In vitro characterization of *STUB1* mutations in recessively inherited spinocerebellar ataxia-16

Yasaman Pakdaman

This thesis is submitted in partial fulfilment of the requirements for the degree of Master in Medical Biology – Medical Cell Biology

Department of Biomedicine,

University of Bergen

Center for Medical Genetics and Molecular Medicine,

Haukeland University Hospital, Bergen

May 2016

Acknowledgements

The presented study was conducted at the Center for Medical Genetics and Molecular Medicine (MGM), Haukeland University Hospital and at the Department of Biomedicine, University of Bergen, during the period August 2015 to May 2016.

I would first like to express my deepest gratitude to my head supervisor Professor Stefan Johansson and co-supervisors Dr. Ingvild Aukrust and Dr. Per Morten Knappskog for the enormous amount of support and patient guidance they have provided at all levels of this work. I am truly grateful for sharing with me your knowledge and expertise. I would not have finished this thesis without your continuous motivation and encouragement.

A very special thanks goes to my beloved friend and former workmate Monica Sanchez Guixe for helping me at the beginning on how to find my way in the laboratory and providing me with a lot of useful information regarding ataxia as well as guiding me throughout the methods and experiments.

Furthermore, I would like to thank Dr. Rune Kleppe for his much appreciated help and kind supervising during the chromatography procedures. I would also like to thank all the laboratory technicians at the MGM, especially Sigrid Erdal for her great advice and nice support in this project. Moreover, I highly acknowledge Professor Ian F. Pryme for his genial assistance in proofreading and correcting my essay. Special thanks go out also to Ehsan Hajjar for sharing his experience and clever ideas while providing the figures.

Finally, I would like to give my sincere thanks to all the new and fantastic friends I got at MGM: Laeya Abdoli Najmi, Khadija El-Jellas, Alba Kaci, and Dinka Smajlagic for their constant love and immediate helps; and to my family, who have always encouraged me to follow my dreams. Thank you for your unflagging love and unconditional support throughout my life and my studies. I am forever grateful to you.

Bergen, May 2016

Yasaman Pakdaman

Content

Abbreviations

ATP	Adenosine triphosphate
BCA	Bicinchoninic acid assay
BP	Base pair
CD	Circular dichroism
CHIP	C-terminus of Hsc70-interacting protein
DTT	Dithiothreitol
ECL	Enhanced chemiluminescence
EDTA	Ethylenediaminetetraacetic acid
EEVD	Amino acid sequence: Glu-Glu-Val-Asp
FBS	Fetal bovine serum
HECT	Homologous to the E6-AP Carboxyl Terminus
HEK	Human embryonic kidney
Hsp/Hsc	Heat shock protein / heat shock cognate
IPTG	Isopropyl-Beta-D-Thiogalactoside
Kd	Dissociation constant
LDS	Lithium dodecyl sulfate
MBP	Maltose binding protein
MST	Microscale thermophoresis
OMIM*	Online Mendelian Inheritance in Man, catalogue number
PBD	Polypeptide binding domain
PBST	Phosphate buffer saline + Tween
PCR	Polymerase chain reaction
RING	Really interesting new gene
SD	Standard deviation
SDS-PAGE	Sodium dodecyl sulfate Polyacrylamide gel electrophoresis
TEV	Tabaco etch virus
Tm	Midpoint of the thermal denaturation
TPR	Tetratricopeptide repeat
UFD	Ubiquitin fusion degradation
UPP	Ubiquitin-proteasome pathway
UV	Ultra violet
WES	Whole exome sequencing
WT	Wild-type
β-ME	β- mercaptoethanol

*The abbreviations for the cerebellar ataxia diseases are given in Figure 1.6.1 and Table 1.7.1.

Abstract

Autosomal **R**ecessive **C**erebellar **A**taxia (ARCA) develops as a result of inefficient protein turnover and further accumulation of damaged proteins inside the cells. Some of mutations associated with ARCA are identified in the *STUB1* gene which encodes CHIP (C-terminus of Heat shock protein 70-Interacting Protein) as a dimeric co-chaperone and E3 ubiquitin ligase protein. Currently, there is limited information available regarding the role of these mutations in the pathogenesis of ARCA. This study aims to characterize the effects of six selected mutations on the structure and function of CHIP, mainly by the use of *in vitro* methods.

Mutant CHIP constructs were created by site-directed mutagenesis, and recombinant proteins were subjected to different assays including ubiquitination activity assay, limited proteolysis, oligomerization analysis, and circular dichroism spectroscopy. Results indicated that mutations N65S and T246M were associated with more severe consequences compared to the rest (E28K, K145Q, M211I, and S236T). Impaired Hsc70-ubiquitination activity was observed for both N65S and T246M. The N65S mutation resulted in more stabilized and compact protein structure, while the T246M mutant presented more flexible and loose structures with a high tendency for aggregation. The other mutations investigated were associated with intact ubiquitination activity, but showed lower protein stability and loss of some secondary structures, although to a lesser extent when compared to T246M.

Altogether, this study provides insights into the development of ARCA through demonstrating the effect of *STUB1* mutations in the expression of damaged CHIP proteins unable to participate efficiently in the protein turnover system, which may lead to toxic accumulation of abnormal proteins and cell death.

1. Introduction

1.1 Protein quality-control system

Being the chief actors within the cells, proteins are involved in nearly every concept of cellular life including shape, structure, biochemical pathways, movement, communication, and signaling. Protein functionality requires achieving a stable, defined native structure as well as maintaining conformational flexibility. Therefore, all the proteins must be folded correctly after being produced by the ribosomes. However, the optimal folded state of the proteome is constantly under the effect of changes in the environment such as heat, oxidation, and ultraviolet stress, causing subjection to various types of protein misfolding [1]. In addition to this, the high density of protein molecules inside the cytosol can increase the likelihood of protein denaturation and aggregation [2]. For this reason, a protein quality control system has been designed inside the cells to ensure that all the proteins are folded properly and to get rid of abnormal proteins. Marques *et al*. introduced a model called the "Triage Model" to describe this system in 2006 [3].

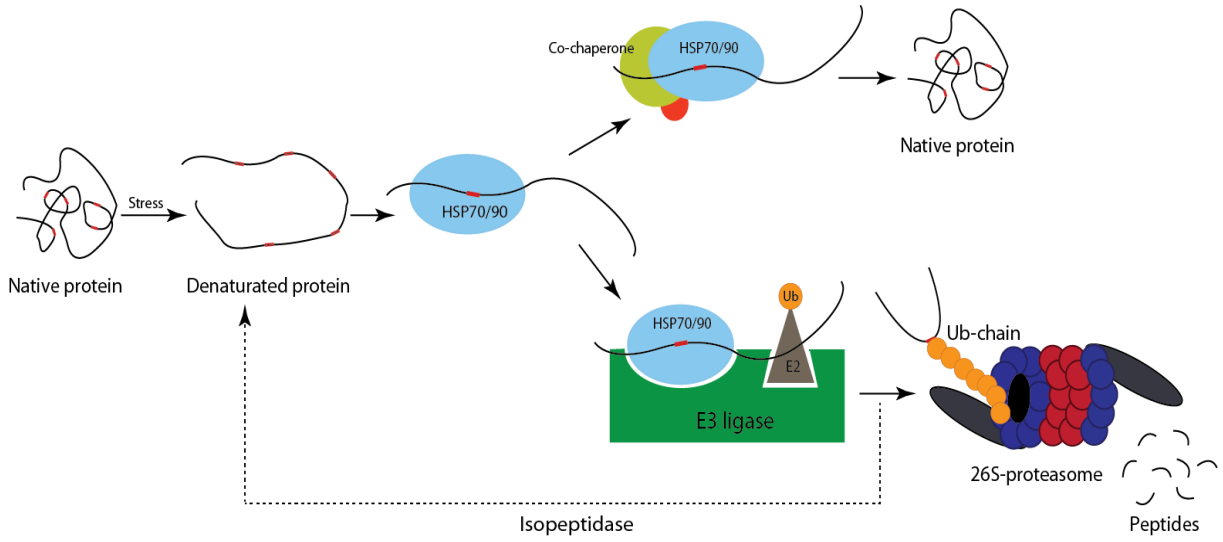


Figure 1.1.1 Triage model of protein quality control system. The intrinsic signals (red) for interaction with molecular chaperones or the components of ubiquitination system are normally hidden inside the protein native structure. Under stressful conditions, such as heat or oxidation, these signals are exposed as a result of protein unfolding. These areas are then recognized by Hsp70 or Hsp90 molecular chaperones which function in collaboration with co-chaperones to refold the protein structure. If the protein cannot be refolded in this way, it will be instead targeted for ubiquitination and proteosomal degradation. However, the ubiquitinated protein might still have a second chance of refolding with deubiquitination activity of isopeptidases. Thus, the efficiency of protein quality control system relies on the competitive relationship existing between the molecular chaperones and the ubiquitin-proteasome pathway (UPP). Ub, Ubiquitin. This Figure is adapted from [3].

According to this model, a protein turnover system functions through the interaction between molecular chaperones and the ubiquitin-proteasome pathway (UPP). Most of the proteins are predicted to have intrinsic signals such as hydrophobic patches hidden in their native conformation which can interact with molecular chaperones and ubiquitin-proteasome system compartments. Once the protein is denatured, these recognition signals will be exposed to molecular chaperones such as Hsp90 or Hsp70 and the protein will be refolded in an ATP-dependent manner. Those proteins which are not able to be refolded rapidly in this pathway will be ubiquitinated and degraded by the 26S proteasome, unless they can be alternatively deubiquitinated by isopeptidases once again and have a second chance of being refolded by molecular chaperones. Therefore, there is a competitive functional relationship between the UPP and molecular chaperones which determines the fate of a damaged protein inside the cell (Figure 1.1.1) [3-5].

1.2 Molecular chaperones help to fold/refold the proteins

The initial folding of proteins necessitates participation of molecular chaperones to help them achieve a stable conformation. Some proteins do not interact with chaperones due to the high stability of their structure. Yet, about 10-20% of newly synthesized proteins are found associated with chaperones once they are out of the ribosome "exit site" [6, 7]. In addition, many proteins may suffer damage as a result of environmental stress including oxidative mechanisms and heat shock, as well as different mutations in their encoding genes. These proteins are recognized and refolded by molecular chaperones in the first level of the protein quality control system. This implies the crucial role of chaperones against protein misfolding and aggregation. There are several highly conserved families of the so-called "Heat shock proteins (Hsp)" involved in protein folding and remodeling. These proteins work as ATP-dependent chaperones together with several regulatory co-factors (co-chaperones). The two most well-studied families of the heat shock proteins are Hsp (Hsc)70 and Hsp90.

The family of Hsp70 chaperones represents the central component of the protein quality control system, assisting a large variety of protein *de novo* folding processes in the cell. It has been shown that ATP cycles of binding and hydrolysis are essential for the activity of these chaperones. In the ATP-bound state, Hsp70 makes short-term interactions with the peptide substrate. Yet, upon ATP hydrolysis, a conformational change in Hsp70 structure traps the

substrate within the amino terminal Polypeptide Binding Domain (PBD) of the chaperone, resulting in longer association of the substrate within the PBD. This allosteric regulation of Hsp70 binding affinity is controlled by specific co-chaperones and nucleotide exchange factors [8].

