relationship between protein solubility and residue compositions. Goh et al. (2004) found that serine composition was the highest-ranked determinant of solubility based on the study of 27,000 proteins from 120 organisms. However, it

was not clearly understood how the decrease in Ser composition affects a protein's solubility. TorsinA and C-terminal AAA+ domain of thermophilus ClpB have respectively a percentage of Ser 6.3% and 2.6%, which was consistent with the fact that torsinA had lower solubility. Furthermore, Relative content of negatively charged residues (DE) was selected as an important attribute by all researchers so far and high content of these residues (DE >18%) were associated with improved solubility (Goh et al., 2004; Smialowski et al., 2007). However, no statistically significant correlation between DE content and protein solubility was observed in torsinA and C-terminal AAA+ domain of thermophilus ClpB. Frequencies of Cys and Met were also found to be important by Goh et al. (2004). The higher amount of two residues in torsinA compared to C-terminal AAA+ domain of thermophilus ClpB increases the probability of the protein to be insoluble. Additionally, Idicula-Thomas and Balaji (2005) suggested that certain amino acids such as Asn, Thr, and Tyr were another critical determinants of solubility of specific. In torsinA sequence, Asn and Tyr content are higher compared with C-terminal AAA+ domain of thermophilus ClpB, which is consistent with results from Idicula-Thomas and Balaji that these three amino acids occurred in higher frequency in insoluble proteins.

Identification of rare codon along torsinA mRNA and their relation to protein folding

A rare codon is usually defined by a low usage frequency. A correlation between the frequency of codon usage, the tRNAs contents, and gene expression levels has been established. Codons that correspond to low-abundance tRNA can form slow-translating regions in the mRNA and cause ribosomal pausing (Zhang et al., 2009). Different factors have been proposed to contribute to codon usage preference, such as selection for higher speed and more accurate translation; gene function type and gene size; GC composition bias; protein hydropathy; even RNA structure and stability (Ciencias, 2008).

In order to increase the expression level of recombinant proteins, rare codons are replaced by more frequently used codons that lead to increased yield of active proteins (Widmann et al., 2008). However, in some cases the substitution of rare codons by frequent ones can result in protein misfolding and inactive biological functions. Slow-translating codons along mRNA can potentially control the efficiency and accuracy of translation process by causing the ribosome

stalling, which allows the newly synthesized peptide chains to adopt well-folded intermediate conformations (Widmann et al., 2008; Zhang et al., 2009).

To determine the functional relevance of the slow-translating stretches for human torsinA folding, we identified the rare codons in the coding region of torsinA by using Graphical Codon Usage Analyzer (GCUA), which is a free online tool (http://gcua.schoedl.de/) to calculate a number of codon usage indices. The relative codon frequency analysis for the torsinA gene was done using the option "each triplet position vs. usage table" of the program. The relative codon frequency analysis had revealed that human torsinA was harboring 17.7% of rare codons for human. Among them, some codons were used less than 20% and 10% that were shown in blue and red bars respectively in Figure 20.

Due to the degeneracy of the genetic code, the number of rare codons varies among different synonymous codon families. Amino acids ranked in descending potential for rare codon use are as follows: (Leu) 5/6 rare; (Arg, Ser) 4/6 rare; (Val) 2/4 rare; (IIe) 1/3 rare; (Ala, Pro, Gly, Thr) 1/ 4 rare; (Phe, Met, Tyr, Asn, Lys, Asp, Glu, Cys, His, Gln, Trp) 0 rare. Based on the fact that slow-translating stretches are essential for folding, it might suggest that folding requires slowing down translation at these amino acids that encoded by rare codons, such as Leu, Arg.

Codon-usage bias in torsinA protein expression in *E. coli*

Codon usage varies among different organisms. The tRNA population in each organism closely reflects the codon bias of the mRNA population (Lee et al., 2009). For example, in *E. coli*, $tRNA^{Arg}$ that corresponding to the infrequently used AGG and AGA codon is at low level. A subset of the rarest codons in E. coli is shown in Table 1 (Lee et al., 2009). High-level expression of heterologous genes with codons that rarely used by *E. coli* can impede translation due to lack of charged tRNAs in the expression host (Terpe, 2006). In many cases, heterologous proteins are expressed in a low level or as insoluble aggregates, which may be attributable to such differences in codon usage between expression and natural hosts (Angov et al., 2008).

To determine how the translation rate changes if translating the human torsinA mRNA in E. coli, the fraction of usage of each codon in the torsinA sequence will be computed and plotted against the fraction of usage of the codon in *E. coli* by using Graphical Codon Usage Analyzer. TorsinA was harboring 22.5% of rare codons for *E. coli*. Similarly, among them, some codons

were used less than 20% and 10% that were shown in blue and red bars respectively in Figure 21.

Codon Adaptation Index (CAI) is a simple, effective measure of synonymous codon usage bias. The index may also give an approximate indication of the likely success of heterologous gene expression (Sharp and Li, 1987). We calculated CAI value for human torsinA expressed in *E. coli* by using Genscript which is a free online website for rare codon analysis (Figure 22). The CAI for our genes is 0.66. A CAI of 1.0 is rated as ideal while a CAI of >0.8 is considered as good for expression in the desired expression organism. The number of occurrences of the six rarest codons of E. coli (AGG, AGA, CGA, AUA, CUA, CCC) (Kane, 1995) in torsinA sequence is listed in Table 2.

The way to circumvent codon-usage bias is to optimize the coding sequence or use tRNA-enhanced host strains. The latter choice is a much more economic and suitable way. Many E. coli strains were engineered to solve this problem. For example, Rosetta (DE3) E. coli strain are BL21 derivatives designed to enhance the expression of eukaryotic proteins that contain codons rarely used in E. coli, such as AUA, AGG, AGA, CGG, CUA, CCC, and GGA. We tried to express C-torsinA (C280S, C319S) in the Rosetta host strain, and expression yield increased by changing strain from E. coli BL21 (DE3) to E. coli Rosetta(DE3). However, efficient translation is necessary but not only parameter to produce a functional protein. The absence of chaperones, normal partners, or correct redox environment may provide additional challenges (Gustafsson et al., 2004).

Future work

In this study, the data provided demonstrates that the 6×His SUMO fusion construct is able to enhance expression and facilitate purification with Ni-NTA chromatography. The lack of an endogenous SUMO protease in *E. coli* facilitates the use of SUMO as a fusion tag to efficiently express the recombinant proteins in host organism (Butt et al., 2005). Although a satisfactory yield of the fusion protein was successfully purified, the fusion protein was found difficult to be cleaved by protease. In this case, (1) addition of serine spacer between SUMO and C-terminal of Δ40TorsinA (C44S, C49S, C50S, C162S, C280S, C319S) would be the first measure to take; (2) go on with searching for an optimal cleavage conditions; (3) try to purify SUMO-Δ40TorsinA protein under denaturing conditions, and perform the cleavage reaction to

see whether the first amino acid of target protein influence the cleavage efficiency. However, it is possible that our target proteins become insoluble after cleavage from the SUMO tag. Therefore, This suggests that alternative strategies may be considered for the future expression and purification, for example, the search for a new expression system, eg. *Saccharomyces cerevisiae* with which to generate active proteins will be an option.

Once we get appreciable quantities of these folded proteins in the future, a series of biochemical and high-resolution structural approaches will be performed. Moreover, it remains important to further explore the torsinA activity in the context of its identification and functional analysis, such as cellular co-localization and possible interaction and/or functional cooperation between different torsins.

Figures

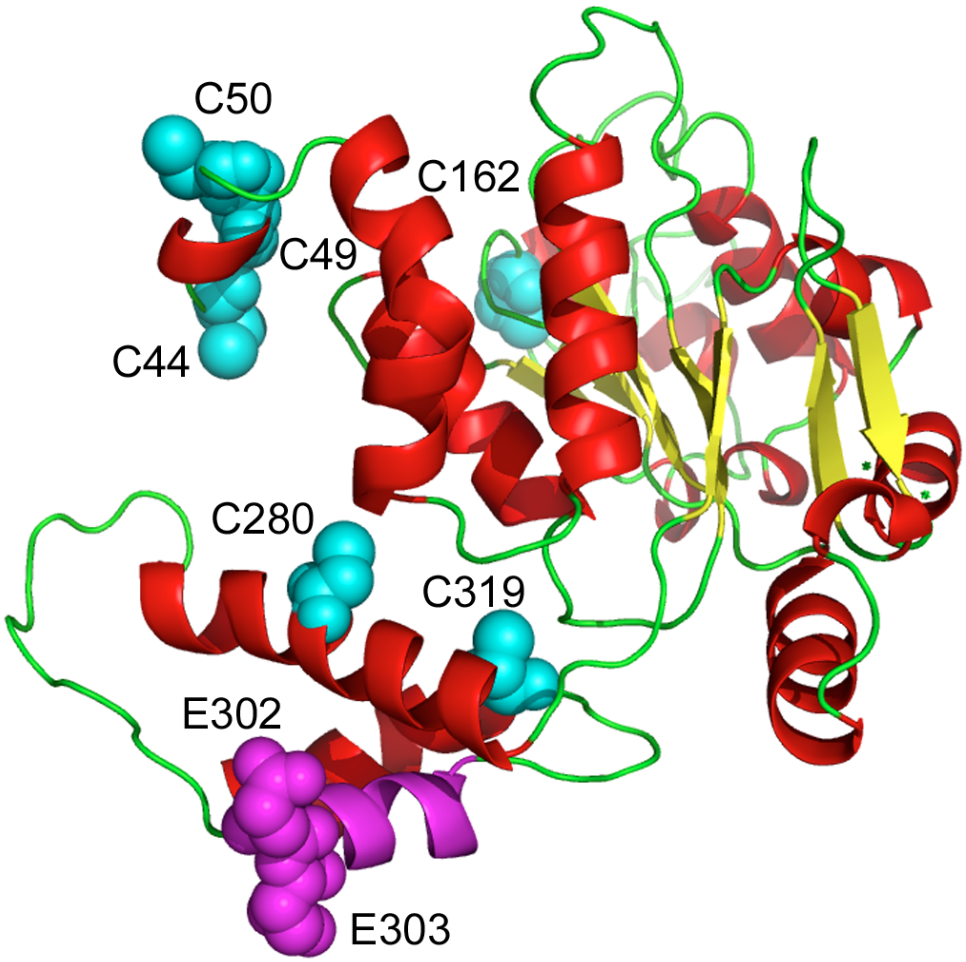


Figure 1: Structural models of human TorsinA, based on an alignment with the closest homologues of known structure, ClpA and Clp B from *E. coli* (Zolkiewski, 2011).
The N-terminal α/β subdomain (on the top) and the C-terminal α-helical subdomain (at the bottom) are shown in the figure respectively. The six conserved cysteines (C44S, C49S, C50S, C162S, C280S, C319S) are colored in blue. The pair of glutamates at position 302/303, one of which is deleted in the dystonia-linked mutant torsinA is marked in purple.

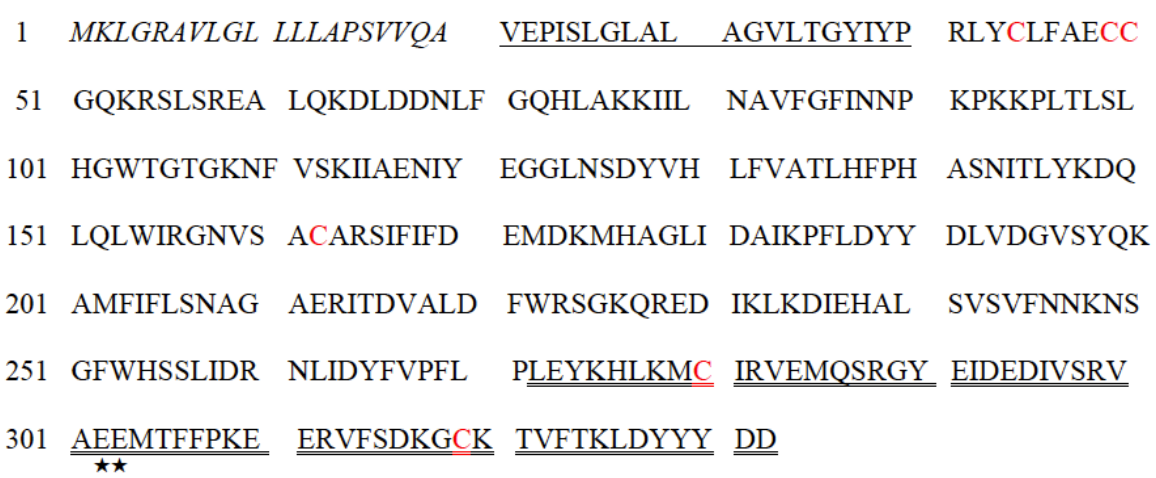


Figure 2: Amino acid sequence of human torsinA

The cleaved signal sequence and hydrophobic segment are shown in italics and underlined, respectively. Six conserved cysteines are colored in red. The dystonia-associated mutation, the deletion of one of glutamate residues at position of 302/303 are marked with ⋆⋆. The C-terminal sequence is double underlined.

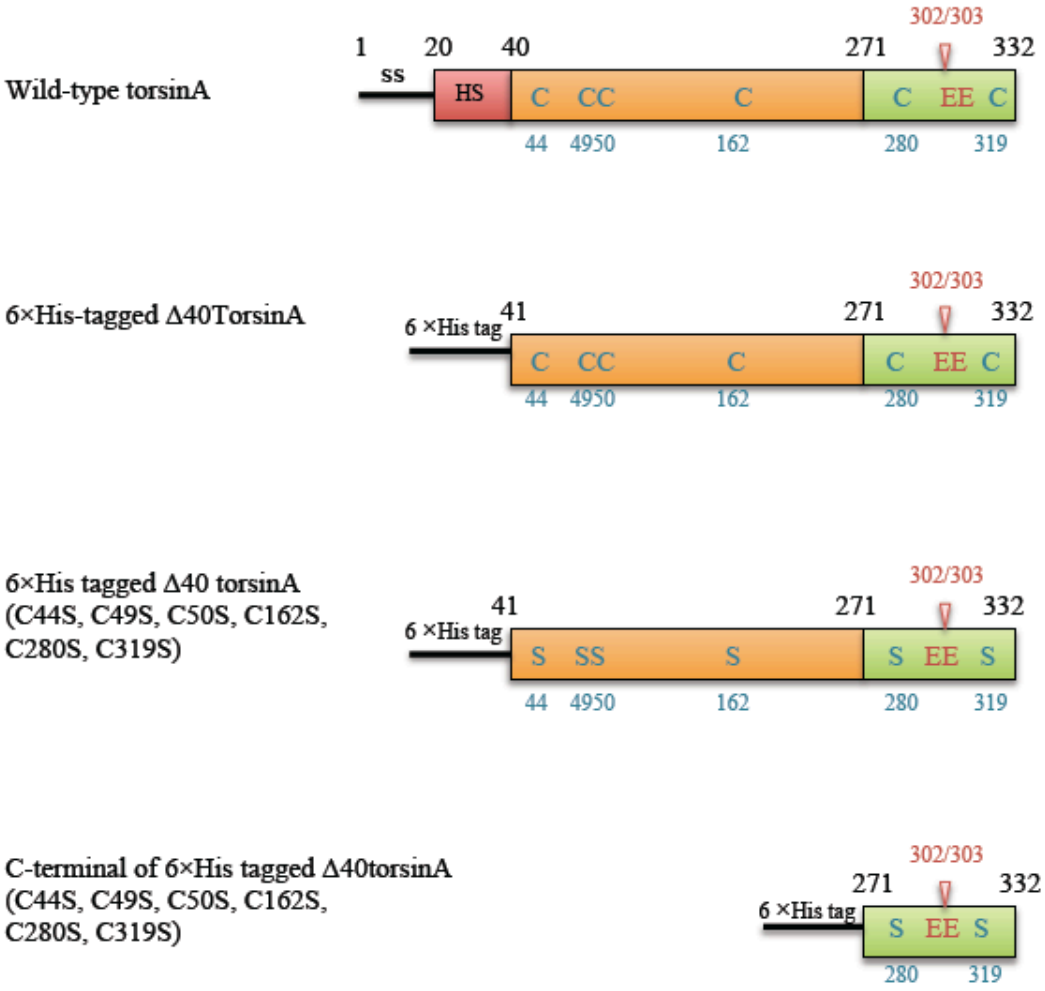


Figure 3: Schematic representation of 6 ×His Tag wild-type torsinA and torsinA mutants. Domains are indicated by abbreviations as follows: SS (signal sequence: residue 1-20), HS (hydrophobic segment: residue 21-40), N-terminal subdomain from residue 41-271, and a C-terminal subdomain from residue 272-332. Asterisk indicates location of ΔE 302/303 mutation. The SS and HD domains were removed from torsinA and replaced with 6*His tag the for purification purposes. Δ40 torsinA (C44S, C49S, C50S, C162S, C280S, C319S) were produced by mutating Ser to Cys.

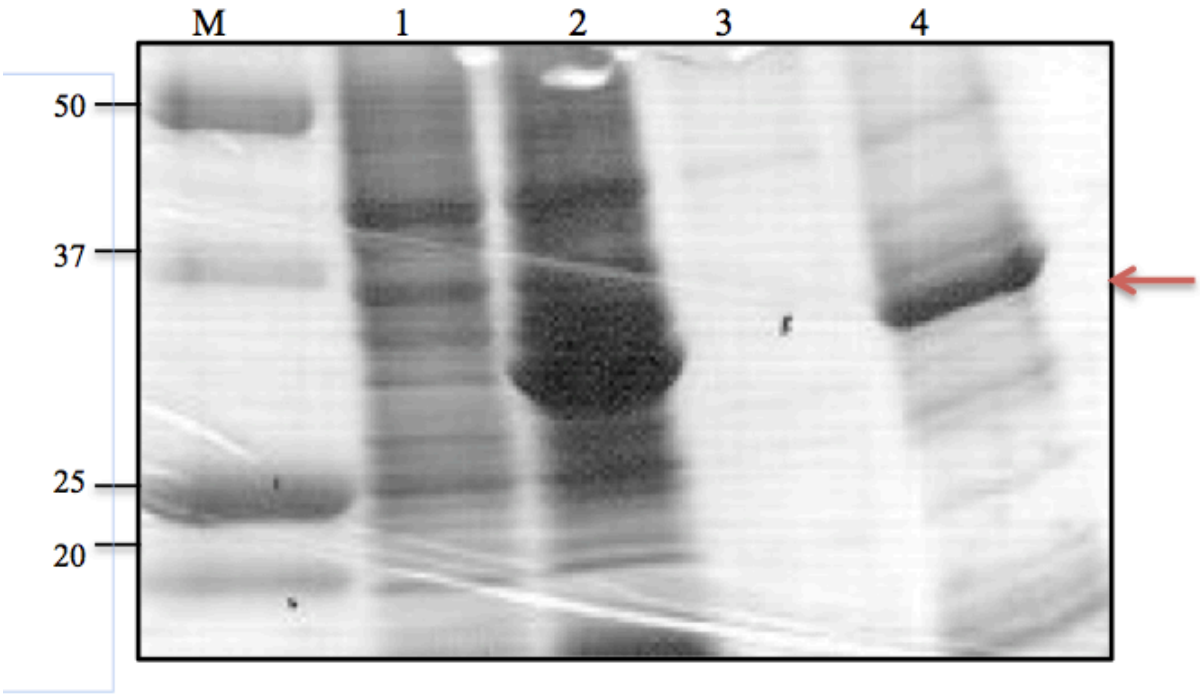


Figure 4: SDS gel electrophoresis analysis of Δ40TorsinA (C44S, C49S, C50S, C162S, C280S) solubility

M: molecular mass ladders.

Lane 1: *E. coli* cells without IPTG induction.

Lane 2: *E. coli* cells with 1mM IPTG induction at 37°C for 2 hs.

Lane 3: Induced *E. coli* supernatant after sonication.

Lane 4: Induced *E. coli* pellet after sonication.

The arrow indicates the band of Δ40TorsinA (C44S, C49S, C50S, C162S, C280S) protein.

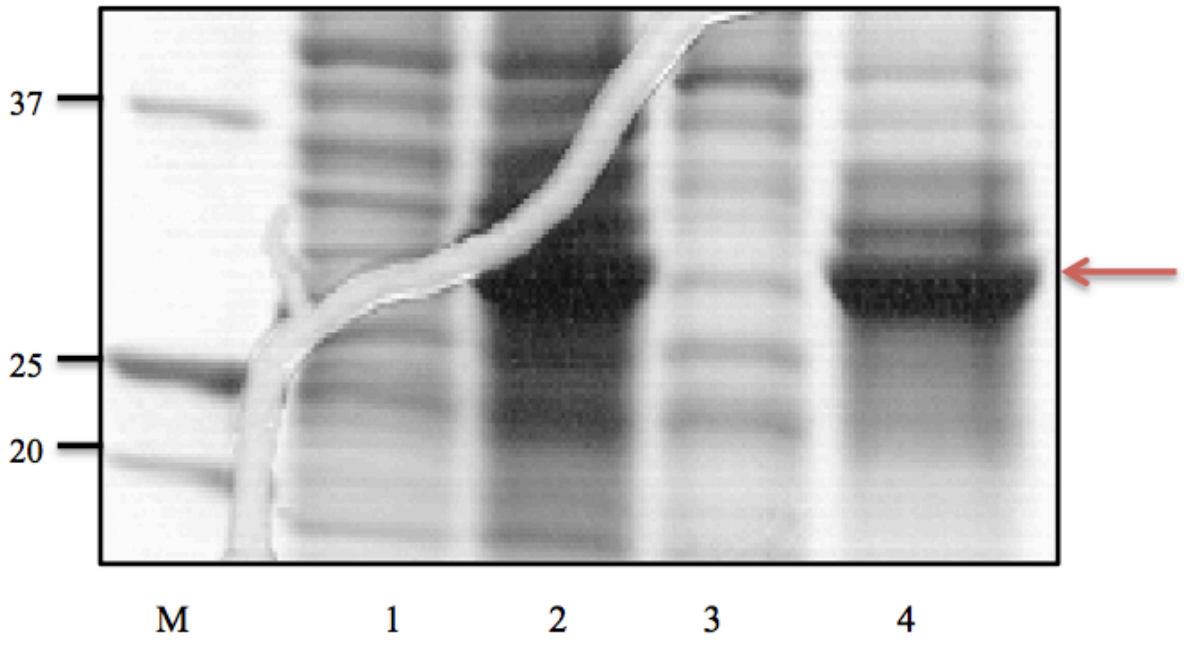


Figure 5: SDS gel electrophoresis analysis of Δ40TorsinA (C44S, C49S, C50S, C162S, C280S, C319S) solubility

M: molecular mass ladders.

Lane 1: *E. coli* cells without IPTG induction.

Lane 2: *E. coli* cells with 1mM IPTG induction at 37°C for 2 h.

Lane 3: Induced *E. coli* supernatant after sonication.

Lane 4: Induced *E. coli* pellet after sonication.

The arrow indicates the band of Δ40TorsinA (C44S, C49S, C50S, C162S, C280S, C319S) protein.

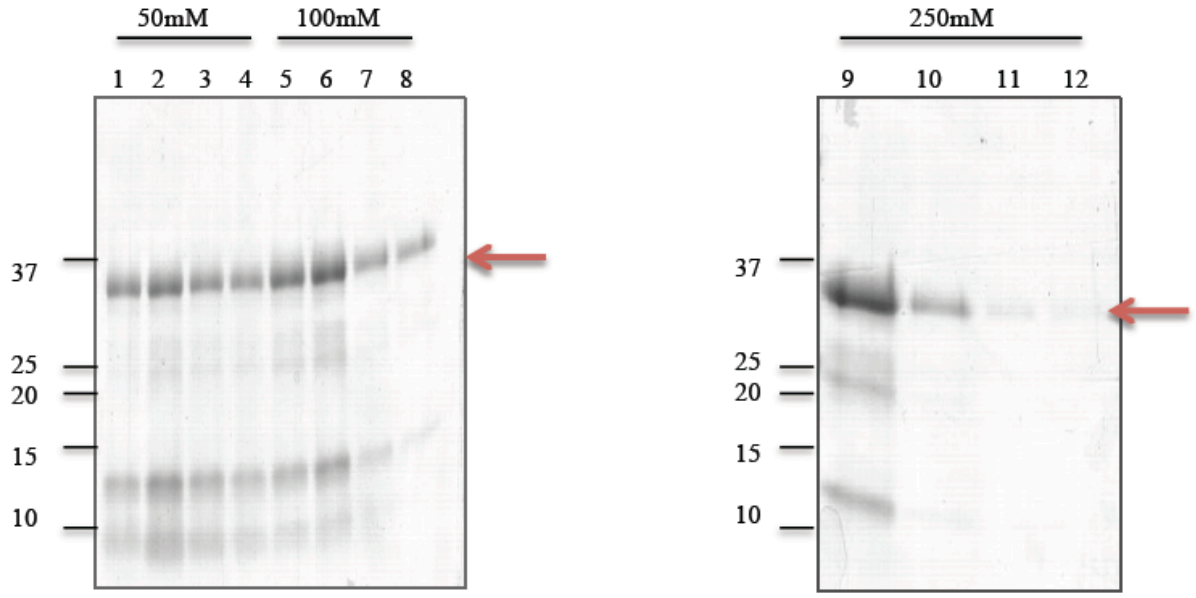


Figure 6: SDS gel electrophoresis analysis of purification of 6×His-tagged Δ40TorsinA (C44S, C49S, C50S, C162S, C280S, C319S) by using Ni-NTA affinity chromatography under denaturing conditions.

The purification was performed in 8 M urea. The 6×His-tagged Δ40TorsinA (C44S, C49S, C50S, C162S, C280S, C319S) proteins were eluted from Ni-NTA column with increasing concentrations of imidazole. 4 mL was for each gradient and fractions (1mL each) were collected and subjected to SDS-PAGE and stained with Coomassie blue stain. The arrow indicates the bands for the target protein.

Lane 1, 2, 3, 4: elution with 50 mM imidazole.

Lane 5, 6, 7, 8: elution with 100 mM imidazole.

Lane 9, 10, 11, 12: elution with 250 mM imidazole.