There are several Hsp70 isoforms present in the cell having significant degrees of sequence conservation and overlapped functions. They are mainly detected in the cytosol, but can also perform essential functions in endoplasmic reticulum and mitochondria. In spite of similarities in their structure and enzymatic activity, these isoforms differ remarkably in the way their expression is regulated in each hosting tissue and during different developmental stages. Genome sequencing studies confirmed the presence of 17 genes and 30 pseudogenes with a high number of mRNA variants that encode Hsp70 isoforms both as organelle-specific and cytosolic chaperones. According to these studies, the organelle-specific Hsp70s are generally encoded by a single gene while those isoforms existing in the same cytosol are very often encoded by different homologous genes. These cytosolic Hsp70s consist of six canonical members including Hsp70-1a, Hsp70-1b, Hsp70-1t, Hsp70-2, Hsp70-6 and Hsc70. Some isoforms, including Hsc70 are constitutively expressed and play important housekeeping functions in the cell. However, the expression of other isoforms such as Hsp70-1a, Hsp70-1b and Hsp70-6 is induced under stress conditions. For each type of harsh condition, there are only specific isoforms capable of efficient binding to the protein substrates and/or co-chaperons and surviving in the cell, which explains the need of the cell for multiple Hsp70 isoforms in a complex organism [9-13].

While the Hsp70 chaperones help with the initial folding of the newly synthesized polypeptides, the Hsp90 family is involved mostly in the final maturation of proteins. For this, the chaperone needs to recognize certain conformations in the client protein rather than the primary sequence. Like Hsp70, the activity of an Hsp90 chaperone is regulated by ATP cycles; once the ATP is bound, the chaperone achieves a “tense” conformational state functioning as a clamp to hold the substrate. Upon ATP hydrolysis and in the “relaxed” state, the clamp is open and the substrate is released [14, 15].

Hsp90 is found to be involved in the UPP and protein degradation as well [5]. This finding was further supported by the discovery of CHIP (Carboxyl terminus of Hsp70-Interacting Protein) as a chaperone-dependent E3 ubiquitin ligase. There are many other proteins that bind to Hsp70 and

Hsp90 and participate in their functions. These co-chaperones can act by presenting a client protein to the chaperones as well as regulating the ATP cycles of binding and release of the chaperone in favor of protein folding or disassembly. To interact with chaperones, most of the co-chaperones adopt one of the two main classes of domains: the "J" domain found in Hsp70 co-chaperones like Hsp40, and the Tetratrico Peptide Repeat (TPR) domain found in co-chaperones interacting with Hsp70 or Hsp90 such as HIP, HOP and CHIP. TPR domain co-chaperones can take part in various biological processes due to the additional domains in their structure that catalyze different reactions such as ubiquitin ligation which is observed in CHIP [16, 17]. These co-chaperones are explained in more detail in Section 1.4.

1.3 Ubiquitin-proteasome pathway mediates protein degradation

As an ultimate mechanism to ensure quality control of proteins, the cell degradative machinery is designed to target damaged proteins for destruction after being labeled by ubiquitin in an enzymatic conjugating cascade. This process can be part of the protein turnover system wherein aged proteins are continually being degraded and replaced by new ones, or it can selectively destruct those abnormal proteins that have resulted from different types of mutations and folding errors. There is a high tendency for misfolded proteins to accumulate and form toxic aggregates. Aggregation happens through association of exposed hydrophobic domains that are normally buried inside the protein structure. In fact, intracellular aggregates are formed when the production of misfolded proteins exceeds the cell capacity for protein degradation. This situation can be found in various inherited neurodegenerative diseases in which inclusions lead to cell death (for more information on protein aggregation diseases, see Section 1.6). Therefore, studying the cellular proteolytic pathways is crucial for understanding the pathogenesis of such diseases as well as developing new therapies [18].

1.3.1 Ubiquitination pathway

Ubiquitin is a 76-residue polypeptide which covalently binds to the intracellular proteins in the process called ubiquitination (ubiquitylation) and targets them to different fates. During ubiquitination, the C terminus of ubiquitin (G76) binds to the Ɛ-amino group of a substrate lysine residue by means of the sequential actions of three enzymes: an activating enzyme (E1) uses ATP to form a thiol ester with the carboxyl group of G76 and activate ubiquitin. The activated

ubiquitin is then transiently carried by a conjugating enzyme named Ubc13 (E2) and finally delivered to a ligase (E3) that transfers the ubiquitin to the lysine residue of the substrate it is bound to. In such a manner, the ubiquitins are added progressively to the substrate until the multiubiquitin chain becomes long enough to be recognized by the proteasome for degradation (Figure 1.3.1.1) [19].

Studying the ubiquitin fusion degradation (UFD) pathway of ß-galactosidase in yeast, Koegl *et al.* discovered a new type of ubiquitination enzymes, called E4, which is required for the assembly of the polyubiquitin chain on the substrate along with the E1, E2, and E3 enzymes [20]. Further studies identified E4's affinity only for those substrates that are modified by ubiquitin. Therefore, the E4 activity was proposed to be a specialized type of E3 enzyme activity in which the oligoubiquitinated artificial fusion proteins serve as substrates [21]. Today, E4 enzymes are known to be required for efficient multiubiquitination and proteosomal targeting of the substrates through affecting the linkage between individual ubiquitin molecules in cells being exposed to stress conditions [22].

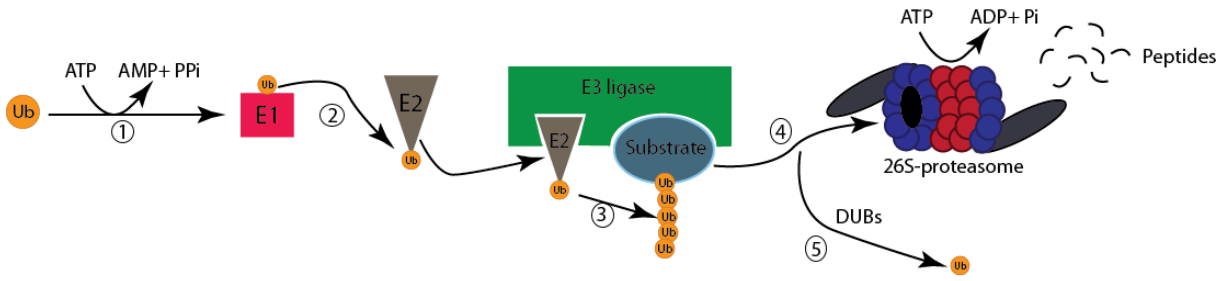


Figure 1.3.1.1 The ubiquitination pathway and its associated enzymes. The E1 enzyme binds to ubiquitin in an ATP-dependent manner (1). The activated ubiquitin is then carried by E2 enzyme and delivered to the E3-substrate complex (2). The E3 ligase transfers ubiquitin to the protein substrate either directly or indirectly (3). This process continues repeatedly until the resulted ubiquitin chain is long enough to be recognized by the 26S proteasome machinery and the protein is degraded (4). Meantime, ubiquitins are released by deubiquitination enzymes (DUBs) for further re-use (5). P_i, inorganic phosphate, pp_i, pyrophosphate; Ub, ubiquitin. The figure is adapted from [23].

Studies focusing on the molecular structure of E3 ligase have identified three families of E3 enzymes. These families act differently while transferring ubiquitin to the substrate: *(a)* a large family of E3 ligases that share the HECT (Homologous to the E6-AP Carboxyl Terminus) domain in their structure. HECT is a ~350-aminoacid region containing a highly conserved

cysteine residue which binds to ubiquitin and forms a thiol ester intermediate, therefore these ligases catalyze ubiquitination of the substrate directly. *(b)* Other E3s share the RING (Really Interesting New Gene) finger domain which is a ~70-residue region with a globular conformation resulting from the two zinc ions in its structure. Enzymes with RING finger domain do not interact with ubiquitin themselves. Instead, they facilitate ubiquitination by bringing the target protein close to the E2 enzyme. *(c)* The last family of E3s shares a modified form of RING domain, termed U-box domain, which functions like the RING domain yet lacks the hallmark metal-chelating residue in its structure. The U-box domain has a conformation similar to that of a RING finger in which the stabilization is achieved by intermolecular forces (electrostatic interactions) rather than metal ion coordination (Figure 1.3.1.2) [19, 21].

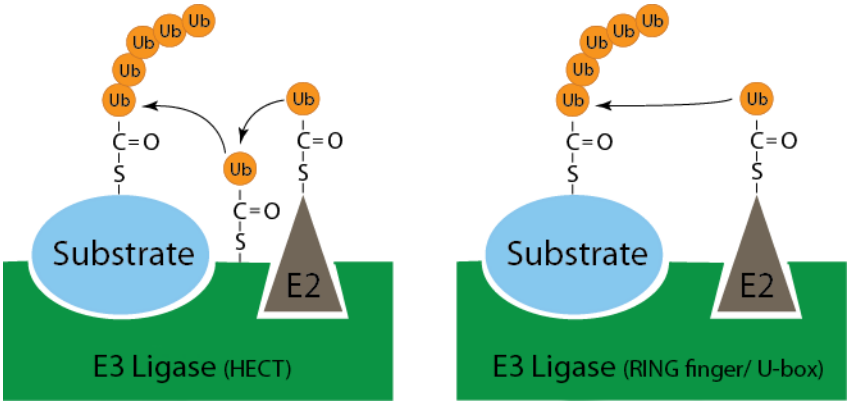


Figure 1.3.1.2 Three classes of E3 ubiquitin ligases act differently while transferring ubiquitin to the protein substrate. E3 enzymes sharing the HECT domain (left) bind to ubiquitin and participate in the delivery process via forming a thiol-ester intermediate. However, the RING finger and U-box domain-containing enzymes (right) do not interact with ubiquitin directly. Instead, they serve as a scaffold, bringing E2 and the protein substrate close together for interaction. These domains differ structurally in the presence of Zn^{2+} coordinating sites which are replaced with stabilizing hydrogen bonds and salt bridges in the U-box domain. Ub, Ubiquitin.

A ubiquitin-labeled protein is destined to different fates based on its subcellular localization, and the number and topology of the conjugated ubiquitins. For example, polyubiquitin chains conjugated by K48-G76 isopeptide bonds often target the substrate to the proteasome, opposite to those linked through K63-G76 bonds that are involved in non-proteolytic signaling. In further contrast, the substrates marked by one or a few ubiquitins are subjected to endocytosis and lysosomal proteolysis [19].

1.3.2 The 26S proteasome

The 26S proteasome is a 2500 kDa multiprotein complex with ~31 different subunits designed for protein degradation in eukaryotic cells. Studies using proteasome inhibitors have reported a huge number of cellular proteins (80-90%) to be degraded by the proteasome. These proteins are first recognized by their covalently attached ubiquitin labels in the 19S regulatory complexes located at both ending sites of the proteasome. Thereby, small conformational changes caused by the binding of a polyubiquitinated substrate will be used for its unfolding and translocation into the core complex known as the 20S proteasome. Proteolytic degradation happens once the ubiquitin molecules are released and the protein passes through the barrel-shaped 20S particle. Finally, small oligopeptides emerge from the proteasome and these are broken down into single, free amino acids by a number of ATP-dependent proteases [24].

From a structural point of view, the 20S proteasome of eukaryotes consists of 28 α- and ß- type subunits composing four stacked rings surrounding a central cavity. The two innermost ß-rings form the central proteolytic chambers with proteolytic sites: two sites that cleave after hydrophobic residues, two after acidic residues and two after basic residues; thus cutting most types of peptide bond. Substrates can enter the 20S complex passing through the narrow channel enclosed by the two outer α-rings. Before entering the channel, proteins need to be fully linearized in an ATP-dependent mechanism provided by the 19S proteasome. For this purpose there are 6 ATPase rings located in the base part of the 19S regulatory complex. The lid complex forms a distal mass on the outermost parts of the proteasome and contains 8 subunits necessary for degradation of the ubiquitinated, target protein [25].

1.4 CHIP: a link between the chaperone and proteasome system

As a co-chaperone with "anti-chaperone" functions, CHIP can interact with Hsp70 and Hsp90 and inhibit their folding activity through interfering with the ATPase cycles of Hsp70 or prevention of Hsp90's interaction with other co-chaperones [26, 27]. CHIP was identified as a TPR domain-containing protein during a human cDNA library screening carried out in 1998 looking for all the proteins with a TPR domain that might be involved in stress regulation. Interestingly, structural analysis showed similarities between the C-terminus of this protein (the U-box) and the E3 ligase component of the ubiquitination proteasome pathway, suggesting an active role in the ubiquitination of the chaperone substrates for CHIP. In this way, CHIP was

discovered as the first ubiquitin ligase that directly associates with molecular chaperones [26-28]. As mentioned earlier, those non-native proteins unable to be refolded by chaperones are directed to the proteasome for degradation. This action is mediated by the interaction of CHIP with the chaperone-substrate complex through which the substrate is prevented from release and labeled for degradation by the ubiquitin ligase activity of CHIP. Although the ubiquitination activity of CHIP via its U-box domain has been confirmed by several studies, the mechanism by which the non-native protein clients are selected for this process is not fully understood. Some studies suggest the presence of an intrinsic polypeptide binding activity that enables CHIP to recognize and selectively bind non-native protein structures in both chaperone-dependent and –independent manners. However, the mechanistic details of how such a selective interaction between CHIP and Hsp70 influences substrate binding and release cycles require further investigation [29].

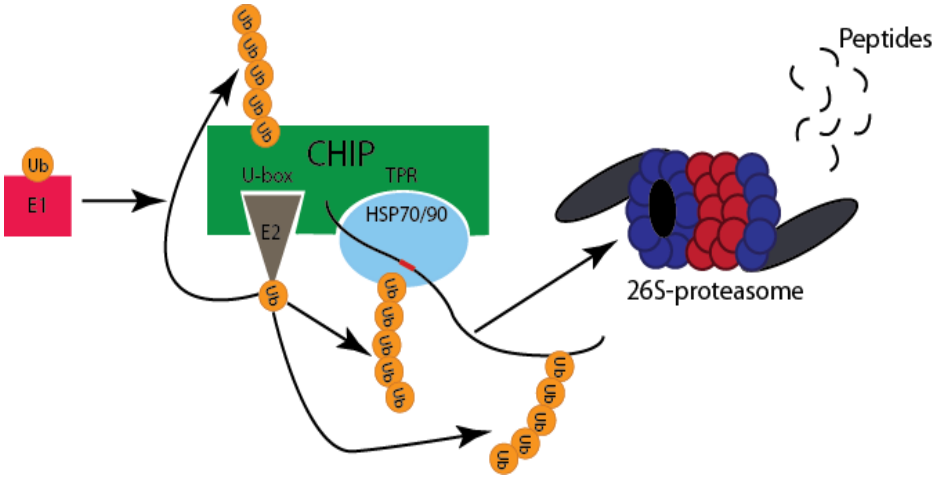


Figure 1.4.1. **The ubiquitination activity of CHIP**. CHIP mediates ubiquitination of misfolded proteins via its U-box domain which binds to E2 ubiquitin ligase and facilitates ubiquitin transfer to the substrate. The TPR domain is used to interact with Hsp70, bridging CHIP to the targeted misfolded protein. CHIP and the Hsp70 molecular chaperone are subjected to ubiquitination as well. Several hypotheses have been put forward to explain this phenomenon. Ub, ubiquitin. Figure adapted from [30].

The ubiquitination activity of CHIP is not limited to the protein substrate. The molecular chaperone to which the substrate is bound and the CHIP itself seem to be subjected to ubiquitination as well (Figure 1.4.1). Different theories have been proposed to explain this phenomenon. It can be considered as an auto-regulation strategy by which the ubiquitinated Hsp70/CHIP molecules are targeted to the proteasome during stress recovery when the non-native protein substrates are depleted [31]. On the other hand, CHIP-mediated ubiquitination of

Hsp70 can be studied regarding the fact that the ubiquitin chains observed on the Hsp70 molecules are both short and atypical (other than Lys 48 linked), suggesting that these chains can act as a targeting sequence during the signaling events rather than recognition signals for the proteasome. In this way, by ubiquitination of Hsp70, CHIP directs the chaperone towards a particular signaling pathway such as the endosome-to-lysosome trafficking [32, 33]. In other studies, CHIP self-ubiquitination is suggested to regulate its ligase activity and the substrate recruitment [34].

In addition to CHIP, the Hsp70 co-factor BAG-1 is also known to regulate chaperone's functions negatively by acting as a link between the molecular chaperones and ubiquitin/proteasome system. Once BAG-1 binds to the Hsp70-protein complex, its N-terminal ubiquitin-like domain is recognized by the proteasome components at the same time, ending up in the facilitated transfer of the complex to the proteasome. Besides, the ubiquitinated substrates are promoted to become separated from Hsp70 in the presence of BAG-1. Therefore, BAG-1 acts in concert with CHIP to target the non-native chaperone substrates for ubiquitination-mediated degradation [26, 27, 35]. On the other hand, several co-chaperones including HIP, HOP, and Hsp40 assist the folding activity of Hsp70 and Hsp90 in different ways. While HIP and Hsp40 increase Hsp70's affinity for substrates by binding to its ATPase domain and stabilizing the ADP-bound conformation, HOP promotes cooperation between Hsp70 and Hsp90 in a way that the substrate is efficiently transferred between them [17]. Maintenance of cellular homeostasis requires a balance between protein folding and degradation pathways. This balance is controlled by the folding kinetics of the protein substrates and the activity network of different co-chaperones that enable them to associate with the molecular chaperones and assist them in either folding or degrading the substrate. The two co-factors CHIP and HOP appear to compete on a common TPR acceptor site on Hsp70 and Hsp90. Similarly, HIP and BAG-1 bind to the molecular chaperones via their C-terminal competitively. Under normal conditions, the lower intracellular concentrations of CHIP and BAG-1, in comparison to the folding-inducing co-chaperones, direct the cell condition in favor of protein folding, whereas the over-expression of CHIP during stress exposure shifts the balance toward protein degradation [26, 27].

1.5 CHIP: structures and interactions

CHIP is an evolutionary conserved ~35 kDa protein which is highly expressed as a dimeric co-chaperone in tissues that are active in terms of metabolism and protein turnover, such as brain, heart, and skeletal muscle. The significant amino acid sequence similarity existing between human and mouse CHIP (98%), makes the latter animal a powerful model for in-depth characterization of the structural properties of CHIP and its interacting partners [35]. There are two main domains involved in the primary structure of CHIP: a C-terminal U-box domain which facilitates ubiquitination of the chaperone substrates through the interaction with different E2 enzymes, and an N-terminal TPR domain which mediates the interaction of CHIP with Hsp70 and Hsp90 molecular chaperones. These domains are separated by a central Helical Hairpin (HH) region (also termed the central Coiled-Coil (CC) domain) which itself plays an important role in the dimerization and stability of the whole protein [36]. In order to become dimerized, the U-box domain and the distal segment of the helical hairpin of the two separated CHIP monomers interact symmetrically through their strongly conserved amino acid chains, providing the two cores of the dimer interface (Figure 1.5.1 A) [37].

Three pairs of TPR motifs together with an elongated seventh helix form a super helical domain with distinct ligand binding grooves at the N-terminal of the CHIP protein. Each TPR motif is ~34 residues in length and consists of two antiparallel α-helices connected by a short turn. At the end of the TPR domain, where it meets the central dimerization region, there is an extended α-helix packing against the third TPR motif with two different positions. This causes the TPR domain to adopt different conformations in each protomer of the homodimeric CHIP (Figure 1.5.1 B). Two characteristic carboxylate clamps are found within the grooves of the TPR domain to interact with the C-terminal EEVD motifs of Hsp70 and Hsp90 selectively. Different amino acid residues located upstream of the EEVD motifs in Hsp70 and Hsp90 chaperones explain the specificity of their interaction with CHIP (Figure 1.5.1 C) [38-40].

At the C-terminal, CHIP is composed of a pair of ß-hairpins continued by two α-helices with another hairpin in between. In contrast to the TPR domain, the structure of the U-box domain is similar in both protomers. As an E3 ligase, each U-box domain of CHIP interacts with Ubc13 (E2) through the association of hydrophobic grooves on the U-box surface and the conserved Ser-Pro-Ala binding motifs of E2-conjugating enzymes. However, protein crystallography studies on

mouse CHIP have reported an asymmetry in this interaction via the detection of only one ubiquitin conjugating system coupled to a dimeric CHIP. In fact, as a consequence of having different conformations, the TPR domain interacts with its associated U-box and prevents the binding of E2 in one protomer of the CHIP dimer, while there is no obstacle for the interaction with E2 in the opposing protomer. Therefore, CHIP displays a "half-of-sites" activity in the formation of polyubiquitin chains on its interacting peptides [37, 41].

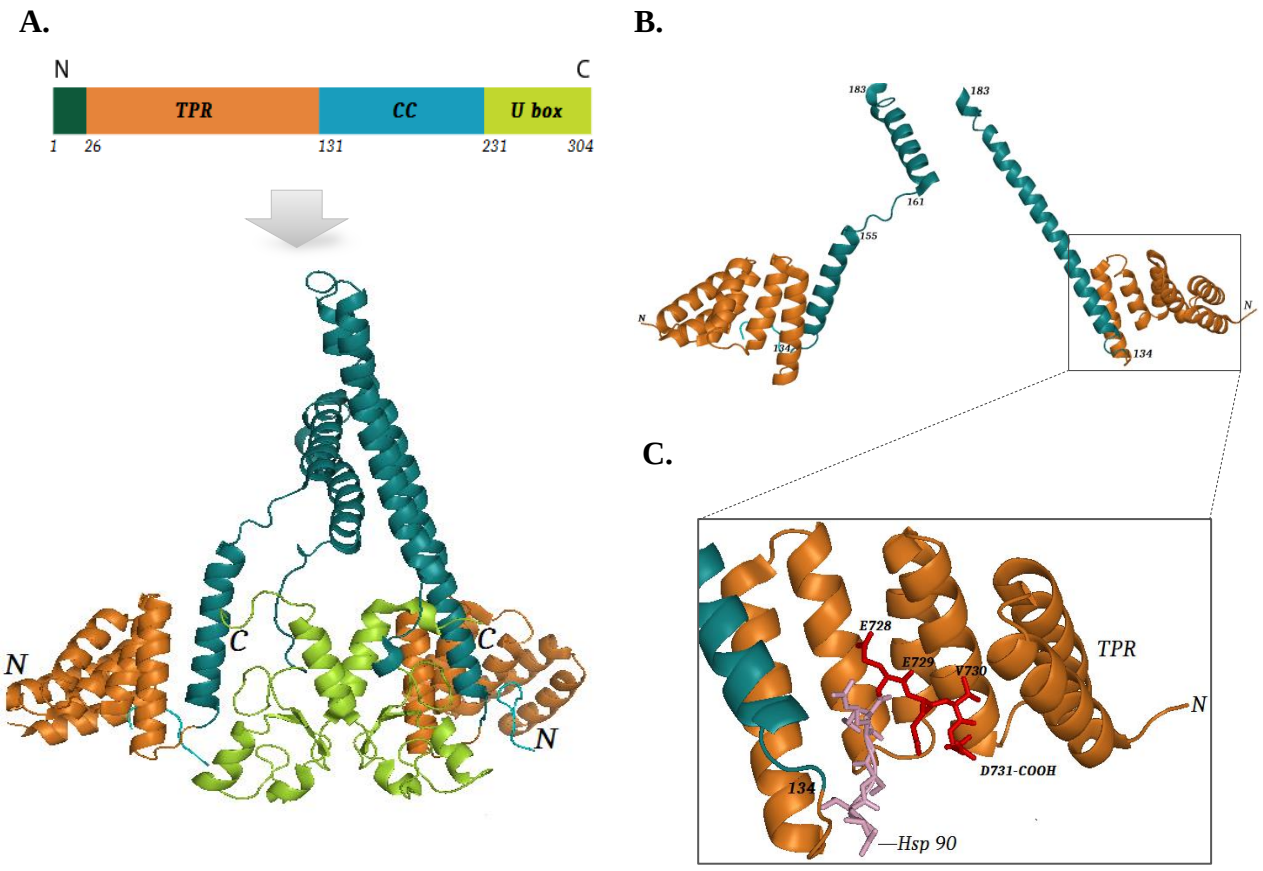


Figure 1.5.1 Architecture of the CHIP protein homodimer. A) CHIP protein domain and secondary structure cartoon of the homodimer from the N (orange) to C (green) terminus: The TPR and U-box domains are linked through the central coiled-coil (CC) region (blue). B) Asymmetric disposition of TPR domains: The two TPR domains of a dimeric CHIP are located asymmetricaly due to different structure of the extended helix 7 (blue) in the two protomers. In one (right), the α-helix is formed straight from Asp 134 to Arg 155. In the other (left), the polypeptide is broken in two separate α-helices: Asp134-Arg155 and Glu161-Arg183. C) detailed view of the Hsp90 C-terminal bound to the CHIP TPR domain: The C-terminal EEVD sequence lies along the TPR channel in both Hsp70 and Hsp90 molecular chaperones. Yet, different amino acids located at the upstream side chains interact with TPR domain specifically. All the molecular graphics were produced by using PyMOL software (https://www.pymol.org).