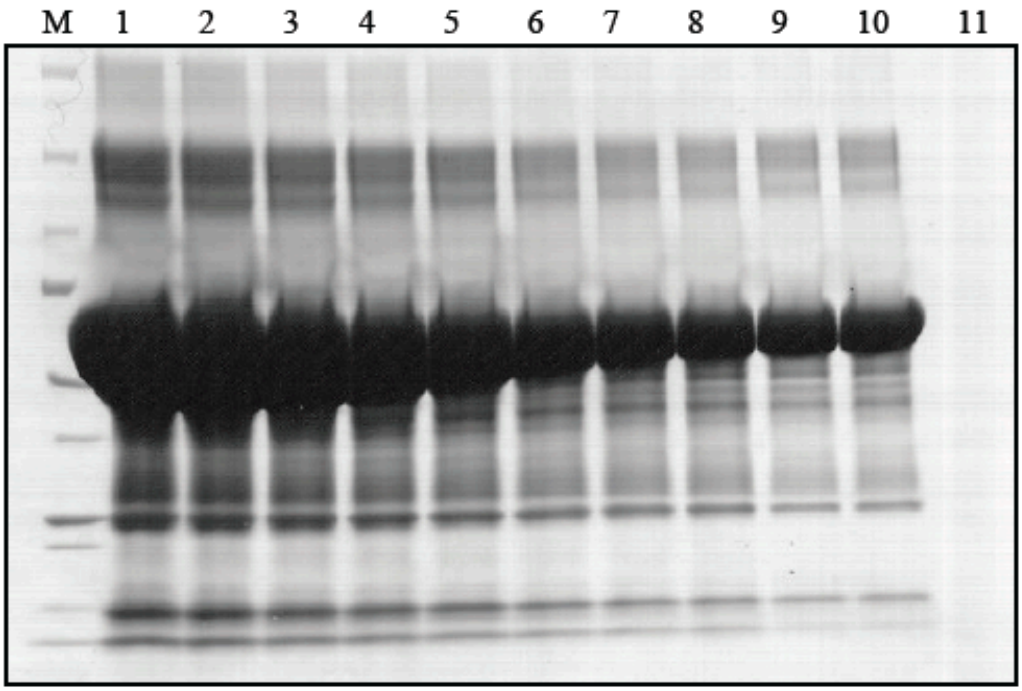


Figure 7: SDS gel electrophoresis analysis of measurement of purified of 6×His-tagged Δ40TorsinA (C44S, C49S, C50S, C162S, C280S, C319S) protein concentration by using a serial dilution of BSA as a standard.

M: molecular mass ladders.

The concentration of BSA (μg/μl) from lane 1 to lane 10 is 30, 20, 10, 8, 5, 4, 3, 2.5, 2, 1.5.

Lane 11: the purified target protein.

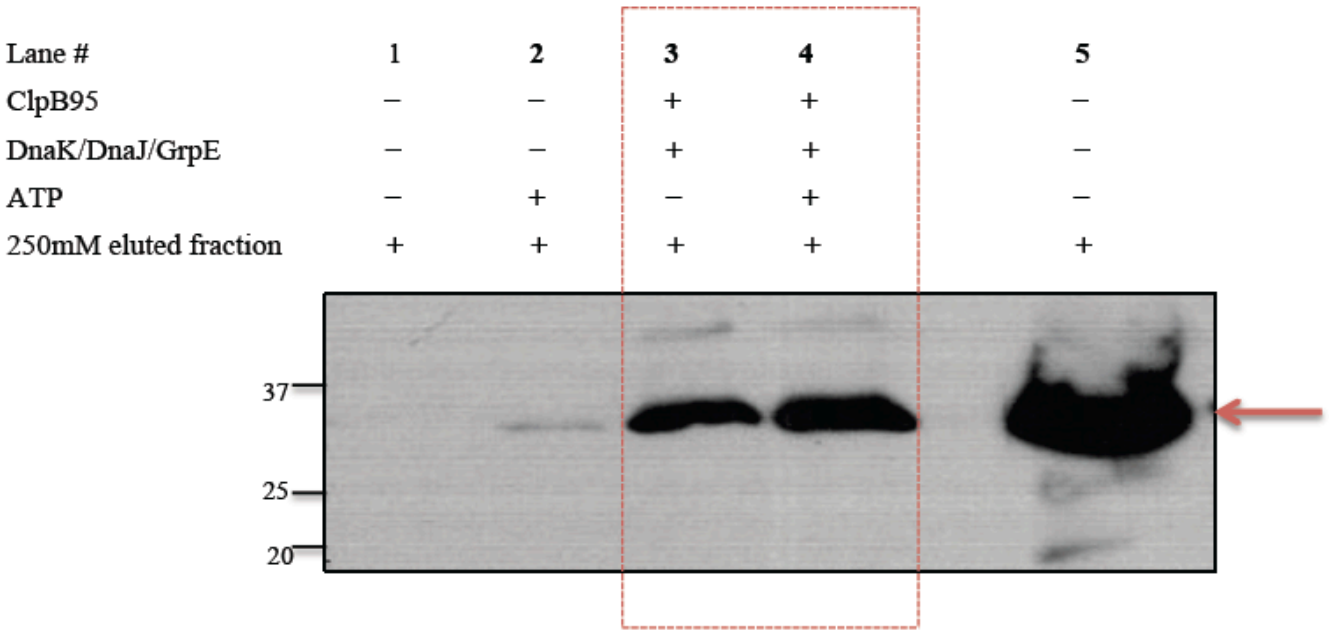


Figure 8: Reactivation of aggregated Δ40TorsinA (C44S, C49S, C50S, C162S, C280S, C319S) in the presence of ClpB 95 and the DnaK chaperone system.

The purified Δ40TorsinA (C44S, C49S, C50S, C162S, C280S, C319S) eluted by 250mM imidazole was diluted 10-fold in buffer A in presence of ATP (lane 2), or in presence of mixtures of DnaK/DnaJ/GrpE (KJE) + ClpB95 without ATP (lane 3) or DnaK/DnaJ/GrpE (KJE) + ClpB95 with ATP (lane 4) for 1 h at 37°C. In case of the control samples the same purified Δ40TorsinA (C44S, C49S, C50S, C162S, C280S, C319S) fraction was diluted 10-fold in 8 M urea buffer under the same condition. The concentrations used were: 1.53 µM of the purified protein sample, 1.5 µM ClpB95, 1 µM DnaK, 0.5 µM DnaJ and 0.5 µM GrpE. After incubation, the sample was centrifuged and supernatant was subjected to Western blot analysis. The results of the Western blot are shown. The arrow indicates the bands for the target protein.

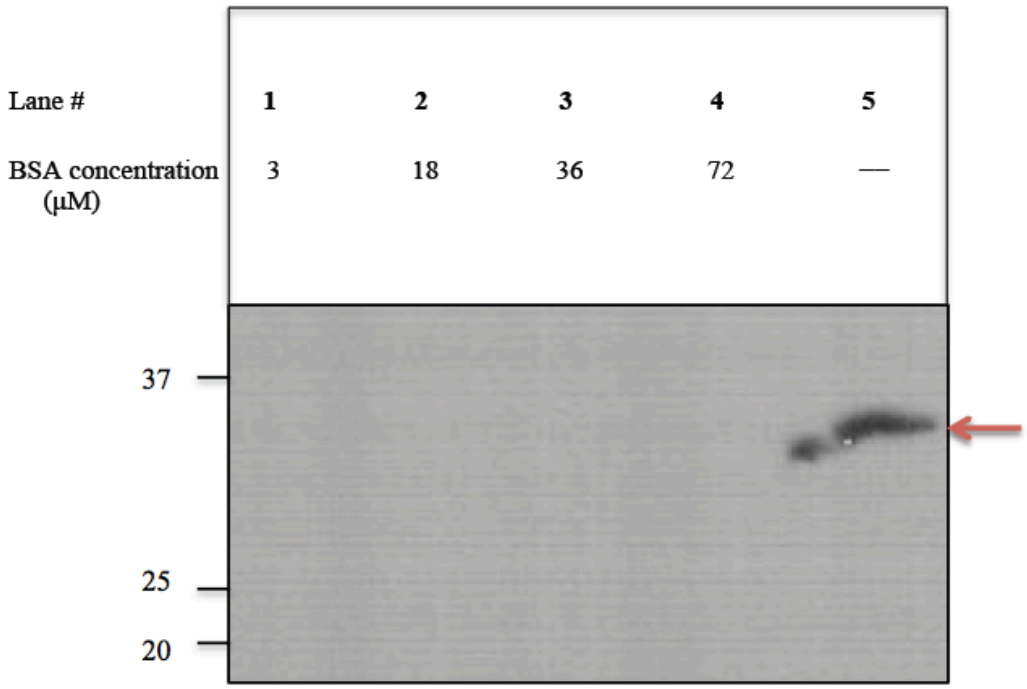
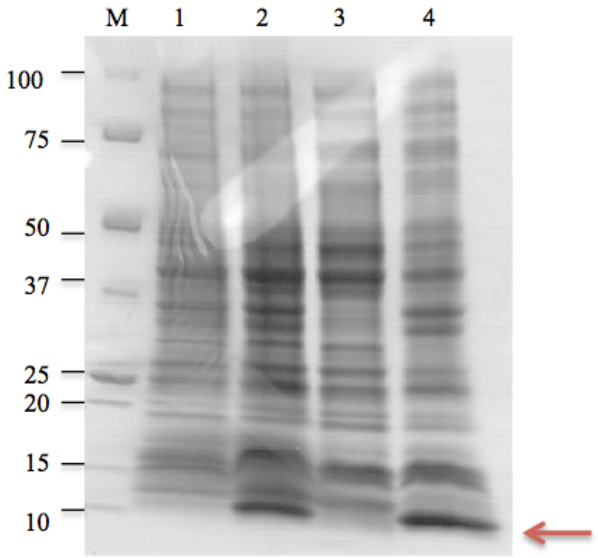


Figure 9: Reactivation of aggregated Δ40TorsinA (C44S, C49S, C50S, C162S, C280S, C319S) in the presence of BSA.

The purified Δ40TorsinA (C44S, C49S, C50S, C162S, C280S, C319S) eluted by 250mM imidazole (1.53 μM) was diluted 10-fold in buffer A containing a series of different concentrations of BSA (3.6 μM, 18 μM, 36 μM, 72 μM) (lane 1-4) for 1 h at 37°C. Control sample was diluted 10-fold in 8M urea buffer under same condition (lane 5). After incubation, the sample was centrifuged and supernatant was subjected to Western blot analysis. The results of the Western blot are shown. The arrow indicates the bands for the target protein.

A



B

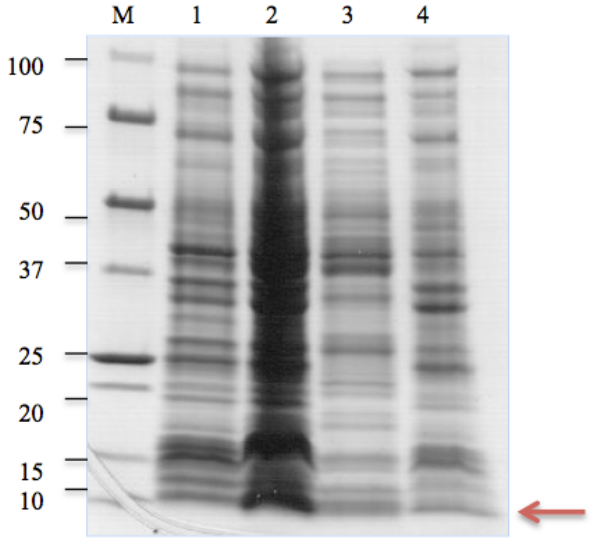


Figure 10: SDS gel electrophoresis analysis of C-terminal of Δ40TorsinA solubility. Protein expression was induced by 1mM IPTG at 37°C for 2 h (A) or 20°C for 16 h (B).

M: molecular mass ladders.

Lane 1: *E. coli* cells without IPTG induction.

Lane 2: *E. coli* cells with 1mM IPTG induction.

Lane 3: Induced *E. coli* supernatant after sonication.

Lane 4: Induced *E. coli* pellet after sonication.

The arrow indicates the band of C- terminal of Δ40TorsinA

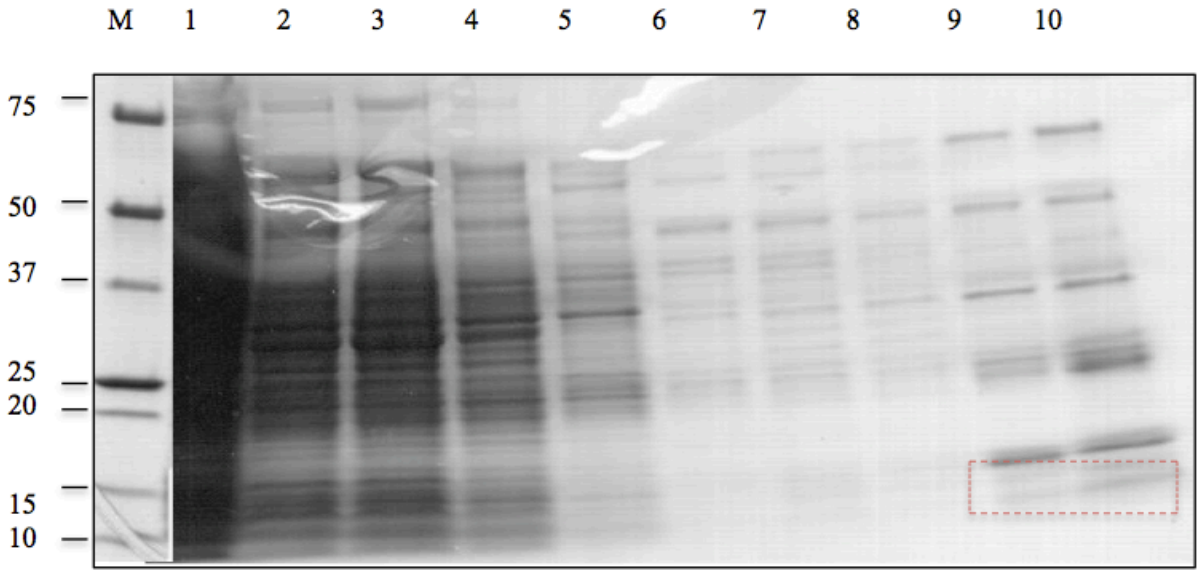


Figure 11: SDS gel electrophoresis analysis of purification of C-terminal of 6×His-tagged Δ40TorsinA by using Ni-NTA affinity chromatography under native conditions.

M: molecular mass ladders.

Lane 1: unbound.

Lane 2, 3: wash.

Lanes 4, 5, 6 and 7: elution with the buffer containing 10mM imidazole.

Lanes 8, 9 and 10: elution with the buffer containing 50mM imidazole.

The red box indicates the band of C- terminal of Δ40TorsinA.