1.6 Protein aggregation and neurodegenerative diseases

Mutations leading to "loss of function" of a single protein are the most probable cause of recessively inherited monogenic diseases [42]. Such mutations can either reduce the activity level or the steady-state concentration of the protein inside the cell. There is evidence suggesting that missense mutations can accelerate the degradation of a protein in two possible ways. First, they can decrease the thermodynamic stability of the protein through enhancing the level of free-energy inside the native conformation and therefore increase its potency for unfolding. Second, these mutations can decrease the rate of correct folding of a protein via affecting its kinetic partitioning in a way that smaller amounts of protein enter the correct folding pathway once they are synthesized as polypeptide chains [42, 43]. The misfolded or damaged proteins are normally removed in cells via the protein degradation system. However, in some cases, when the load of unfolded proteins overreaches the amount of available chaperones and cell capacity for protein degradation, these proteins start to accumulate and form toxic aggregates inside the cell. This notion is supported by the observation of chaperones and components of ubiquitin proteasome pathway inside the aggregations. The accumulation of abnormal, damaged proteins appear as a type of disorders with gain of (abnormal) function phenotype, named "aggregation disorders" [42]. A large number of dominantly inherited disorders and many neurodegenerative diseases can develop in this way. The presence of protein aggregates as intra-neuronal inclusions is a common morphological feature and the hallmark of all neurodegenerative diseases. These lesions were first detected in neurofibrillary tangles and further recognized in the cytoplasm of neurons in patients hosting diseases such as Alzheimer's, Parkinson, and Huntington [44].

Further studies in this area have provided insight into the pathophysiological importance of aggregated proteins: once the protein is unfolded, exposed hydrophobic structures on its surface interact with each other in order to minimize the contact area to the hydrophilic environment. This, together with other non-covalent and cross-linking interactions, form a protein aggregate which increases in size by addition of more proteins to the surface over time. The resulting aggregates trigger a set of intracellular reactions ending up in toxicity and cell death [45].

Surprisingly, there has been an increasing number of studies over the past few years reporting little correlation between the level of protein aggregation and toxicity in cells associated with neurodegenerative diseases. According to these studies, the toxicity is related to those soluble

misfolded protein intermediates which are produced on the pathway to inclusions and can be detected by proteasomes for degradation. Yet, due to the high amounts of ß-sheets in their structure, their full entry to the proteasome is resisted, resulting in steric occlusion of the proteasome. In this way, protein aggregation is considered as a new mechanism of protein quality control system through which unfolded proteins are directed to specific cellular compartments for subsequent removal [46-48].

Different factors are known to be responsible for abnormal protein folding, such as mutations, defects in protein biogenesis, environmental stresses including the presence of reactive oxygen species (ROS) inside the cell, and aging, which is characterized by reduced capacity of the protein quality control system for eliminating misfolded proteins. There are two sets of mutations resulting in protein unfolding and subsequent aggregation. Some mutations are believed to modify protein structure, and increase its tendency towards misfolding. Expanded CAG trinucleotide repeats in Huntington disease (HD), and mutations responsible for the development of various conformational diseases such as type 2 diabetes are examples of this type. Other mutations are found in genes encoding components of the protein quality control system. For instance, mutations in *PARKIN* (encoding an E3 ligase) are known to be strongly associated with Parkinson disease (PD) [48]. CHIP, as another E3 ligase enzyme in the protein ubiquitination pathway is subjected to various mutations in the Cerebellar Ataxia Diseases [49]. Cerebellar Ataxias (CA) consist of a wide range of inherited and sporadic neurodegenerative disorders in which ataxia, or the loss of coordination, is considered as the main phenotype. A general classification of the common ataxia subtypes is given in Figure 1.6.1.

Patients with CA may suffer from balance and gait impairments as well as speech, vision, and cognitive disabilities [49, 50]. Repeat expansions in either coding or non-coding parts of the genome are responsible for the majority of inherited ataxias. For example, a gene called *FXN* contains a tri-nucleotide sequence (GAA) that is normally repeated between 7 and 22 times. However, when defective, the repeat can expand to hundreds. Alleles containing GAA expansions of 90-1700 repeats are associated with the development of Friedreich Ataxia (FA) - the most common type of Autosomal Recessive Cerebellar Ataxia (ARCA) [51]. Other rare ARCA forms are caused as a result of conventional mutations in related genes [51-53].

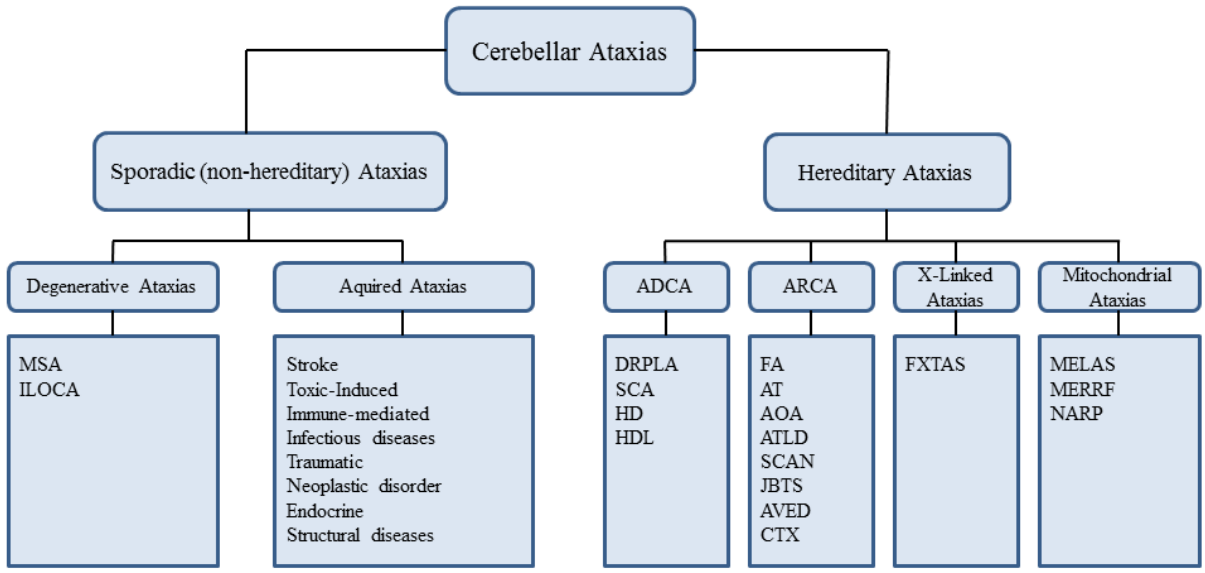


Figure 1.6.1 General classification of Cerebellar Ataxia diseases. MSA, multiple system atrophy/ OMIM #146500; ILOCA, idiopathic late-onset cerebellar ataxia; DRPLA, dentatorubral pallidoluysian atrophy/ OMIM #125370; SCA, spinocerebellar ataxia/ OMIM #164400; HD, Huntington disease/ OMIM #143100; HDL, Huntington disease-like/ OMIM #603218; FA, Friedreich ataxia/ OMIM #229300; AT, ataxia telangiectasia/ OMIM #208900; AOA, ataxia with oculomotor apraxia/ OMIM #208920; ATLD, ataxia-telangiectasia-like disorder/ OMIM #604391; SCAN, spinocerebellar ataxia with axonal neuropathy/ OMIM #607250; JBTS, Joubert syndrome/ OMIM #614464 ; AVED, ataxia with isolated vitamin E deficiency/ OMIM #277460; CTX, Cerebrotendinous xanthomatosis/ OMIM #213700; FXTAS, fragile X-associated tremor/ataxia syndrome/ OMIM #300623; MELAS, mitochondrial encephalopathy, lactic acidosis and stroke-like episodes/ OMIM #540000; MERRF, myonucleus epilepsy with ragged-red fibres/ OMIM #545000; NARP, neuropathy, ataxia and retinitis pignemtosa/ OMIM #551500.

The focus of this thesis concerns one of these rare ARCA diseases caused by recessive mutations in *STUB1* encoding the CHIP ligase protein.

1.7 Autosomal Recessive Cerebellar Ataxia (ARCA): the current state of affairs

The recessively inherited group of cerebellar ataxias is characterized by an early onset and gradual worsening of gait and balance with the development of hypotonia or excessive clumsiness over months and years. Friedrich Ataxia (FA) with the estimated prevalence of 2-4/100,000 is known to be the most frequent type of ARCA. This is followed by Ataxia Telangiectasia (AT) (1-2/100,000) and early onset cerebellar ataxia with retained tendon reflexes (1/100,000) [53, 54]. Many new ARCA genes have been identified recently, owing to the

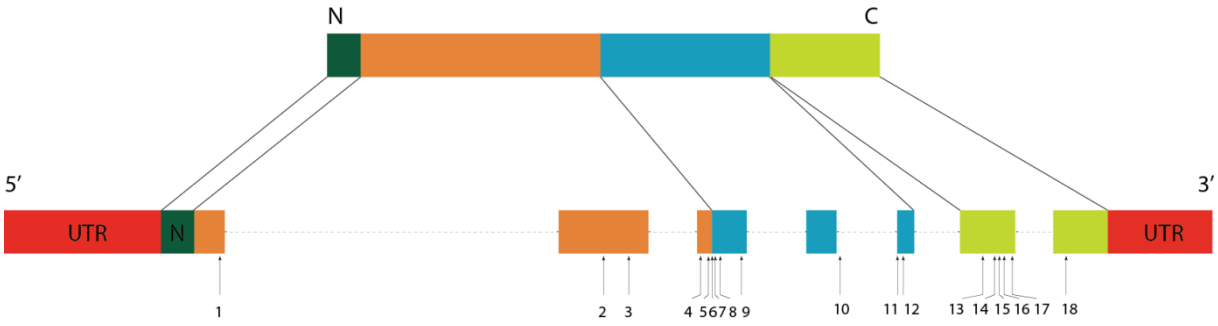
development of the whole exome sequencing (WES) approach. WES is a powerful tool for studying those rare Mendelian disorders for which the causative mutations could not be identified by conventional methods such as linkage mapping and candidate gene sequencing. Exome sequencing allows for a quick, cost-efficient sequencing of the entire protein-coding region of the genome through parallel sequencing of an enriched, amplified library of short DNA fragments containing approximately 3-10 billion base pairs (Gbp) of data. Candidate causal variants called by mapping and alignment of the sequencing data will be filtered subsequently against a set of polymorphisms available in public databases, and a minimum number of high priority variants will be detected for each individual, among which, mutations leading to dominant and recessive disorders are expected to be seen [55, 56].

Table 1.7.1 Classification of ARCA based on molecular pathogenesis. IOSCA, infantile onset spinocerebellar ataxia/ OMIM# 271245; SANDO, sensory ataxic neuropathy, dysarthria and ophthalmoparesis/ OMIM #607459; MIRAS, mitochondrial recessive ataxia syndrome; ARSACS, autosomal recessive spastic cerebellar ataxia of Charlevoix-Saguenay/ OMIM #270550; MSS, Marinesco-Sjogren syndrome/ OMIM #248800.

Disorder	Gene	Protein
Mitochondrial dysfunction		
▪ FA	*FXN*	Frataxin
▪ AVED	*TTPA*	α- tocopherol transfer protein
▪ I0SCA	*C10orf2*	Twinkle
▪ SAND0 or MIRAS	*POLG1*	DNA polymerase subunit γ-1
▪ ARCA2	*CABC1*	Chaperone-activity of bc1 complex-like
DNA repair dysfunction		
▪ AT	*ATM*	Serine protein kinase
▪ ATLD	*MRE11*	Meiotoc recombination 11
▪ AOA1 or EAOH	*APTX*	Aprataxin
▪ AOA2	*SETX*	Senataxin
▪ SCAN1	*TDP1*	Tyrosyl-phosphodiesterase-1
Protein misfolding and chaperone dysfunction		
▪ ARSACS	*SACS*	Sacsin
▪ MSS	*SIL1*	Nucleotide exchange factore SIL1
Mis-localization of synaptic myonuclei		
▪ ARCA1	*SYNE1*	Nesprin-1
Altered function of calcium mediated chloride channel		
▪ ARCA3	*ANO10*	Anoctamin-10
Altered vesicular trafficking		
▪ Salih ataxia	*KIAA0226*	Rundataxin

Different types of ARCAs can be distinguished based on the underlying disease mechanism: (1) mitochondrial dysfunction, (2) DNA repair deficiency, (3) mis-localization of synaptic myonuclei, (4) altered function of calcium-mediated chloride channels, (5) altered vascular trafficking, and (6) protein misfolding and chaperone dysfunctions. Common diseases and causative mutations related to each pathway are listed in Table 1.7.1.

Some of these mutations are considered to be very rare, being restricted to isolated populations or single families, like mutations in *C10orf2*, *TDP1*, and *KIAA0226* [51]. In addition, recessive mutations in *STUB1* (encoding CHIP) have been discovered lately to cause another subtype of ARCA named autosomal recessive spinocerebellar ataxia-16 (SCAR16) in a limited number of families. These mutations affect different domains of CHIP, mostly through the substitution of a single amino acid in the protein sequence (missense mutations) (Figure 1.7.1) [30, 49, 57-63].



TPR

# Mutation	Ref.
1 c.82G>A/p.E28K	30
2 c.194A>G/p.N65S	30
3 c.235G>A; 236C>A/p.A79T,A79D	63
4 c.367C>G/p.L123V	63

CC

# Mutation	Ref.
5 c.389A>T/p.N130I	62
6 c.430A>T/p.K144*	30
7 c.433A>C/K145Q	60
8 c.441G>T/p.W147C	62
9 c.493C>T/p.L165F	62
10 c.612+1G>C	59
11 c.621C>G/p.Y207*	62
12 c.633G>A/p.M211I	58

U-Box

# Mutation	Ref.
13 c.687-690 del /p.Y230Cfs*8	60
14 c.707G>C/p.S236T	62
15 c.712G>T/E238*	58
16 c.719T>C/p.M240T	63
17 c.737C>T/p.T246M	61
18 c.823C>G/p.L275V	59

Figure1.7.1 *STUB1* mutations and corresponding CHIP protein domains. The locations of all the mutations associated with SCAR16 are specified. The inserted tables indicate the respective nucleic acid and amino acid change for each mutation. This figure is adapted from [49]. *STUB1* reference sequence (RefSeq) = NM_005861.3.

2. Aims of the study

This project is based on identification of three mutations in *STUB1* in two families with SCAR 16 using a combination of homozygosity mapping and exome sequencing by our group. These mutations include p.E28K, p.N656S, and p.K144* [30]. Recent studies have characterized the molecular properties of these mutations to some extent, yet more investigations are required to achieve a clear picture of changes in CHIP structure and function under the effect of different mutations. Thus, the two mutations E28K and N65S together with four additional mutations on CHIP were selected for further analysis. The location of each mutation is illustrated in Figure 2.1.

Mutations	Reference
TPR domain	
▪ c.82G>A/ p.E28K	54
▪ c. 194A>G/ p.N65S	54
Coiled-coil domain	
▪ c.433A>C/ p.K145Q	53
▪ c.633G>A/ p.M211I	51
U-box domain	
▪ c.707G>C/ p.S236T	56
▪ c.737C>T/ p.T246M	55

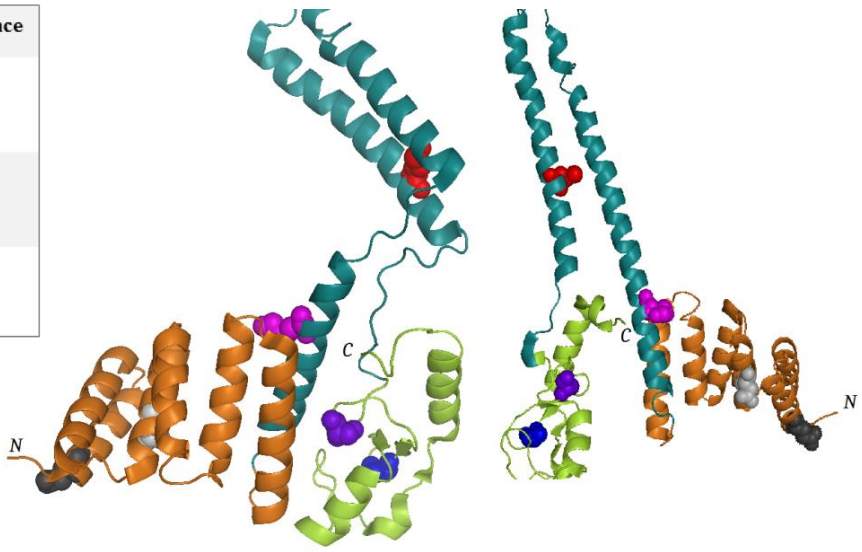


Figure 2.1 The position of each selected mutation is shown in the 3D structure of the dimeric CHIP with colorful spheres. Mutations E28K (dark gray) and N65S (light gray) are located at the TPR domain (orange). Mutations K145Q (pink) and M211I (red) are placed at the central domain (light blue) and the two last mutations S236T (purple) and T246M (dark blue) are positioned in the U-box domain (green) of the protein. This figure was created by using PyMOL software (https://www.pymol.org).

The overall aim of this study was to characterize the molecular properties of mutations in *STUB1* -a newly identified gene responsible for SCAR 16, following the specific objectives:

1) To study *in vitro* ubiquitination activity of the wild type and mutant encoded CHIP proteins.
2) To check the effect of mutations on CHIP protein susceptibility to limited proteolysis, *in vitro.*
3) To identify different conformational states of mutant CHIP proteins and compare them with the wild type, *in vitro.*

4) To study the secondary structure and thermal stability of CHIP proteins for both wild type and mutant forms, *in vitro.*
5) To investigate if the expression level of exogenous CHIP is affected by the mutations, *in cellulo.*

3. Materials and Methods

3.1 Site-directed mutagenesis

The QuickChange® XL Site-directed Mutagenesis Kit (Agilent Technologies, Santa Clara, California, USA) was used to introduce three specific mutations including p.K145Q (c.433A>C), p.M211I (c.633G>A), and p.S236T (c.707G>C) into the WT *STUB1* cDNA (reference sequence: NM_005861.3, [64]) in the mammalian expression vector pcDNA3.1-V5/HisB and bacterial expression vector pETM-41-MBP/His [30]. Forward and reverse primers (provided by Sigma-Aldrich, Missouri, USA) are listed in Table 3.1.1. Mutant plasmids for three mutations including p.E28K (c. 82C>G), p.N65S (c.194 C>A), and p.T246M (c.737C>T) were previously constructed [30] and, therefore, were not included here.

Table 3.1.1 Mutagenesis primers

Primer	Sequence
K145Q Forward	5´-GCT-GTT-CCA-GCG-CTG-CTT-CTT-CGC-GAT-TCG-3´
K145Q Reverse	5´-CGA-ATC-GCG-AAG-AAG-CAG-CGC-TGG-AAC-AGC-3´
M211I Forward	5´AAA-AAG-CTC-GTC-TAT-GTC-CGC-CAT-GTA-CTT-GTC-GT-3´
M211I Reverse	5´-ACGACA-AGT-ACA-TGG-CGG-ACA-TAG-ACG-AGC-TTT-TT-3´
S236T Forward	5´-CCG-CAT-CAG-CTC-AAA-GGT-GAT-CTT-GCC-ACA-CAG-3´
S236T Reverse	5´-CTG-TGT-GGC-AAG-ATC-ACC-TTT-GAG-CTG-ATG-CGG-3

Each tube for a mutagenesis reaction was prepared by adding 1 µl of PfuUltra High Fidelity DNA polymerase to 50 µl master mix containing 5 µl 10x reaction buffer, 10 ng plasmid DNA, 125 ng forward and reverse primers, 1 µl dNTP mix, 3 µl QuickSolution™ Reagent, and MilliQ water. The mutagenesis was carried out by running Polymerase Chain Reaction (PCR) (Gene Amp® PCR Systems 9700, Applied Biosystems, Thermo Fisher Scientific, Waltham, Massachusetts, USA) for the control and sample reactions using the program:

Pre-denaturation	95 °C	1 min	
Denaturation	95 °C	50 sec	
Annealing	60 °C	50 sec	x18 cycles
Extension	68 °C	6 min	
Final Extension	68 °C	7 min	
End	4 °C	∞	

Following the PCR, the tubes were cooled down to ≤37 °C for 2 min on ice, and the parent DNA plasmids were digested by adding 1 µl of diphosphopyridine nucleotidase *Dpn* I restriction enzyme directly to the tubes and incubating at 37 °C for one hour (BINDER Laboratory Incubator, Tuttlingen, Germany). The mutant DNA plasmids were further transformed into the bacterial competent cells for sequencing and analysis as described in the sections below.

3.2 Transformation of XL10-Gold competent cells

The XL10-Gold ultracompetent cells were gently thawed on ice and prepared in 45 µl aliquots for transformation with the control and mutated vectors. Two µl of XL10-Gold β-mercaptoethanol mix (β-ME) was first added to the cells, swirling every 2 min during a 10 min pre-incubation on ice. The cells were then mixed with 2 µl of DpnI-treated DNA plasmid and incubated for 30 min on ice. Transformation was completed by subjecting the cells to a 30 sec heat-pulse in a water bath at 42 °C, and followed by a further 2 min incubation on ice. A volume of 500 µl pre-heated S.O.C Medium (Sigma-Aldrich) was added to the transformed cells, and they were grown at 37 °C for one hour with agitation (Innova 4300 Incubator Shaker, Thermo Fisher Scientific). After incubation, 250 µl of the cells were spread on room-tempered ImMedia^{TM} KanAgar (for pETM-41 plasmid transfected cells) and ImMedia^{TM} AmpAgar (for pcDNA3.1-V5/HisB plasmid transfected cells) LB-agar plates (Invitrogen, Thermo Fisher Scientific). Colonies were analyzed and selected after an overnight incubation at 37 °C.