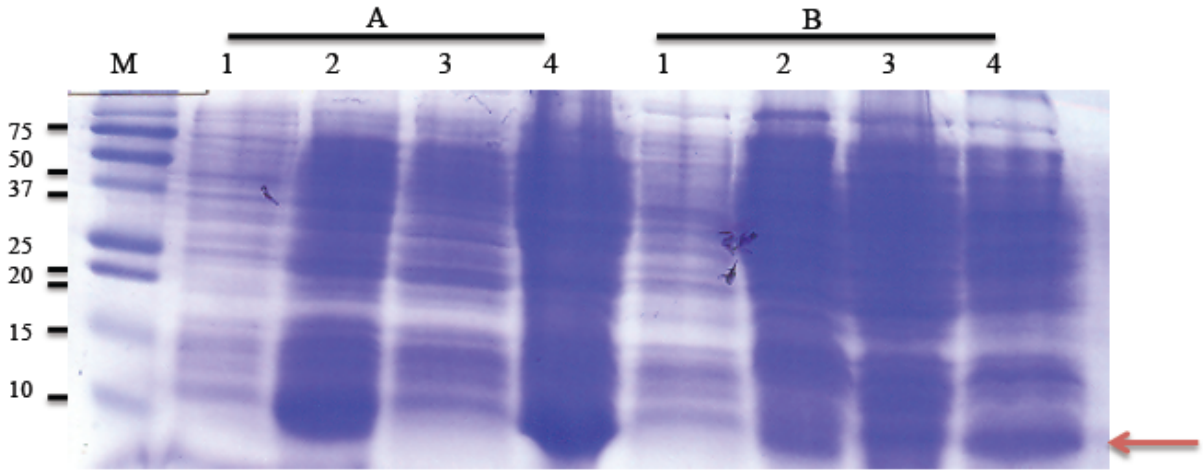


Figure 12: SDS gel electrophoresis analysis of C-terminal of Δ40TorsinA (C44S, C49S, C50S, C162S, C280S, C319S) solubility.

Protein expression was induced by 1mM IPTG at 37°C for 2 h (A) or 20°C for 16 h (B).

M: molecular mass ladders.

Lane 1: *E. coli* cells without IPTG induction.

Lane 2: *E. coli* cells with 1mM IPTG induction.

Lane 3: Induced *E. coli* supernatant after sonication.

Lane 4: Induced *E. coli* pellet after sonication.

The arrow indicates the band of C- terminal of Δ40TorsinA (C44S, C49S, C50S, C162S, C280S, C319S)

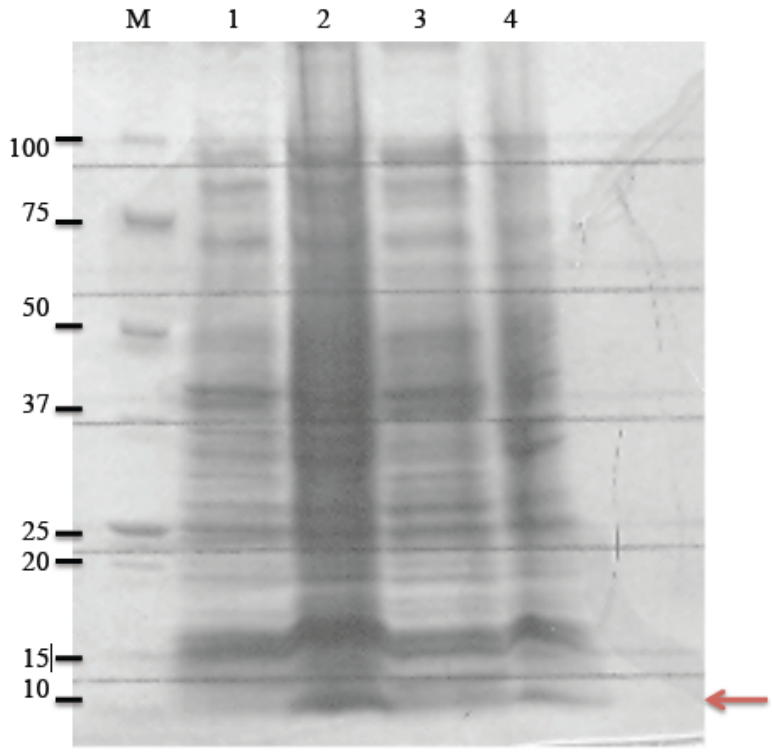


Figure 13: SDS gel electrophoresis analysis of C-terminal of Δ40TorsinA (ΔE) solubility. Protein expression was induced by 1mM IPTG at 20°C for 16 h.

M: molecular mass ladders.

Lane 1: *E. coli* cells without IPTG induction.

Lane 2: *E. coli* cells with 1mM IPTG induction.

Lane 3: Induced *E. coli* supernatant after sonication.

Lane 4: Induced *E. coli* pellet after sonication.

The arrow indicates the band of C- terminal of Δ40TorsinA (ΔE).

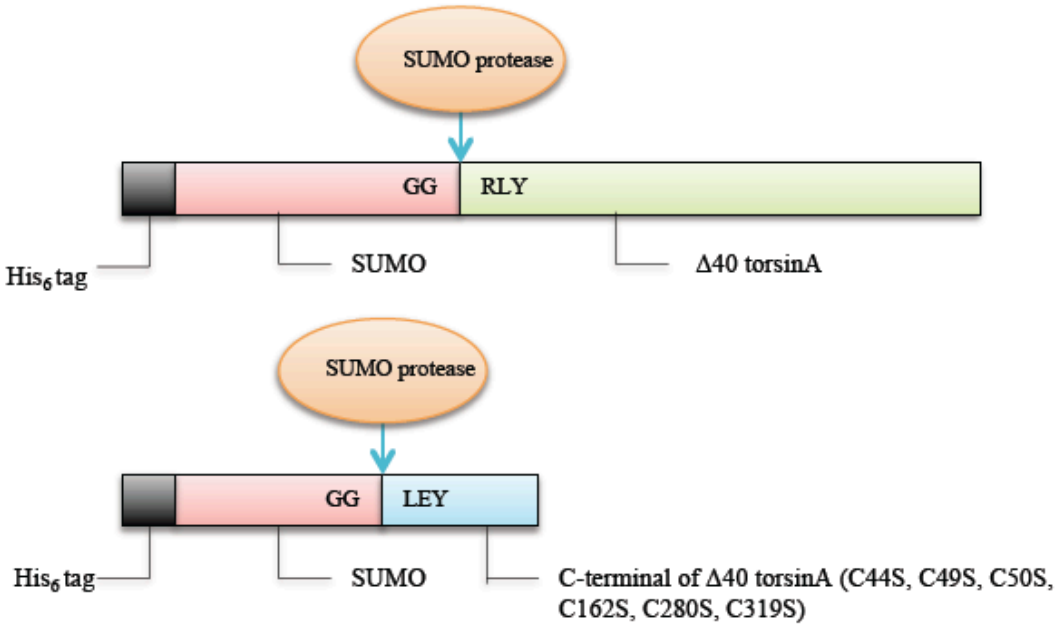


Figure 14: Schematic representation of 6×His-SUMO tag Δ40TorsinA and 6×His-SUMO tag C-terminal of Δ40TorsinA (C44S, C49S, C50S, C162S, C280S, C319S) expression constructs.

The expression construct contains a His-6tag for IMAC and SUMO as the fusion partner. Protein sequence at the junction between the SUMO and target protein is shown, with the SUMO protease cleavage site indicated.

A

MGSSHHHHHGSGLVPRGSASMSDSEVNQEAKPEVKPEVKPETHINLKVSDGSSEIFFKIKKT
TPLRRLMEAFAKRQGKEMDSLRFLYDGIRIQADQTPEDLDMEDNDIIEAHREQIGGRLYCL
FAECCGQKRSLSREALQKDLDDNLFGQHLAKKIILNAVFGFINNPKPKKPLTLSLHGWTGT
GKNFVSKIIAENIYEGGLNSDYVHLFVATLHFPHASNITLYKDQLQLWIRGNVSACARSIFIF
DEMDKMHAGLIDAIKPFLDYYDLVDGVSYQKAMFIFLSNAGAERITDVALDFWRSGKQRE
DIKLKDIEHALSVSVFNNKNSGFWHSSLIDRNLIDYFVPFLPLEYKHLKMCIRVEMQSRGYE
IDEDIVSRVAEEMTFFPKEERVFSDKGCKTVFTKLDYYYDD

B

MGSSHHHHHGSGLVPRGSASMSDSEVNQEAKPEVKPEVKPETHINLKVSDGSSEIFFKIKKT
TPLRRLMEAFAKRQGKEMDSLRFLYDGIRIQADQTPEDLDMEDNDIIEAHREQIGGLEYKH
LKMSIRVEMQSRGYEIDEDIVSRVAEEMTFFPKEERVFSDKGSKTVFTKLDYYYDD

Figure 15: The amino acid sequence of 6×His-SUMO tag Δ40TorsinA and 6×His-SUMO tag C-terminal of Δ40TorsinA (C44S, C49S, C50S, C162S, C280S, C319S).

The polyhistidine region is dotted underlined and the sumo tag is thick underlined. The target protein is marked as red.

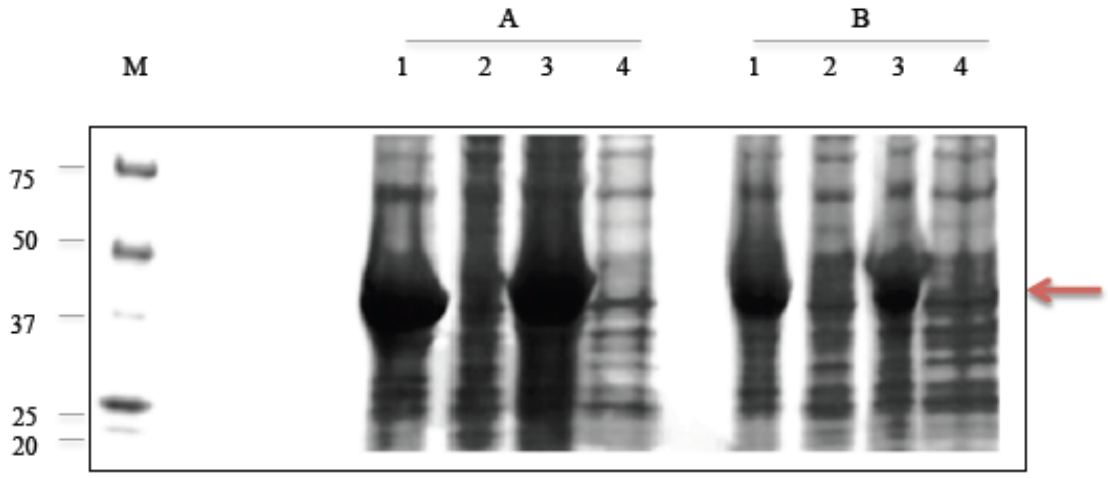


Figure 16: SDS gel electrophoresis analysis of SUMO tag Δ40TorsinA solubility.

Protein expression was induced by 1mM IPTG at 20°C for 16 h (A) or 37°C for 2 h (B).

M: molecular mass ladders.

Lane 1: Induced *E. coli* pellet after sonication.

Lane 2: Induced *E. coli* supernatant after sonication.

Lane 3: *E. coli* cells with 1mM IPTG induction.

Lane 4: *E. coli* cells without IPTG induction.

The arrow indicates the band of SUMO tag Δ40TorsinA.

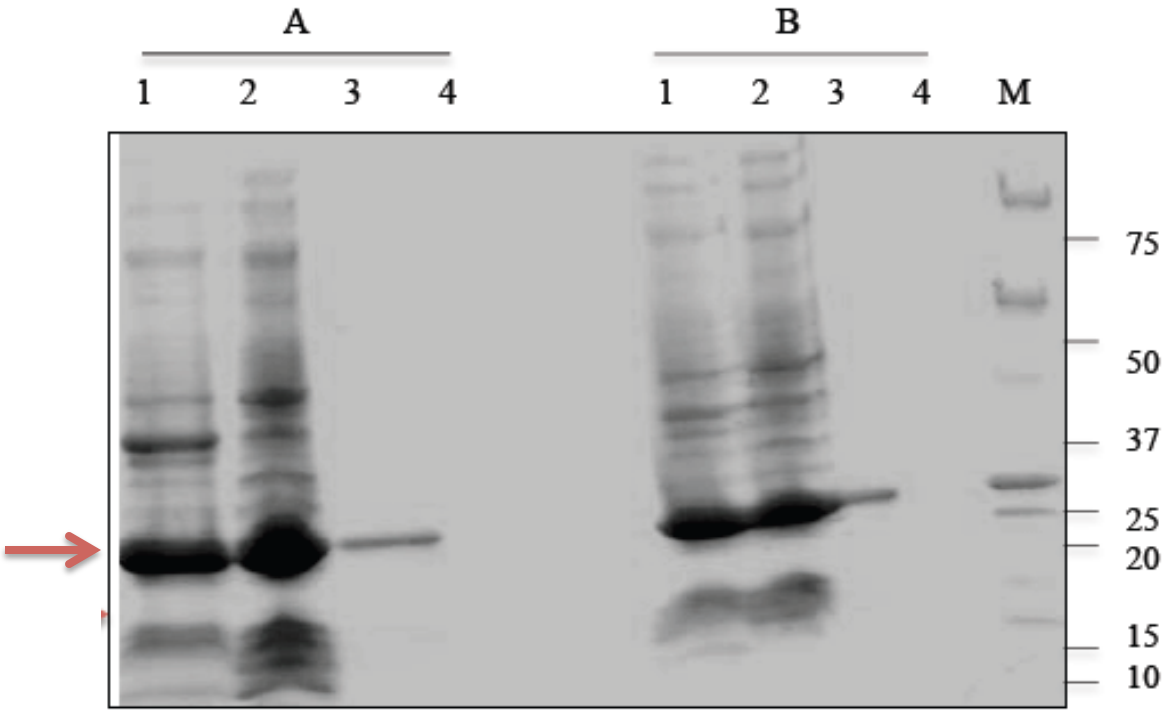


Figure 17: SDS gel electrophoresis analysis of SUMO tag C-terminal of Δ40TorsinA (C44S, C49S, C50S, C162S, C280S, C319S) solubility.