3.3 Colony PCR and sequencing

A few colonies were selected and screened by sequencing to confirm that they contained plasmids with the desired mutations. Each colony was suspended in 10 µl of MilliQ water of which 1 µl was used for the PCR reaction. The PCR mixture was prepared in a final volume of 20 µl by adding 1 µl 360 GC Enhancer (Thermo Fisher Scientific), 1 µl of each forward and reverse primer (Sigma-Aldrich) and 1 µl of the colony suspension to 10 µl of AmpliTaq Gold® 360 Master Mix (Thermo Fisher Scientific). A negative control tube was prepared by adding 1 µl of MilliQ water in substitution for the bacterial suspension to the reaction mixtures. The PCR was conducted with the following thermocycling conditions:

Pre-denaturation	94 °C	10 min	
Denaturation	94 °C	30 sec	
Annealing	58 °C	30 sec	x35 cycles
Extension	72 °C	30 sec	
Final Extension	72 °C	7 min	
End	4 °C	∞	

For all the colonies and control reactions, T7 Promoter Primer was used as the forward primer. In addition, BGH and T7 Terminator primers were used as the reverse primer for pcDNA3.1-V5/His and pETM-41 mutant plasmids, respectively. For primer sequences, see Table 3.3.1.

After the PCR, 5 µl of each reaction was mixed with 1 µl of DNA Gel loading Dye (6x) (Thermo Fisher Scientific) and analyzed for the presence of the amplified DNA fragment by agarose gel electrophoresis using 1 kbp DNA Ladder (New England BioLabs, Massachusetts, USA) as described in Section 3.15.1.

Colonies with correct DNA size were subjected to DNA sequencing. In this step, the additional dNTPs and primers in the reaction tubes were first removed by adding 2 µl illustra™ ExoProStar™ (Sigma-Aldrich) to 5 µl of PCR product in new sets of tubes using the PCR-program:

1	94 °C	15 min
2	80 °C	15 min
End	4 °C	∞

This was continued with another PCR process in which 1 µl of the previous PCR product together with 0.5 µl of different internal *STUB1* primers covering the area of mutations (Table 3.1.1) were added to an 8.5 µl mixture containing 2 µl of 5x Sequencing Buffer (Applied Biosystems), 1 µl of Big Dye Terminator v1.1 cycle sequencing buffer (Applied Biosystems) and MilliQ water. The following PCR-program was used in this step:

Pre-denaturation	96 °C	1 min	
Denaturation	96 °C	10 sec	
Annealing	58 °C	5 sec	x25 cycles
Extension	60 °C	4 min	
End	4 °C	∞	

Sequencing was performed by the DNA sequencing facility core (Center for Medical Genetics and Molecular Medicine, Haukeland University Hospital) using an Applied Biosystems 3730 capillary sequencer (Applied Biosystems) and the results were analyzed by Finch TV v.1.4.0 software (Geospiza Inc., Seattle, Washington, USA).

Table 3.3.1 PCR and cDNA sequencing primers

Primer	Sequence
STUB1-1F	5´-ATG-AAG-GGC-AAG-GAG-GAG-3´
STUB1-1R	5´-CTC-CTC-CTT-GCC-CTT-CAT-3´
STUB1-2F	5´-GGA-GAT-GGA-GAG-CTA-TGA-TGA-G-3´
STUB1-2R	5´-CTC-ATC-ATA-GCT-CTC-CAT-CTC-C-3´
STUB1-3F	5´-TTT-CTC-AGG-TGG-ATG-AGA-AGA-G-3´
STUB1-4F	5´-TGG-CTG-GGT-GGA-GGA-CTA-CTG-A-3´
STUB1-4R	5´-TCA-GTA-GTC-CTC-CAC-CCA-GCC-A-3´
T7 promoter	5´-TAA-TAC-GAC-TCA-CTA-TAG-GG-3´
T7 terminator	5´-GCT-AGT-TAT-TGC-TCA-GCG-G-3´
BGH	5´-TAG-AAG-GCA-CAG-TCG-AGG-3´

3.4 Plasmid purification

Those bacterial colonies in which the desired mutations were verified by sequencing were selected and grown for plasmid purification using QIAfilter™ Plasmid Maxi Kit (25) (Qiagen, Hilden, Germany). The first cultures were prepared by an 8 hour incubation of the bacterial colony in 5 ml LB Medium containing one LB Broth 1.1 G Tablet (Sigma-Aldrich) per 50 ml MilliQ water and 100 µg/ml kanamycin. Five hundred ml of the same medium was then inoculated with 500 µl of the grown culture and incubated overnight at 37 °C with agitation. On the next day, bacterial cultures were centrifuged at 4600 rpm for 40 min at 4 °C (Multifuge 3S-R, Thermo Fisher Scientific), and the resulting pellets were completely resuspended in 10 ml buffer P1 containing 100 µg/ml RNAase A and 1:1000 LyseBlue reagent®. The solution was then mixed well with 10 ml buffer P2 and incubated for 5 min at room temperature. Using LyseBlue reagent, the lysate turned blue after sufficient mixing. In order to precipitate non-plasmid material, 10 ml of pre-chilled buffer P3 was added to the lysate and mixed immediately until a homogeneous, colorless suspension was appeared. The lysate was poured into the barrel of the QIAfilter Cartridge and incubated for 10 min at room temperature. Thereafter, the cap was removed from the cartridge and the cell lysate was gently filtered into the QIAGEN-tip that had previously been

equilibrated with 10 ml buffer QBT. The tip was washed twice with 30 ml buffer QC, and further eluted with 15 ml buffer QF. To precipitate DNA, 10.5 ml room-temperature Isopropanol (Arcus, Oslo, Norway), was added to the eluted samples, followed by 1 hour centrifugation at 4600 rpm at 4 °C. DNA pellets were then washed with 5 ml room tempered 70% Ethanol (Sigma-Aldrich) and centrifuged at 4600 rpm for 30 min at 4 °C. The supernatant was carefully removed and the pellets were air-dried for 15 min before dissolving in 250 µl Tris-EDTA (TE) buffer (1 M Tris-HCl, pH 8.0, 0.5 M EDTA, pH 8.0). The DNA concentration in the samples was measured using Nanodrop ND-1000^{TM} (Saveen Werner, Limhamn, Sweden), and the samples stored at 20 °C.

Finally, the entire sequence of *STUB1* cDNA was verified by DNA sequencing, using the same primers as used before (Section 3.3, Table 3.3.1). Sequencing was performed in an Applied Biosystems 3730 capillary sequencer at the DNA sequencing facility core at the Center for Medical Genetics and Molecular Medicine at Haukeland University Hospital.

For in vitro recombinant protein studies

3.5 Recombinant protein production

3.5.1 Transformation of BL21 bacterial cells with recombinant vectors

WT and mutant CHIP cDNA in His6-MBP-tagged vectors pTEM-41 were transformed into BL21-CodonPlus (DE3)-RP competent cells (Agilent Technologies). One µl plasmid was added to 50 µl BL21 cells and incubated on ice for 30 min. In order to increase transformation efficiency, 2 µl of a 1:10 dilution of the XL10-Gold β-mercaptoethanol mix (Agilent Technologies) had been previously added to the cells and gently swirled for 10 min on ice. The cells were then heat-pulsed in a water bath at 42 °C for 20 sec, and placed on ice for another 10 min. Transformed cells received 450 µl of pre-heated S.O.C medium and were incubated with shaking (225 rpm) at 37 °C for one hour. Fifty µl of the cell suspension together with 50 µl of S.O.C medium was further plated out on ImMedia^{TM} KanAgar LB-agar plates, and incubated overnight at 37 °C. Colonies were analyzed and selected the next day.

3.5.2 Recombinant protein expression in BL21 cells (*E.coli*)

A start culture was prepared by picking a single colony of transformed BL21 cells and adding it to 20 ml LB medium containing 100 µg/ml kanamycin, 0.2% filtered glucose and 50 µg/ml chloramphenicol (Sigma-Aldrich). The culture was incubated overnight at 37 °C with agitation.

Five hundred ml of the same medium was inoculated with 5 ml of the overnight culture and left to incubate at 37 °C with shaking until it had grown to 0.6-0.8 OD at 600 nm (measured in a spectrophotometer; PerkinElmer, Massachusetts, USA) which is an ideal value of cell density for initiating protein expression based on the Qiagen protocol for high-level expression and purification of 6X His-tagged proteins. After adding Isopropyl ß-D-1-thiogalactopyranoside (IPTG) (GE healthcare, little Chalfont, UK) to a final concentration of 0.5 mM, the cells were induced to express recombinant protein during an overnight incubation at 25 °C with agitation (Innova 4230 Refrigerated Incubator Shaker, Thermo Fisher Scientific). In order to check whether protein expression had been sufficiently induced before proceeding with protein purification, 1 ml culture was removed before and after IPTG addition and centrifuged at 3000 rpm for 2 min (Centrifuge 5424, Eppendorf, Hamburg, Germany). The bacterial pellet was further resuspended in 400 µl MilliQ water, and a 10 µl aliquot was analyzed by SDS-PAGE and Coomassie blue staining as described in Sections 3.15.2 and 3.16.1.

3.5.3 Recombinant protein purification

The overnight culture was centrifuged at 4000x g for 45 min at 4 °C. Bacterial pellets were dissolved in 5 ml/g wet weight lysis buffer (50 mM $NaH_2PO_4.H_2O$, 300 mM NaCl, 10 mM Imidazole, 20 mM 2-Mercaptoethanol, 0.1% Tween-20, pH 8) and sonicated 10 times (each for 20 sec with 10 sec cooling time in-between) on ice at 300-400 W power output (VirSonic 300, VirTis, Gardiner, USA). Aliquots of 2 ml harvested cells were further centrifuged at 15000x g for 40 min at 4 °C (Centrifuge 5417C, Eppendorf), and the supernatant was collected. In order to purify His-MBP-tagged proteins from the protein pool of the cells, 1 ml of Ni-NTA Agarose Nickel Resin (Qiagen, Venlo, Netherlands) was added as a 50% slurry in lysis buffer to the supernatant and left rotating for 1 hour (Labinco Beher B.V., Breda, Netherlands). The unbound material was removed by centrifugation at 1500x g for 2 min, and the protein-bound resin was loaded to the Micro Bio-Spin™ Size Exclusion Spin Columns (Bio-Rad Laboratories, Hercules, California, USA) and washed with 1 ml washing buffer (50 mM $NaH_2PO_4.H_2O$, 300 mM NaCl,

20 mM Imidazole, 20 mM 2-Mercaptoethanol, 0.1% Tween-20, pH 8) three times. The resin was centrifuged at 1500x g after each wash. Finally, 500 µl of elution buffer (50 mM $NaH_2PO_4.H_2O$, 300 mM NaCl, 250 mM Imidazole, 20 mM 2-Mercaptoethanol, 0.1% Tween-20, pH 8) was added and allowed to pass through the column by gravity to elute His-MBP-tagged proteins. The elution step was repeated three times and the protein fractions were collected in separate Eppendorf tubes each time.

3.5.4 Determination of recombinant protein concentration

The protein concentrations were measured on a Nanodrop ND-1000^{TM} using the molar extinction coefficient number of 12.29 ($mg.mL^{-1}.cm^{-1}$) for MBP-CHIP, decided according to the method of Gill and von Hippel [65]. Blank measurements were made with elution buffer, and a minimum of two parallel samples were measured for each protein sample.

3.5.5 Recombinant protein buffer exchange

Using Zeba™ Spin Desalting Columns (Thermo Fisher Scientific), the elution buffer was changed to the suitable buffer for protein storage (100 mM Hepes, pH 8, 5 mM DTT, 100 mM NaCl, 10% Glycerol). The columns were first put in collection tubes and centrifuged at 1000x g for 2 min with twisted-off bottom closures and loosened caps. The storage buffer was next added to the columns and centrifuged at 1000x g for another 2 min. This step was repeated 4 times to ensure optimal buffer exchange. Finally, the columns were placed in new collection tubes and protein samples were applied to the center of the compact resin bed. After 2 min centrifugation at 1000x g, the MBP-CHIP proteins were collected in the new storage buffer. The concentrations of the protein samples were measured once again by Nanodrop ND-1000^{TM} using the storage buffer as the blanking solution.