Protein expression was induced by 1mM IPTG at 20°C for 16 h (A) or 37°C for 2 h (B).

M: molecular mass ladders.

Lane 1: Induced *E. coli* pellet after sonication.

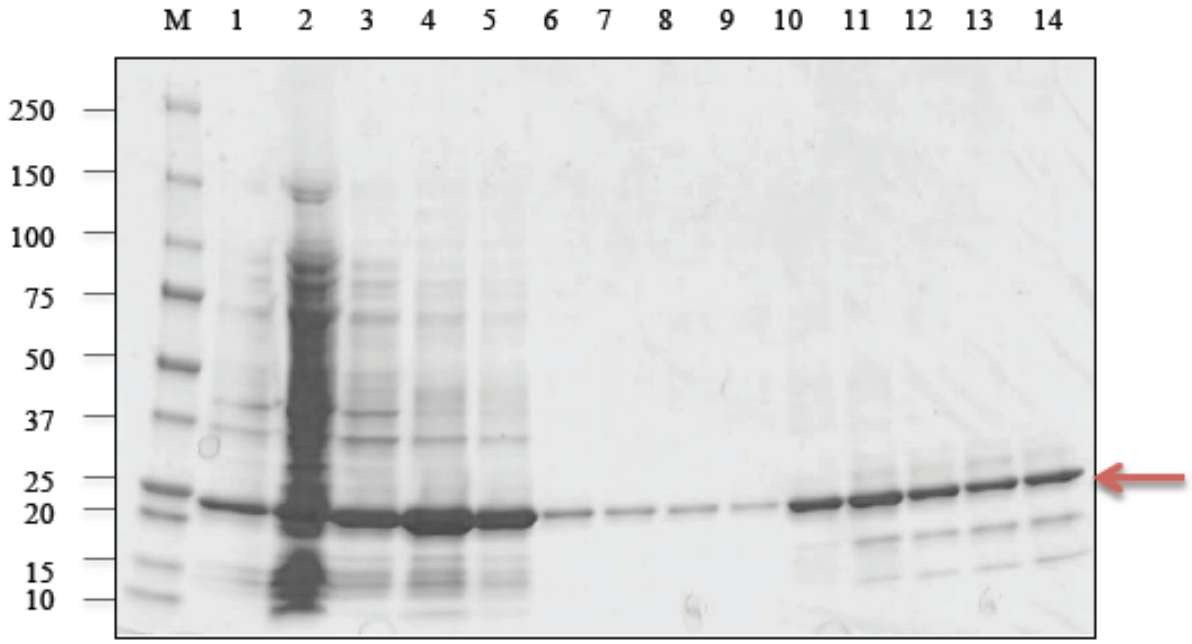
Lane 2: Induced *E. coli* supernatant after sonication.

Lane 3: *E. coli* cells with 1mM IPTG induction.

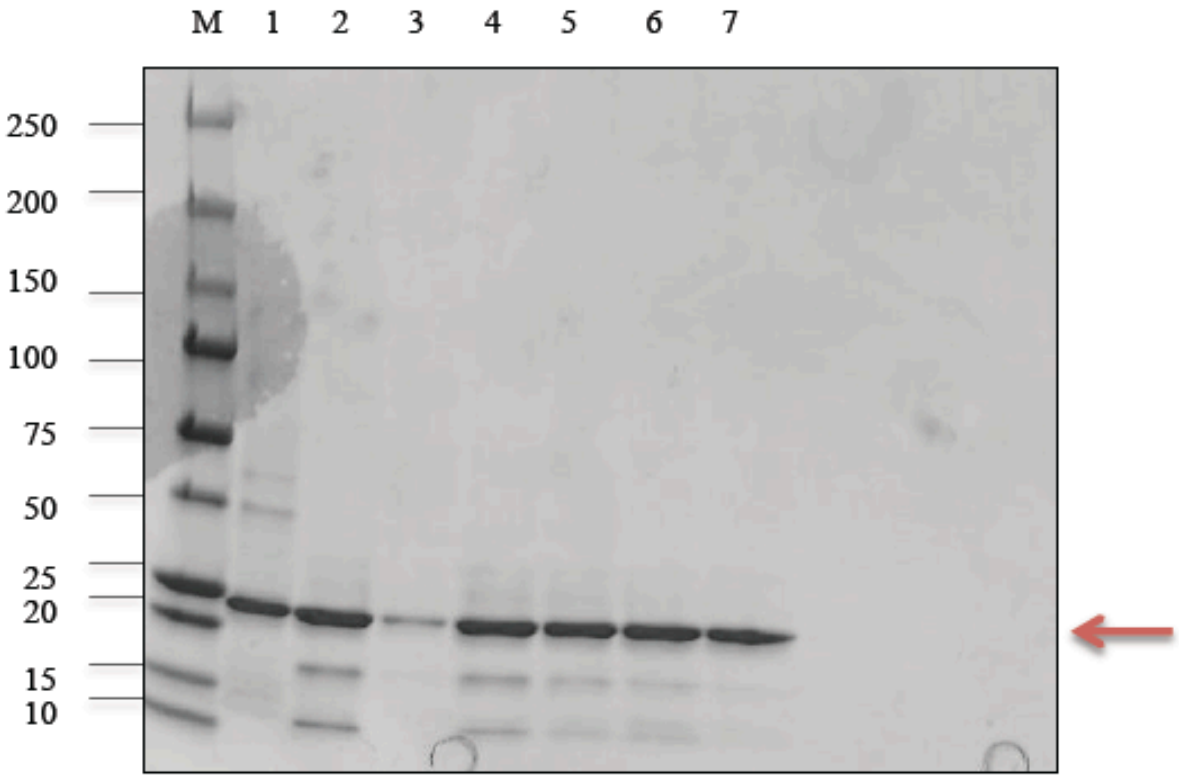
Lane 4: *E. coli* cells without IPTG induction.

The arrow indicates the band of SUMO tag C-terminal of Δ40TorsinA (C44S, C49S, C50S, C162S, C280S, C319S).

A



B



C

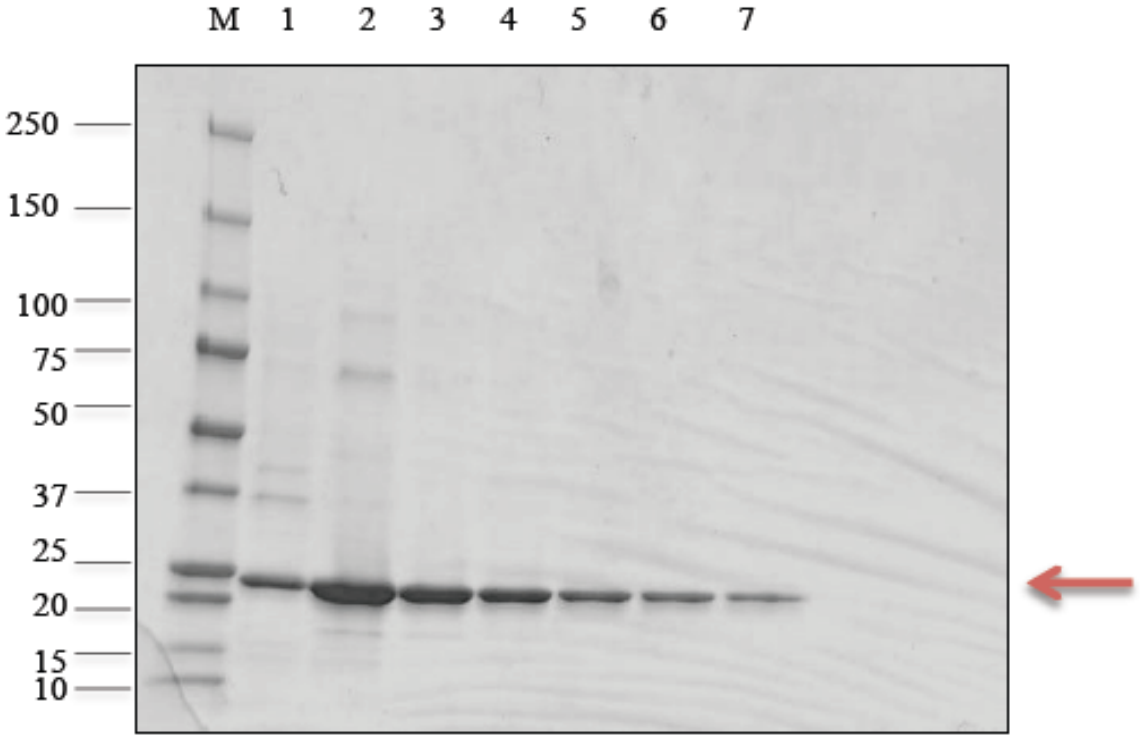


Figure 18: SDS gel electrophoresis analysis of purification of SUMO tag C-terminal of Δ40TorsinA (C44S, C49S, C50S, C162S, C280S, C319S) by using Ni-NTA affinity chromatography under native conditions.

The target proteins were eluted from Ni-NTA column with increased concentration of imidazole. The red arrow indicates the band of SUMO tag C-terminal of Δ40TorsinA (C44S, C49S, C50S, C162S, C280S, C319S).

M: molecular mass ladders.

A: Lane 1: E. coli cells with 1mM IPTG induction at 20°C for 16 hs.

Lane 2: unbound

Lane 3,4: wash

Lane 5, 6, 7, 8: elution with the buffer containing 10mM imidazole.

Lane 9, 10, 11, 12: elution with the buffer containing 50mM imidazole.

Lane 13, 14: elution with the buffer containing 100mM imidazole.

B: Lane 1: E. coli cells with 1mM IPTG induction at 20°C for 16 hs.

Lane 2, 3: elution with the buffer containing 100mM imidazole.

Lane 4, 5, 6, 7: elution with the buffer containing 250mM imidazole.

C: Lane 1: E. coli cells with 1mM IPTG induction at 20°C for 16 hs.

Lane 2, 3, 4, 5, 6, 7: elution with the buffer containing 250mM imidazole.

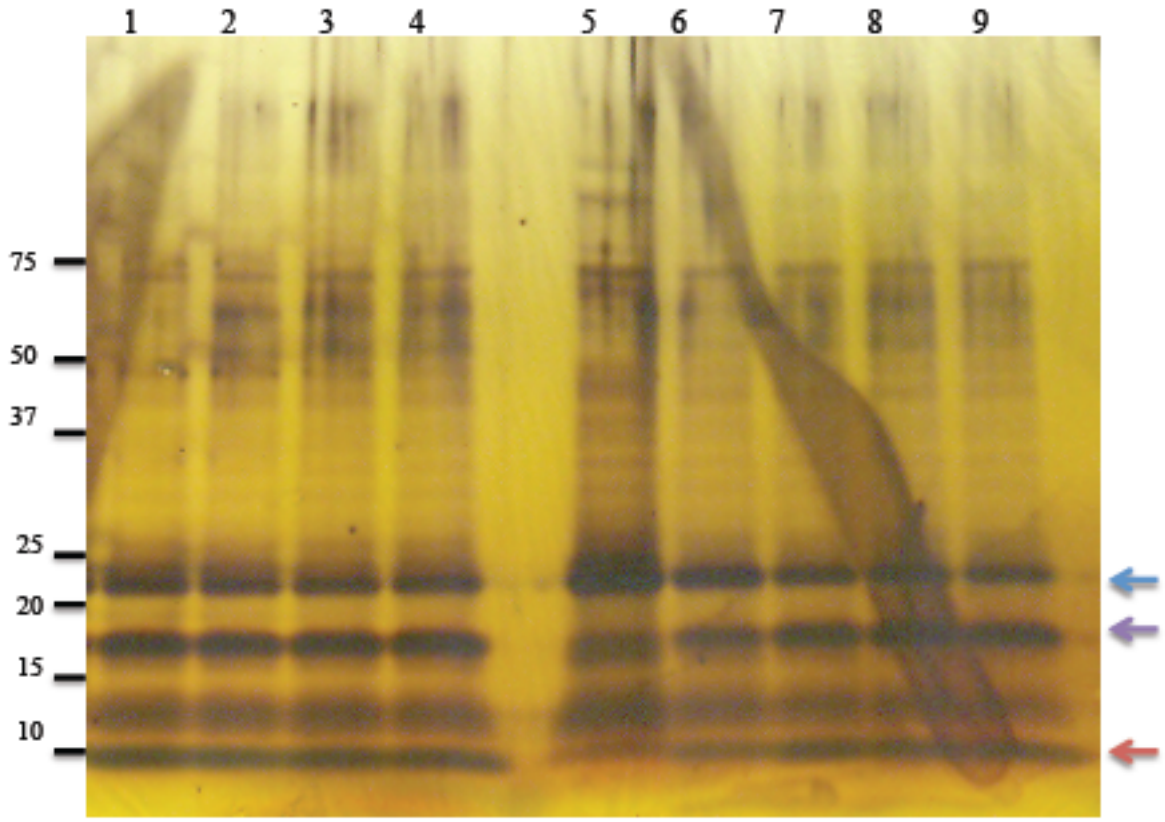
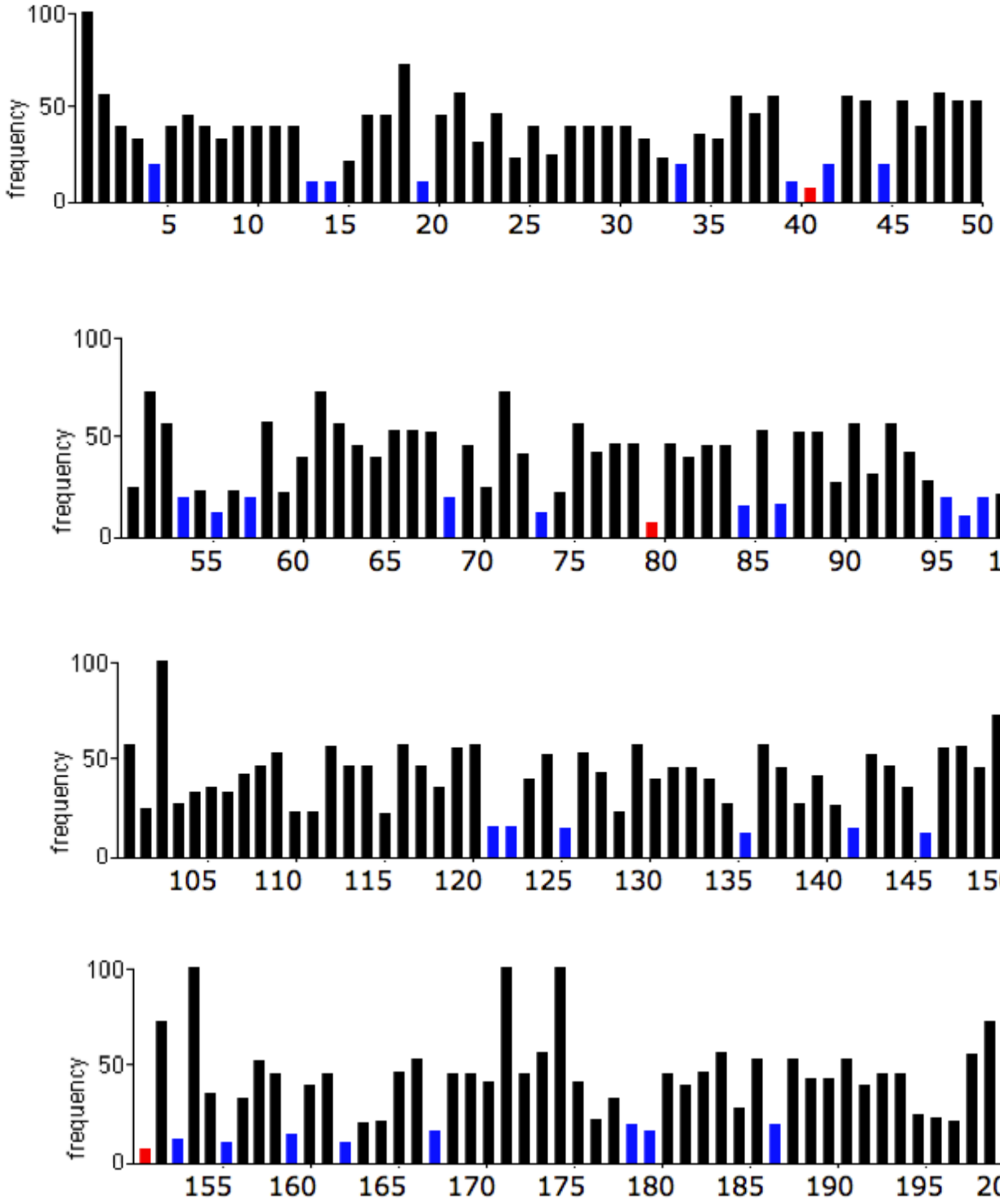


Figure 19: small-scale cleavage reaction of 6×His-SUMO-C-torsinA (C280S, C319S) protein by SUMO protease.

Purified 6×His-SUMO-C-torsinA (C280S, C319S) (20 µg) was incubated in the presence of 10 units (Lane 1, 2, 3, 4) or 20 unites (Lane 6, 7, 8, 9) of SUMO protease at 30°C. 20µl aliquots were withdrawn at 0 (Lane 5), 0.5, 1, 2 and 3 h, and subjected to SDS-PAGE followed by silver staining.

Lane 5: Purified 6×His-SUMO-C-torsinA (C280S, C319S) before cleavage reaction as control.
Lane 1, 2, 3, 4: results from cleavage by 10 units SUMO protease during 0.5, 1, 2, 3h.
Lane 6, 7, 8, 9: results from cleavage by 20 units SUMO protease during 0.5, 1, 2, 3h.
The blue, purple and red arrows indicate likely the position of fusion protein, SUMO and C-torsinA (C280S, C319S).


100
50
0
frequency
5
10
15
20
25
30
35
40
45
50
55
60
65
70
75
80
85
90
95
100
105
110
115
120
125
130
135
140
145
150
155
160
165
170
175
180
185
190
195
200

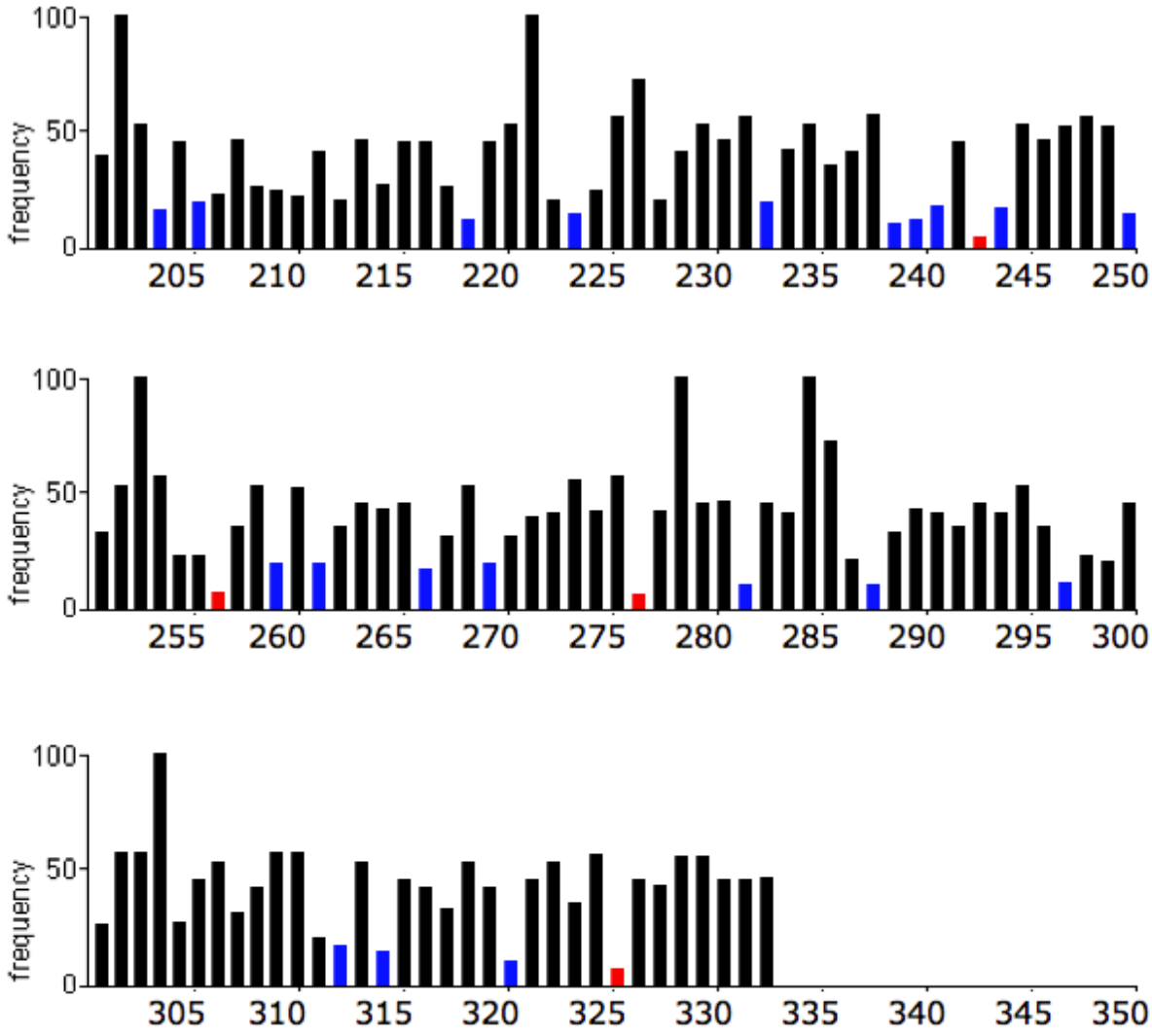
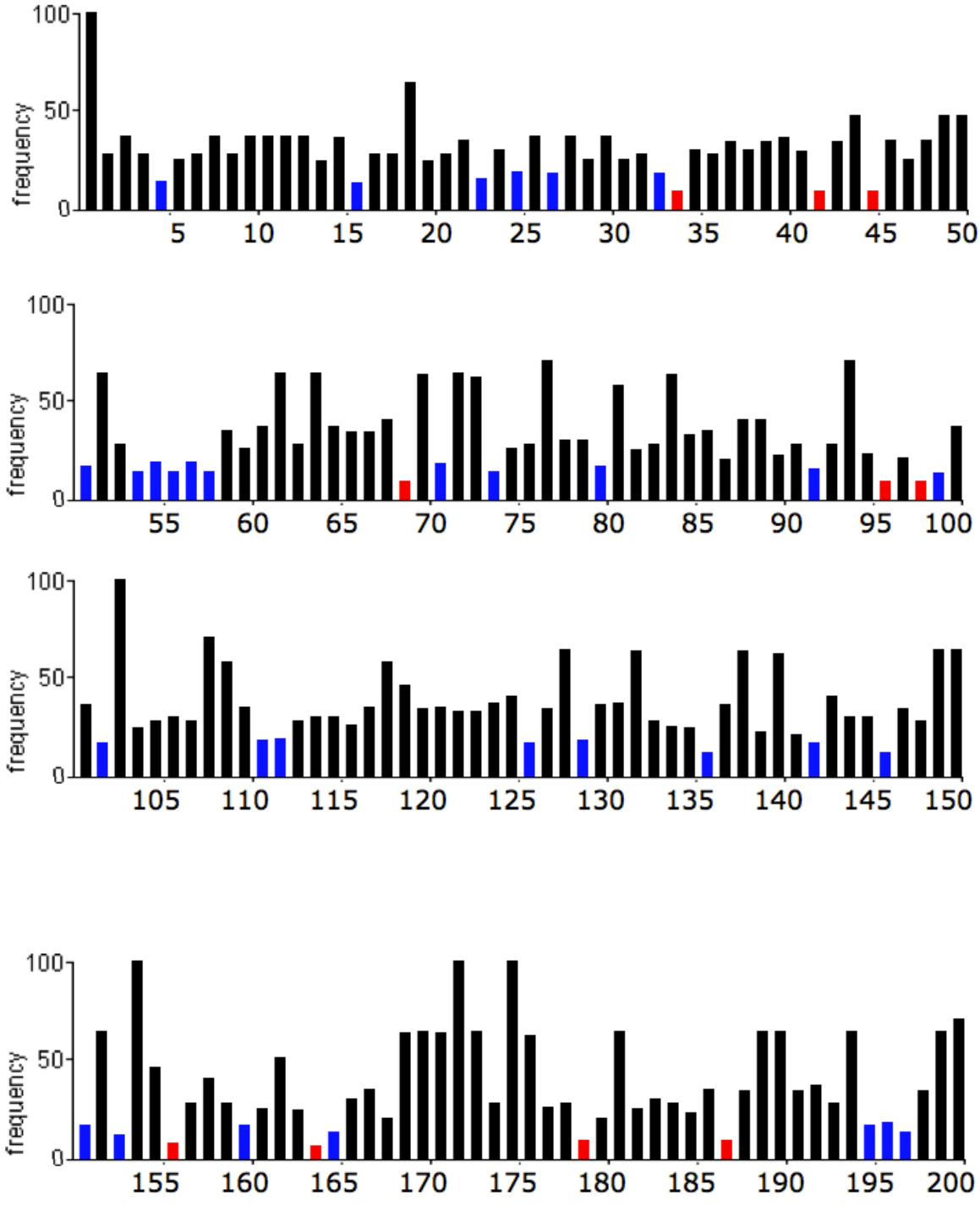


Figure 20: Graph shows the relative codon frequency (RCF) of torsinA.

The RCF index of torsinA was analyzed by comparing the nucleotide sequence with the codon table of human. The graph shows the percentage of usage of every codon of human torsinA. The blue and red bars indicate that are used less than 20% and 10%, respectively.

Amino acid	Rare codon(s)
Arginine	AGG, AGA, CGG, CGA
Leucine	CUA, CUC
Isoleucine	AUA
Serine	UCG, UCA, AGU, UCC
Glycine	GGA, GGG
Proline	CCC, CCU, CCA
Threonine	ACA

Table 1: Rarely used codons in *E. coli* (Lee et al., 2009).