3.6 Separation of MBP tags by TEV protease

In order to generate CHIP, the Hisx6-MBP-tagged proteins were subjected to TEV protease cleavage at 25 °C. The optimal conditions for cleavage was decided between three different ratios of MBP-CHIP protein to TEV 10:1, 20:1, 50:1 (by mass), and two different incubation periods of 1 and 2 hours. For this, the protein concentrations of TEV Protease (produced from pTH24 vector with C-terminal His tag, a kind gift from Rune Kleppe, Dept. of Biomedicine, University of

Bergen) and WT MBP-CHIP fusion protein were first measured on Nanodrop ND-1000^{TM}, and the proper volumes for each protein solution were decided accordingly. The samples were then analyzed for complete cleavage by SDS-PAGE and Coomassie staining as described in Sections 3.15.2 and 3.16.1.

3.7 CHIP ubiquitination activity assay

In vitro ubiquitination activity assay was set up for both MBP-fusion and MBP-cleaved CHIP recombinant proteins. The reactions were prepared in the total volume of 20 µl, containing 2.5 µM (MBP-) CHIP (E3), 2.5 µM UbcH5c (E2) (Boston Biochem, Cambridge, Massachusetts, USA), 0.05 µM Ube1 (E1, a kind gift from Lise Bjørkhaug Gundersen, KG Jebsen Center for Diabetes Research, University of Bergen), 250 µM Ubiquitin (Boston Biochem), and 0.8 µM His-$HSPA_8$ (HSC71) Recombinant Human Protein (Life Technologies AS, Thermo Fisher Scientific) in an ubiquitination buffer (50 mM Tris HCl, pH 7.5, 0.6 mM DTT, 2.5 mM Mg-ATP). MilliQ water was added at the end to reach the final solution volume. A one hour incubation at 37 °C was allowed to complete each reaction before the samples were placed on ice. Protein samples were explored for self- and Hsc70-ubiquitination activities by Western blot analysis and SDS-PAGE electrophoresis (see Sections 3.17 and 3.15.2), using primary antibodies against CHIP (LS-C137950, LifeSpan Bioscience, Seattle, WA, USA) and Hsc70 (ADI-SPA-815-F, Enzo Life Sciences, NY, USA), respectively. The goat anti-rabbit IgG-HRP (sc-2030, Santa Cruz Biotechnology, Dallas, Texas, USA) antibody was used as the secondary antibody in both cases.

3.8 Limited proteolysis assay

Limited proteolysis assay was carried out using 30 µg of the CHIP recombinant proteins separated from the MBP tag by TEV protease beforehand (Section 3.6). To start proteolysis, trypsin was added to each protein sample at a CHIP to trypsin ratio of 1:600 (by mass) in a 100 µl reaction mixture (50 mM NaCl, 20 mM Hepes, and 2 mM DTT) and incubated at 25 °C over a 30 min period of time. The reactions were terminated at different time points (0, 5, 10, 20, 30 min) by adding 18.5 µl of the reaction mixture to 5 µl NuPAGE LDS Sample Buffer (4X) (Thermo Fisher Scientific), 1.4 µl NuPAGE Sample Reducing Agent (10X) (Thermo Fisher Scientific),

and 1 µl of the trypsin inhibitor (previously prepared using a protease to inhibitor mass ratio of 1:1.5) on ice. Finally, all the samples were analyzed by SDS-PAGE and SYPRO Ruby staining as described in Sections 3.15.2 and 3.16.2. Full-length CHIP protein gel bands were quantified by Multi Gauge v3.0 software (Fujifilm, Tokyo, Japan) and Image Processing and Analysis in Java (Image J, National Institutes of Health, Bethesda, Maryland, USA) software. Rates of proteolysis were further plotted using Microsoft® Office Excel 2010 (Microsoft ® Corporation, Redmond, WA, USA).

3.9 CHIP protein purification on amylose resins

To remove MBP tag from the CHIP protein solution, ~400 µg of the WT MBP-CHIP fusion protein was tested for amylose purification after being separated from MBP by TEV protease cleavage (Section 3.6). The Amylose Resin stored in 20% Ethanol (New England Biolabs) was prepared in a 200 µl volume as a 50% slurry with the protein storage buffer (for the buffer composition, see Section 3.5.5). The MBP-cleaved protein sample was added to the resins, and rotation was performed for 30 min at 4 ˚C. Next, the sample was transferred to the Micro Bio-Spin™ Size Exclusion Spin Columns, and washed twice with 150 µl storage buffer. The flow-through was collected in separate Eppendorf tubes each time for further analysis by SDS-PAGE and Coomassie blue staining as described in sections 3.15.2 and 3.16.1. For a complete recovery of the flow-through, the resin was centrifuged for 2 min at 1500x g following each wash.

3.10 Purification of CHIP by size-exclusion chromatography

The purification of CHIP from MBP-CHIP protein solution was tested by size-exclusion chromatography using BioLogic DuoFlow™ Medium-Pressure Chromatography System (Bio-Rad Laboratories, Hercules, California, USA) after it had been separated from MBP tag by TEV protease cleavage (see Section 3.6). Two hundred µg of the WT MBP-cleaved CHIP protein was injected in a final volume of 600 µl protein solution to the top of the Superdex 200 Increase 10/300 GL (GE Healthcare) chromatography column, passing through the column at a constant flow rate of 0.5 ml/min for 50 min at 4 °C (performed under supervision). To avoid possible aggregation, the protein sample was centrifuged for 10 min at 16000 rpm at 4 °C before injection. In addition, the column was equilibrated overnight with two column volumes (50 ml) of a buffer

containing 20 mM NaH_2PO_4, pH 7.4 before starting separation. The purification process was repeated twice using different conditions so that optimized purification among the three trials could be achieved. For the second chromatography, the amount of protein was increased by as much as 2 mg, with a flow rate of 0.25 ml/min through the column that was equilibrated overnight by using another buffer composition (50 mM NaH_2PO_4, 5 mM DTT, pH 7.2). The last chromatography was performed using Superdex 75 Increase 10/300 GL (GE Healthcare) as the gel-filtration column material. A buffer containing 100 mM Hepes, pH 8, 5 mM DTT, 100 mM NaCl was used for column equilibration, and the same amount of protein (2 mg) passed through the column at a flow rate of 0.5 ml/min. At the end of each run, using BioLogic DuoFlow software (Bio-Rad Laboratories), those fractions corresponding to the recorded elution peaks were collected and analyzed for protein content with SDS-PAGE and Coomassie blue staining as described in Sections 3.15.2 and 3.16.1.

3.11 Oligomerization study of MBP-CHIP fusion proteins

Different oligomeric states of the WT and mutant MBP-CHIP fusion proteins were studied by gel-filtration analysis, using the BioLogic DuoFlow™ Medium-Pressure Chromatography System. One mg of MBP-CHIP protein was run on the Superdex 200 Increase 10/300 GL chromatography column at a constant flow rate of 0.3 ml/min as described in Section 3.10. The column was equilibrated overnight with 50 ml phosphate buffer (20 mM NaH_2P0_4, pH 7.4) in advance. BioLogic DuoFlow software was used to monitor the elution peaks at the end of each run. Protein fractions associated with the recorded peaks were collected and analyzed using SDS-PAGE and Coomassie blue staining as described in Sections 3.15.2 and 3.16.1.

3.12 Circular Dichroism (CD) spectroscopy

The secondary structure and thermal unfolding of MBP-CHIP fusion proteins (WT and mutants) as well as MBP protein were monitored with a spectropolarimeter. All the samples were prepared at a concentration of 6 µM in a CD compatible buffer containing 10 mM potassium phosphate (pH 7.4), and 100 mM sodium fluoride. The buffer was exchanged following the protocol provided by Zeba™ Spin Desalting Columns (see Section 3.5.5). Protein concentrations were precisely determined by the protein $A_{(280)}$ method, in which the protein absorption at 280 nm was

measured in a spectrophotometer and used to calculate the protein concentration ($mg.ml^{-1}$), from the following formula:

$$C=A/\mathcal{E}\,(0.1\%) * I$$

Where $\mathcal{E}$ is the extinction coefficient number for a 0.1% protein solution (1.224 for MBP-CHIP and 0.512 for MBP, given by [66]), and I is the cuvette path length in cm. The protein absorbance was measured over the range of 220-340 nm, using a 1 cm absorption cuvette (Hellma Analytics, Müllheim, Germany), and Agilent 8453 UV-Visible spectrophotometer (Agilent Technologies) at a temperature maintained at 20 °C, controlled by a Hewlett Packard 89090A Peltier Temperature Controller (Hewlett-Packard, California, USA).

Far-UV spectra and thermal denaturation analysis were performed for 300 µl of each protein sample in a 1 mm quartz cell (Hellma Analytics), using the Spectra Manager software v.1.53.04 (Jasco Corporation, MD21601, USA) and a J-810 Jasco spectropolarimeter equipped with a CDF-426S Peltier element for temperature control (Jasco Products Company, Oklahoma City, Oklahoma, USA). The initial nitrogen flow rate of 10 l/min was used to measure the far-UV spectra in the wavelength range between 185–260 nm at a scan rate of 50 nm/min at 20 °C. Four scans were accumulated for each spectrum and averaged. The final spectra were recorded after buffer-subtraction. Thermal denaturation profiles were obtained at 5 l/min nitrogen flow rate through recording the ellipticity at 222 nm as a function of temperature in the range 20–90 °C with a scan rate of 40 °C per hour. Online Dichroweb software (Dr. L. Whitmore, University of London, London, UK) was used to calculate secondary structure content of proteins from their far-UV spectra [67]. Figure 3.12.1 demonstrates a typical thermal unfolding curve for a globular protein, consisting of at least one melting transition region (two-state unfolding pattern) where the unfolded (U) form of protein is resulted from denaturation of the native (N) protein at a temperature reported as transition midpoint temperature or T_m.

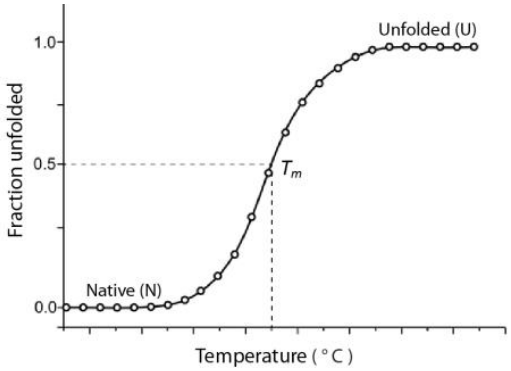


Figure 3.12.1 Two-state thermal unfolding profile of globular proteins. The transition region indicates that the native protein unfolds while the temperature is increasing. T_m or denaturation midpoint is defined as the temperature at which both the unfolded (U) and folded (N) states of protein are presented equally at equilibrium.

For in cellulo protein studies

3.13 Cell cultivation

Human Embryonic Kidney 293 (HEK293) (ATCC, Manassas, Virginia, USA) cells were cultured in Dulbecco's modified eagles medium (DMEM) (4.5 g/L glucose) (Invitrogen) containing 10% heat-inactivated FBS (Invitrogen) and 0.05 mg/ml of active Gentamycin® (50 mg/ml) (Life technologies). The cells were grown at 37 °C with 5% CO_2 (Steri-Cycle CO_2 Incubator, Thermo Scientific); using 25 cm^2 and 75 cm^2 Nucleon™ Delta Surface cell culture flasks (Thermo Scientific). The media was changed every second day. All cell procedures were carried out in a Scanlaf Mars Flowbench (Labogene ApS, 3540 Lynge, Denmark), and all the volumes correspond to 75 cm^2 flasks, if not otherwise stated.