100
50
0
frequency
5 10 15 20 25 30 35 40 45 50
100
50
0
frequency
55 60 65 70 75 80 85 90 95 100
100
50
0
frequency
105 110 115 120 125 130 135 140 145 150
100
50
0
frequency
155 160 165 170 175 180 185 190 195 200

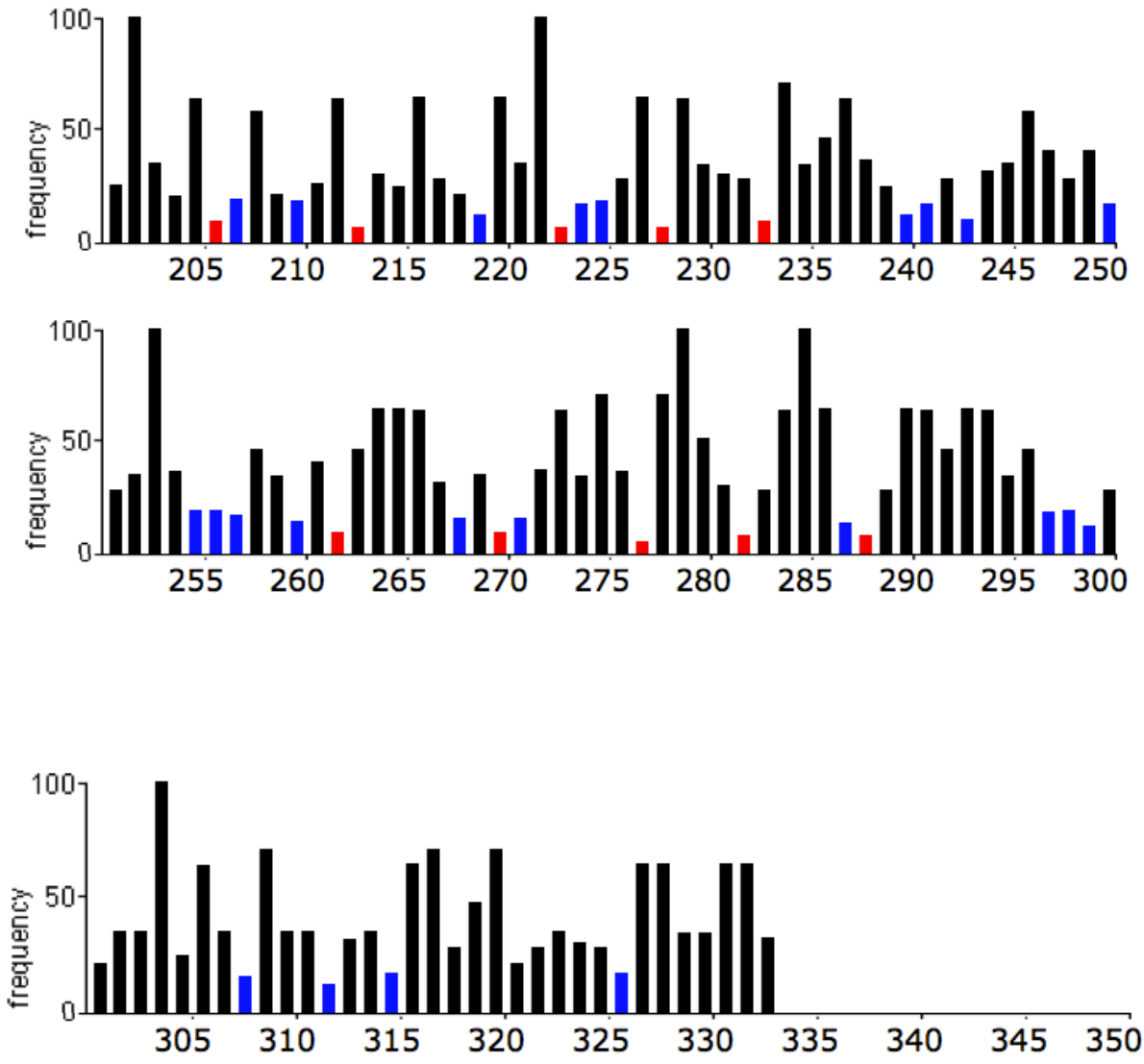


Figure 21: Graph shows the relative codon frequency (RCF) of torsinA expressed in *E. coli.*
The RCF index of torsinA was analyzed by comparing the nucleotide sequence with the codon table of *E. coli.* The graph shows the percentage of usage of every codon of human torsinA expressed in *E. coli*. . The blue and red bars indicate that are used less than 20% and 10%, respectively.

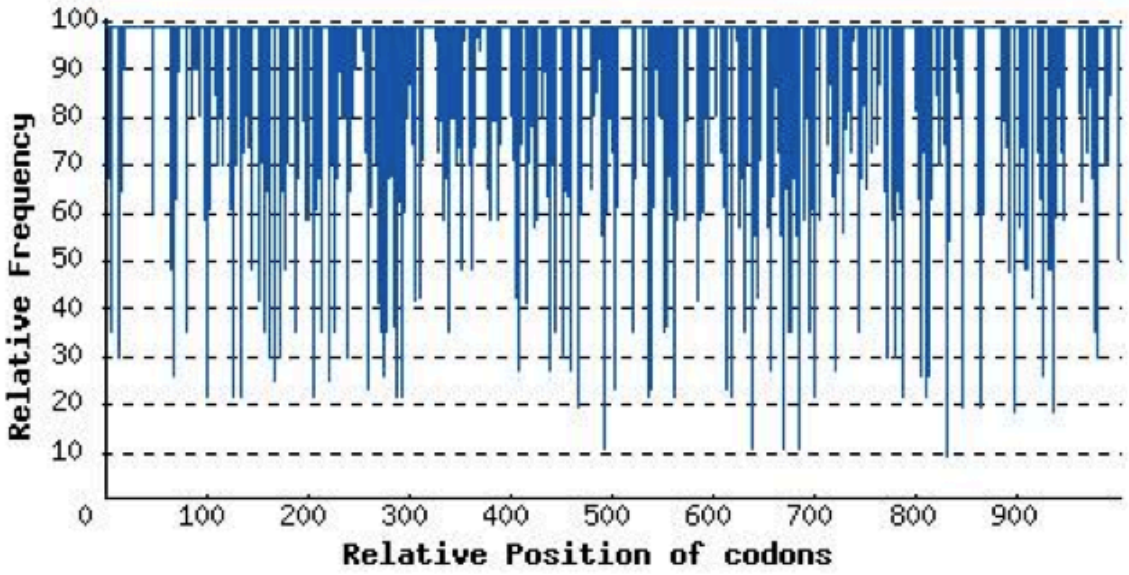


Figure 22: The distribution of codon usage frequency along the length of torsinA expressed in *E. coli*.

CAI is 0.66.

Rare Arg codons (AGG/AGA, CGA)	Rare Ile codons (AUA)	Rare Ieu codons (CUA)	Rare Pro codons (CCC)	Total
7	4	1	4	16

Table 2: the occurrence of the six rarest codons of *E. coli* in torsinA gene.

	TorsinA	C-terminal of AAA+ domain of thermophilus ClpB
Molecular Weight	37808.7	34213.3
Number of amino acids	332	303
Theoretical pI	6.51	5.90
Total number of negatively charged residues (Asp + Glu)	40	47
Total number of positively charged residues (Arg + Lys)	38	43
Aliphatic index	98.98	104.59
Instability index	37.73 (stable)	47.73 (unstable)
Grand average of hydropathicity (GRAVY)	-0.025	-0.214
Amino acid composition		
Ala (A)	6.6%	8.6%
Arg (R)	4.2%	9.9%
Asn (N)	4.2%	1.7%
Asp (D)	6.9%	5.6%
Cys (C)	1.8%	0.0%
Gln (Q)	2.7%	4.0%
Glu (E)	5.1%	9.9%
Gly (G)	6.3%	7.3%
His (H)	2.7%	1.7%
Ile (I)	6.9%	6.9%

Leu (L)	12.0%	11.6%
Lys (K)	7.2%	4.3%
Met (M)	2.1%	1.0%
Phe (F)	6.3%	4.3%
Pro (P)	3.3.%	5.3%
Ser (S)	6.3%	2.6%
Thr (T)	3.0%	4.6%
Trp (W)	1.2%	0.3%
Tyr (Y)	4.5%	2.3%
Val (V)	6.3%	8.3%
Pyl (O)	0.0%	0.0%
Sec (U)	0.0%	0.0%

Table 3: The primary sequence characteristics of torsinA and C-terminal AAA+ domain of *T. thermophilus* ClpB

References

Aguiar, P.M.D., and Ozelius, L.J. (2002). Classification and genetics of dystonia. Lancet Neurol*,* 1, 316-325.

Albanese, A., Barnes, M.P., Bhatia, K.P., Fernandez-Alvarez, E., Filippini, G., Gasser, T., Krauss, J.K., Newton, A., Rektor, I., Savoiardo, M.*, et al.* (2006). A systematic review on the diagnosis and treatment of primary (idiopathic) dystonia and dystonia plus syndromes: report of an EFNS/MDS-ES Task Force. European Journal of Neurology : the Official Journal of the European Federation of Neurological Societies*,* 13, 433-444.

Angov, E., Hillier, C.J., Kincaid, R.L., and Lyon, J.A. (2008). Heterologous Protein Expression Is Enhanced by Harmonizing the Codon Usage Frequencies of the Target Gene with those of the Expression Host. PloS one *3*.

Breakefield, X.O., Blood, A.J., Li, Y., Hallett, M., Hanson, P.I., and Standaert, D.G. (2008). The pathophysiological basis of dystonias. Nature reviews Neuroscience*, 9*, 222-234.

Breakefield, X.O., Kamm, C., and Hanson, P.I. (2001). TorsinA: movement at many levels. Neuron*,* 31, 9-12.

Butt, T.R., Edavettal, S.C., Hall, J.P., and Mattern, M.R. (2005). SUMO fusion technology for difficult-to-express proteins. Protein Expression and Purification*,* 43, 1-9.

Caldwell, G.A., Cao, S., Sexton, E.G., Gelwix, C.C., Bevel, J.P., and Caldwell, K.A. (2003). Suppression of polyglutamine-induced protein aggregation in Caenorhabditis elegans by torsin proteins. Human Molecular Genetics*,* 12, 307-319.

Callan, A.C., Bunning, S., Jones, O.T., High, S., and Swanton, E. (2007). Biosynthesis of the dystonia-associated AAA+ ATPase torsinA at the endoplasmic reticulum. The Biochemical Journal, 401, 607-612.

Cao, S., Hewett, J.W., Yokoi, F., Lu, J., Buckley, A.C., Burdette, A.J., Chen, P., Nery, F.C., Li, Y., Breakefield, X.O.*, et al.* (2010). Chemical enhancement of torsinA function in cell and animal models of torsion dystonia. Disease Models & Mechanisms*,* 3, 386-396.

Chen, X.P., Shang, H.F., and Luo, Z.M. (2008). Genetic classification and molecular mechanisms of primary dystonia. Neural Regen Res, 3, 296-300.

Ciencias, F.d. (2008). Folding at the rhythm of the rare codon beat. Biotechnol J*,* 3, 1047–1057.

Dang, M.T., Yokoi, F., McNaught, K.S.P., Jengelley, T.A., Jackson, T., Li, J.Y., and Li, Y.Q. (2005). Generation and characterization of Dyt1 Delta GAG knock-in mouse as a model for early-onset dystonia. Exp Neurol*,* 196, 452-463.

Dang, M.T., Yokoi, F., Pence, M.A., and Li, Y. (2006). Motor deficits and hyperactivity in Dyt1 knockdown mice. Neurosci Res*,* 56, 470-474.

Davis, G.D., Elisee, C., Newham, D.M., and Harrison, R.G. (1999). New fusion protein systems designed to give soluble expression in Escherichia coli. Biotechnol Bioeng , 65, 382-388.

de Carvalho Aguiar, P.M., and Ozelius, L.J. (2002). Classification and genetics of dystonia. Lancet Neurol, 1, 316-325.

Elmore, Z.C., Donaher, M., Matson, B.C., Murphy, H., Westerbeck, J.W., and Kerscher, O. (2011a). Sumo-dependent substrate targeting of the SUMO protease Ulp1. Bmc Biol, 9, 74.

Elmore, Z.C., Donaher, M., Matson, B.C., Murphy, H., Westerbeck, J.W., and Kerscher, O. (2011b). Sumo-dependent substrate targeting of the SUMO protease Ulp1. Bmc Biol *9*.

Espargaro, A., Castillo, V., de Groot, N.S., and Ventura, S. (2008). The in vivo and in vitro aggregation properties of globular proteins correlate with their conformational stability: the SH3 case. Journal of Molecular Biology, 378, 1116-1131.

Geyer, H.L., and Bressman, S.B. (2006). The diagnosis of dystonia. Lancet Neurol, 5, 780-790.

Giles, L.M., Chen, J., Li, L., and Chin, L.S. (2008). Dystonia-associated mutations cause premature degradation of torsinA protein and cell-type-specific mislocalization to the nuclear envelope. Human Molecular Genetics 17, 2712-2722.

Giles, L.M., Li, L., and Chin, L.S. (2009a). Printor, a novel torsinA-interacting protein implicated in dystonia pathogenesis. The Journal of Biological Chemistry 284, 21765-21775.

Giles, L.M., Li, L., and Chin, L.S. (2009b). TorsinA protein degradation and autophagy in DYT1 dystonia. Autophagy, 5, 82-84.

Goh, C.S., Lan, N., Douglas, S.M., Wu, B.L., Echols, N., Smith, A., Milburn, D., Montelione, G.T., Zhao, H.Y., and Gerstein, M. (2004). Mining the structural genomics pipeline: Identification of protein properties that affect high-throughput experimental analysis. Journal of Molecular Biology*,* 336, 115-130.

Gomes, E.R., Jani, S., and Gundersen, G.G. (2005). Nuclear movement regulated by Cdc42, MRCK, myosin, and actin flow establishes MTOC polarization in migrating cells. Cell, 121, 451-463.

Gonzalez-Alegre, P., Bode, N., Davidson, B.L., and Paulson, H.L. (2005). Silencing primary dystonia: Lentiviral-mediated RNA interference therapy for DYT1 dystonia. Journal of Neuroscience, 25, 10502-10509.

Goodchild, R.E., and Dauer, W.T. (2004). Mislocalization to the nuclear envelope: An effect of the dystonia-causing torsinA mutation. Proceedings of the National Academy of Sciences of the United States of America, 101, 847-852.

Goodchild, R.E., and Dauer, W.T. (2005). The AAA+ protein torsinA interacts with a conserved domain present in LAP1 and a novel ER protein. The Journal of Cell Biology, 168, 855-862.

Goodchild, R.E., Kim, C.E., and Dauer, W.T. (2005). Loss of the dystonia-associated protein torsinA selectively disrupts the neuronal nuclear envelope. Neuronscience, 48, 923-932.

Gordon, K.L., and Gonzalez-Alegre, P. (2008). Consequences of the DYT1 mutation on torsinA oligomerization and degradation. Neuroscience, 157, 588-595.

Granata, A., Schiavo, G., and Warner, T.T. (2009). TorsinA and dystonia: from nuclear envelope to synapse. Journal of Neurochemistry, 109, 1596-1609.

Granata, A., Watson, R., Collinson, L.M., Schiavo, G., and Warner, T.T. (2008). The dystonia-associated protein torsinA modulates synaptic vesicle recycling. Journal of Biological Chemistry, 283, 7568-7579.

Gustafsson, C., Govindarajan, S., and Minshull, J. (2004). Codon bias and heterologous protein expression. Trends in Biotechnology, 22, 346-353.

Hewett, J., Gonzalez-Agosti, C., Slater, D., Ziefer, P., Li, S., Bergeron, D., Jacoby, D.J., Ozelius, L.J., Ramesh, V., and Breakefield, X.O. (2000). Mutant torsinA, responsible for early-onset torsion dystonia, forms membrane inclusions in cultured neural cells. Human Molecular Genetics, 9, 1403-1413.

Hewett, J., Ziefer, P., Bergeron, D., Naismith, T., Boston, H., Slater, D., Wilbur, J., Schuback, D., Kamm, C., Smith, N., *et al.* (2003a). TorsinA in PC12 cells: localization in the endoplasmic reticulum and response to stress. Journal of Neuroscience Research, 72, 158-168.

Hewett, J., Ziefer, P., Bergeron, D., Naismith, T., Boston, H., Slater, D., Wilbur, J., Schuback, D., Kamm, C., Smith, N., *et al.* (2003b). TorsinA in PC12 cells: Localization in the endoplasmic reticulum and response to stress. Journal of Neuroscience Research, 72, 158-168.

Hewett, J.W., Zeng, J., Niland, B.P., Bragg, D.C., and Breakefield, X.O. (2006). Dystonia-causing mutant torsinA inhibits cell adhesion and neurite extension through interference with cytoskeletal dynamics. Neurobiology of Disease, 22, 98-111.

Idicula-Thomas, S., and Balaji, P.V. (2005). Understanding the relationship between the primary structure of proteins and its propensity to be soluble on overexpression in Escherichia coli. Protein Science, 14, 582-592.

Jankovic, J. (2006). Botulinum toxin therapy for cervical dystonia. Neurotox Res, 9, 145-148.

Jinnah, H.A., and Hess, E.J. (2008). Experimental therapeutics for dystonia. Neurotherapeutics, 5, 198-209.

Johanssen, V.A., Barnham, K.J., Masters, C.L., Hill, A.F., and Collins, S.J. (2012). Generating recombinant C-terminal prion protein fragments of exact native sequence. Neurochemistry International, 60, 318-326.

Jungwirth, M., Dear, M.L., Brown, P., Holbrook, K., and Goodchild, R. (2010). Relative tissue expression of homologous torsinB correlates with the neuronal specific importance of DYT1 dystonia-associated torsinA. Human Molecular Genetics, 19, 888-900.

Kamm, C. (2006). Early onset torsion dystonia (Oppenheim's dystonia). Orphanet J Rare Dis *1*.

Kamm, C., Boston, H., Hewett, J., Wilbur, J., Corey, D.P., Hanson, P.I., Ramesh, V., and Breakefield, X.O. (2004). The early onset dystonia protein torsinA interacts with kinesin light chain 1. The Journal of Biological Chemistry, 279, 19882-19892.

Kane, J.F. (1995). Effects of rare codon clusters on high-level expression of heterologous proteins in Escherichia coli. Current Opinion in Biotechnology, 6, 494-500.

Koh, Y.H., Rehfeld, K., and Ganetzky, B. (2004). A Drosophila model of early onset torsion dystonia suggests impairment in TGF-beta signaling. Human Molecular Genetics, 13, 2019-2030.

Konakova, M., and Pulst, S.M. (2005a). Dystonia-associated forms of torsinA are deficient in ATPase activity. Journal of Molecular Neuroscience, 25, 105-117.

Konakova, M., and Pulst, S.M. (2005b). Dystonia-associated forms of torsinA are deficient in ATPase activity. Journal of Molecular Neuroscience, 25, 105-117.

Kustedjo, K., Deechongkit, S., Kelly, J.W., and Cravatt, B.F. (2003). Recombinant expression, purification, and comparative characterization of torsinA and its torsion dystonia-associated variant Delta E-torsinA. Biochemistry, 42, 15333-15341.

Kyte, J., and Doolittle, R.F. (1982). A simple method for displaying the hydropathic character of a protein. Journal of Molecular Biology, 157, 105-132.

Lange, N., Hamann, M., Shashidharan, P., and Richter, A. (2011). Behavioural and pharmacological examinations in a transgenic mouse model of early-onset torsion dystonia. Pharmacology, Biochemistry, and Behavior, 97, 647-655.

Lee, S.F., Li, Y.J., and Halperin, S.A. (2009). Overcoming codon-usage bias in heterologous protein expression in Streptococcus gordonii. Microbiology, 155, 3581-3588.

Liu, Z., Zolkiewska, A., and Zolkiewski, M. (2003). Characterization of human torsinA and its dystonia-associated mutant form. The Biochemical Journal, 374, 117-122.

Lu, W., Cai, X., Gu, Z., Huang, Y., Xia, B., and Cao, P. (2012). Production and Characterization of Hirudin Variant-1 by SUMO Fusion Technology in E. coli. Molecular Biotechnology.

Lu, W., Cao, P., Cai, X., Yu, J., Hu, C., Cao, M., and Zhang, S. (2009). Molecular cloning, expression, and bioactivity of dove B lymphocyte stimulator (doBAFF). Veterinary Immunology and Immunopathology, 128, 374-380.

Marsden, C.D. (1976). Dystonia: the spectrum of the disease. Research publications - Association for Research in Nervous and Mental Disease, 55, 351-367.

McLean, P.J., Kawamata, H., Shariff, S., Hewett, J., Sharma, N., Ueda, K., Breakefield, X.O., and Hyman, B.T. (2002). TorsinA and heat shock proteins act as molecular chaperones: suppression of alpha-synuclein aggregation. Journal of Neurochemistry, 83, 846-854.

Melchior, F. (2000). SUMO--nonclassical ubiquitin. Annual Review of Cell and Developmental Biology, 16, 591-626.

Muller, U. (2006). Dystonia. The New England Journal of Medicine, 355, 1935; author reply 1935.

Nery, F.C., Zeng, J., Niland, B.P., Hewett, J., Farley, J., Irimia, D., Li, Y.Q., Wiche, G., Sonnenberg, A., and Breakefield, X.O. (2008). TorsinA binds the KASH domain of nesprins and participates in linkage between nuclear envelope and cytoskeleton. Journal of Cell Science, 121, 3476-3486.

O'farrell, C., Hernandez, D.G., Evey, C., Singleton, A.B., and Cookson, M.R. (2002). Normal localization of Delta F323-Y328 mutant torsinA in transfected human cells. Neurosci Lett, 327, 75-78.

Rostasy, K., Augood, S.J., Hewett, J.W., Leung, J.C.O., Sasaki, H., Ozelius, L.J., Ramesh, V., Standaert, D.G., Breakefield, X.O., and Hedreen, J.C. (2003). TorsinA protein and neuropathology in early onset generalized dystonia with GAG deletion. Neurobiology of Disease, 12, 11-24.

Roubertie, A., Echenne, B., Cif, L., Vayssiere, N., Hemm, S., and Coubes, P. (2000). Treatment of early-onset dystonia: update and a new perspective. Child's nervous system : ChNS : Official Journal of the International Society for Pediatric Neurosurgery, 16, 334-340.

Sharma, N., Baxter, M.G., Petravicz, J., Bragg, D.C., Schienda, A., Standaert, D.G., and Breakefield, X.O. (2005). Impaired motor learning in mice expressing TorsinA with the DYT1 dystonia mutation. Journal of Neuroscience, 25, 5351-5355.

Sharp, P.M., and Li, W.H. (1987). The codon Adaptation Index--a measure of directional synonymous codon usage bias, and its potential applications. Nucleic Acids Research, 15, 1281-1295.

Shashidharan, P., Kramer, B.C., Walker, R.H., Olanow, C.W., and Brin, M.F. (2000). Immunohistochemical localization and distribution of torsinA in normal human and rat brain. Brain Research, 853, 197-206.

Shashidharan, P., Sandu, D., Potla, U., Armata, I.A., Walker, R.H., McNaught, K.S., Weisz, D., Sreenath, T., Brin, M.F., and Olanow, C.W. (2005). Transgenic mouse model of early-onset DYT1 dystonia. Human Molecular Genetics, 14, 125-133.

Siegert, S., Bahn, E., Kramer, M.L., Schulz-Schaeffer, W.J., Hewett, J.W., Breakefield, X.O., Hedreen, J.C., and Rostasy, K.M. (2005). TorsinA expression is detectable in human infants as young as 4 weeks old. Dev Brain Res, 157, 19-26.

Smialowski, P., Martin-Galiano, A.J., Mikolajka, A., Girschick, T., Holak, T.A., and Frishman, D. (2007). Protein solubility: sequence based prediction and experimental verification. Bioinformatics, 23, 2536-2542.

Tarsy, D., and Simon, D.K. (2006). Dystonia. The New England Journal of medicine, 355, 818-829M

Tegel, H., Tourle, S., Ottosson, J., and Persson, A. (2010). Increased levels of recombinant human proteins with the Escherichia coli strain Rosetta(DE3). Protein Expression and Purification, 69, 159-167.

Terpe, K. (2006). Overview of bacterial expression systems for heterologous protein production: from molecular and biochemical fundamentals to commercial systems. Applied Microbiology and Biotechnology, 72, 211-222.

Torres, G.E., Sweeney, A.L., Beaulieu, J.M., Shashidharan, P., and Caron, M.G. (2004). Effect of torsinA on membrane proteins reveals a loss of function and a dominant-negative phenotype of the dystonia-associated DeltaE-torsinA mutant. Proceedings of the National Academy of Sciences of the United States of America, 101, 15650-15655.

Vander Heyden, A.B., Naismith, T.V., Snapp, E.L., and Hanson, P.I. (2011). Static retention of the lumenal monotopic membrane protein torsinA in the endoplasmic reticulum. The EMBO Journal, 30, 3217-3231.

Vander Heyden, A.B., Naismith, T.V., Snapp, E.L., Hodzic, D., and Hanson, P.I. (2009). LULL1 Retargets TorsinA to the Nuclear Envelope Revealing an Activity That Is Impaired by the DYT1 Dystonia Mutation. Molecular Biology of the Cell, 20, 2661-2672.

Walker, R.H., Brin, M.F., Sandu, D., Good, P.F., and Shashidharan, P. (2002). TorsinA immunoreactivity in brains of patients with DYT1 and non-DYT1 dystonia. Neurology, 58, 120-124.

Widmann, M., Clairo, M., Dippon, J., and Pleiss, J. (2008). Analysis of the distribution of functionally relevant rare codons. Bmc Genomics, 9, 207.

Xiao, H.F., Gong, S.Z., Zhao, Y., and LeDoux, M.S. (2004). Developmental expression of rat torsinA transcript and protein. Dev Brain Res, 152, 47-60.

Yokoi, F., Dang, M.T., Li, J., Standaert, D.G., and Li, Y. (2011). Motor deficits and decreased striatal dopamine receptor 2 binding activity in the striatum-specific Dyt1 conditional knockout mice. PloS one *6*, 24539.

Yokoi, F., Dang, M.T., Mitsui, S., Li, J., and Li, Y. (2008). Motor deficits and hyperactivity in cerebral cortex-specific Dyt1 conditional knockout mice. Journal of Biochemistry, 143, 39-47.

Zhang, G., Hubalewska, M., and Ignatova, Z. (2009). Transient ribosomal attenuation coordinates protein synthesis and co-translational folding. Nature Structural & Molecular Biology, 16, 274-280.

Zhu, L., Wrabl, J.O., Hayashi, A.P., Rose, L.S., and Thomas, P.J. (2008). The torsin-family AAA+ protein OOC-5 contains a critical disulfide adjacent to Sensor-II that couples redox state to nucleotide binding. Molecular Biology of the Cell, 19, 3599-3612.

Zolkiewski, M. (1999). ClpB cooperates with DnaK, DnaJ, and GrpE in suppressing protein aggregation. A novel multi-chaperone system from Escherichia coli. The Journal of Biological Chemistry, 274, 28083-28086.

Zolkiewski, M. (2006). A camel passes through the eye of a needle: protein unfolding activity of Clp ATPases. Molecular microbiology, 61, 1094-1100.

Zolkiewski, M.a.W., H.-C. (2011). Emerging Area: TorsinA, a Novel ATP-Dependent Factor Linked to Dystonia. Protein Chaperones and Protection from Neurodegenerative Diseases (ed S N Witt), 10.1002.

Appendix A - Sequence of the primers used in mutagenesis

C280S:

5’- CAAACACCTAAAAATGTCTAT**CC**GAGTGGAAATGCAGTCCCG -3’

C319S: -

5’ - TTCTCAGATAAAGGCT**CC**AAAACGGTGTTCACC -3’

ΔE302/303:

5’ – GTAAGCAGAGTGGCT**GAG**ATGACATTTTTCCCC -3’