3.13.1 Thawing

Frozen cells were thawed quickly by placing Cryotubes™ Vials (Nunc™, Thermo Scientific) in a water bath (Lauda Aqualine water bath) (97912 Lauda-Konigshofen, Germany) at 37 °C for 2-3 min. To the thawed cells in a 15 ml Falcon tube, 4 ml growth media was added, and the tube was centrifuged (Heraeus Megafuge 1.0R centrifuge) (DJB Labcare Ltd., Buckinghamshire, UK) for 6 min at 1300 rpm. The cell pellet obtained was resuspended in 5 ml media and transferred to a 25 cm^2 flask for further growth.

3.13.2 Splitting

Cells were passaged into new flasks after reaching 80-90% confluency. The media was removed and the cells were detached from the flask surface after a quick wash with 10 ml PBS (Invitrogen). Two ml 0.05% Trypsin-EDTA (Invitrogen) was used to trypsinize the cells. After 5 min incubation at room temperature cells were monitored under the microscope for complete detachment. Trypsin was next neutralized by adding 8 ml media to the flask, and resuspended cells were transferred into new flasks together with fresh growth media in a total volume of 10 ml, using various ratios of cells to media.

3.13.3 Freezing

In order to freeze cells, a flask of ~80% confluent cells was washed with 10 ml PBS and trypsinized as described in Section 3.13.2. Following trypsin inactivation, the cell suspension was centrifuged in a 15 ml Falcon tube for 5 min at 1300 rpm. The pellet was resuspended in 4 ml BAMBANKER™ serum-free cell freezing medium (Lymphotec, Inc., Tokyo, Japan) and aliquoted into four cryotubes. The tubes were kept at room temperature for 20 min before a long-term freezing at -80 °C.

3.14 Determination of exogenous protein expression level

3.14.1 Cell transfection

Transiently transfected HEK293 cells expressing WT and mutant CHIP were prepared by using Metafectene. Cells were seeded into Nucleon™ Delta Surface 6-well plates (Thermo Scientific) at a concentration of one million cells per ml (10^6 cells/ml), counted by using a Scepter™ Handheld Automated Cell Counter (Merk Millipore, Massachusetts, USA). Each well was coated with 1 ml Poly-L-Lysine 0.01% solution (Sigma-Aldrich) (1:10 diluted in PBS) and washed with PBS three times before seeding. One million cells (1 ml) were added to 1 ml of media (DMEM without gentamycin) in each well, and grown overnight at 37 °C. In order to transfect the cells, 3 µg of each WT and mutant *STUB1* pcDNA/V5-HisB plasmid as well as a negative control (pcDNA/V5-HisB plasmid without *STUB1* cDNA) was mixed with 2 µl Metafectene® Easy$^+$ and 75 µl 1x Easy$^+$ buffer (Biontex Laboratories GmbH, Martinsreid, Germany), and incubated for 15 min at room temperature. Cells were further added by 50 µl of the prepared mixture in each well, and grown at 37 °C overnight. The media was changed to the one containing gentamycin antibiotic after 4-6 hours to avoid possible contamination. Cells were collected and lysed the following day as is described in the next section.

3.14.2 Cell lysis

Cells were washed twice with 1 ml cold PBS prior to being lysed. Two hundred µl Pierce® RIPA buffer (Thermo Scientific) containing 1x Halt Protease Inhibitor Cocktail (PIC) (100X) (Thermo Scientific) was used to lyse the cells in each well. Cell lysates were next collected and transferred

into Eppendorf tubes by using a cell scraper (TPP Techno Plastic Products AG, Trasadingen, Switzerland). This was followed by a 15 min centrifugation at 15000x g whereby a cleared lysate was achieved and stored at -80 °C until further use (see the next section).

3.14.3 Determination of cellular protein concentration

The Pierce™ Bicinchronic acid (BCA) protein assay kit (Thermo Scientific) was used to determine the concentrations of exogenously expressed proteins in HEK293 cell lysate fractions by measuring absorbance at 562 nm ($A_{(562)}$). Measurements were performed in 96-well microplates (Greiner Bio-One GmbH, 72636 Frickenhausen, Germany) on a Synergy™ HT plate reader (Bio Tek Instruments, Inc., VT, USA), using BSA standards and working reagents. Standard reagents were prepared by diluting the BSA standard stock solution in RIPA buffer to a series of concentrations from 0 to 2000 µg/ml according to the supplier's protocol. The working reagent was made by mixing 50 parts of BCA Reagent A with 1 part of BCA Reagent B, and 200 µl volume was added to 25 µl of each unknown and standard sample replicates in the microplate wells. Samples were mixed on a plate shaker for 30 second and incubated for 30 min at 37 °C. Protein absorbance was measured two times for both BSA standards and protein samples in parallel, and the average concentration for each sample was obtained from a linear regression standard curve using the Gen5 Microplate Reader and Imager Software (Bio Tek).

Cell lysates were next analyzed by SDS-PAGE (see Section 3.15.2) and immunoblotting (see Section 3.17), using a primary antibody against CHIP and secondary antibody against rabbit. Samples were loaded in volumes equal to 10 µg cellular protein. To provide an internal loading control, the membrane was stripped and blotted against actin as described in Section 3.17.3. A set of antibodies against actin (sc-1615, Santa Cruz Biotechnology) and rabbit anti-goat IgG-HRP (sc-2768, Santa Cruz Biotechnology) was used in this step. CHIP protein bands were finally quantified using Multi Gauge and Image J softwares, and were finally normalized against actin.

For protein detection and analysis

3.15 Electrophoresis

*If not otherwise stated, all the materials in the electrophoresis techniques were purchased from the Thermo Fisher Scientific Company.

3.15.1 Agarose gel electrophoresis

For separation and analysis of DNA fragments by molecular weight, agarose gel electrophoresis was performed using 50 µl of 1% (w/v) Nusieve agarose gel placed in 1x TBE buffer (Sigma-Aldrich). Three drops of ethidium bromide were added to the gel prior to solidification to visualize the bands on the gel. Six µl of loading sample was prepared by mixing 3 µl of DNA samples and DNA Gel Loading Dye (6x) to a final concentration of 1x, and loaded onto the gel along with 5 µl of DNA standard size marker (New England BioLabs). Following electrophoresis at 80 V, DNA fragments were visualized by UV-light using a Molecular Imager® Gel Doc™ XR System (BioRad Laboratories).

3.15.2 SDS polyacrylamide gel electrophoresis (SDS-PAGE)

Using sodium dodecyl sulfate polyacrylamide gels, proteins were linearized and separated by size during SDS-PAGE. Ten µl of protein samples were added with NuPAGE® LDS Sample Buffer (4x) and NuPAGE® Sample Reducing Agent (10x) to a final concentration of 1x followed by denaturation at 70 °C for 10 min. Maximum recommended volumes of the denatured proteins were applied together with 5 µl of SeeBlue® Plus2 Pre-stained Protein Standard ladder on the wells of precast Bolt™ 10% Bis-Tris Plus Gels. In order to prevent sample reoxidation, and to maintain reducing conditions, 500 µl of Nupage® Antioxidant was added to the 1x MOPS running buffer (prepared in a total volume of 1 liter NuPAGE® MOPS SDS Running Buffer (20x) in the Upper Buffer chamber of the XCell™ *sureLock*™ Mini-Cells. The gel was run at 200 V for 1 hour, and the bands were visualized using the SimplyBlue SafeStain Coomassie staining technique as described in Section 3.16.1.

3.15.3 Native polyacrylamide gel electrophoresis (Native-PAGE)

To study different conformational states of the WT and mutant CHIP proteins, native gel

electrophoresis was performed using NativePAGE™ Bis-Tris Gels (12%) whereby the proteins get separated with maximum stability according to their molecular weight and electric charge. Wild type and mutant protein samples were prepared in 30 µl loading volume by adding 10 µg of MBP fusion protein to Native PAGE™ sample buffer (4x), at a final concentration of 1x. To avoid protein denaturation, all the samples were prepared on ice. The lower (outer) buffer chamber of the XCell™ *sureLock*™ Mini-Cell was filled with 300 ml of the 1x NativePAGE™ Anode Buffer previously prepared in 1000 ml volume by diluting 50 ml of the pre-chilled NativePAGE™ Running Buffer (20x) in deionized water. For the upper (outer) chamber, 200 ml of the Light Blue 1x NativePAGE™ Cathode Buffer was prepared in similar way using the NativePAGE™ Running Buffer (20x) and 1 ml of the NativePAGE™ Cathode Additive (20x). In order to provide easy visualization, protein samples were applied to the wells in the recommended maximum load volume prior to filling the inner chamber. Seven µl of the NativeMark™ Unstained Protein Standard was loaded as the marker reference for estimation of the size of the bands. Electrophoresis was performed at room temperature for 110 min applying a voltage of 150 V. The gel was further prepared for SimplyBlue SafeStain Coomassie staining as described in Section 3.16.1.

3.16 Staining techniques

*All the steps in the following staining methods were carried out at room temperature with slight agitation.

3.16.1 Coomassie blue staining

Following SDS-PAGE, the polyacrylamide gels were rinsed three times for 5 min with 150 ml MilliQ water. The gels were then incubated in 20 ml SimplyBlue™ SafeStain solution (Thermo Fisher Scientific) for one hour. A clear gel background was achieved by an overnight washing with 150 ml MilliQ water before gel documentation and analysis.

3.16.2 SYBRO Ruby staining

The Molecular Probes SYBRO® Ruby protein gel stain was used as an ultrasensitive luminescent stain to detect the proteins separated on the gels after SDS-PAGE. The gels were first placed into a clean container and fixed with 100 ml of fix solution (100 ml Methanol and 14 ml Acetic acid added to 86 ml MilliQ water) for 30 min. This step was repeated once more with fresh fix

solution. Afterwards, the gels were stained with 60 ml of SYBRO® Ruby gel stain (Thermo Fisher Scientific) and incubated overnight. In order to minimize the background staining, the gels were transferred to a clean container the next day and washed with 100 ml of wash solution (10 ml Methanol and 7 ml Acetic acid added to 83 ml MilliQ water) for 30 min. The proteins were visualized by UV-light using the Molecular Imager® Gel Doc™ XR System (BioRad Laboratories). Before imaging, the gels were washed twice for 5 min with ultrapure water to prevent possible corrosive damage to the imager.

3.17 Immunoblot (western blot) analysis

The SDS-PAGE-separated proteins were transferred on to a nitrocellulose membrane in order to be detected specifically by antibodies. The Nitrocellulose (NC) Membrane Filter Paper Sandwich (0.45 µm pore size, Novex, Life Technologies) was soaked together with the blotting pads and filter papers in 1x NuPAGE® Transfer Buffer (Thermo Fisher Scientific), 10% Methanol, and 0.1% Antioxidant and used to prepare a gel/membrane blotting sandwich following the supplier´s protocol (Invitrogen). The sandwich system was then placed vertically in a tank filled with 1x transfer buffer, where the proteins moved from the gel to the membrane at an electrical voltage of 30 V, for one hour. The transferring process was cooled down using ice in the outer chamber of the tank. The presence of a protein ladder on the nitrocellulose membrane confirms a complete protein transfer step. Transferred proteins were further analyzed by antibody binding (immunoblotting) as described in the next two sections.

3.17.1 Immunodetection

The NC membrane was blocked for unspecific binding reactions by a one hour incubation in 10 ml SuperBlock™ (PBS) Blocking Buffer (Thermo Fisher Scientific) with 0.05% Tween 20 (Sigma-Aldrich). The primary antibody (CHIP and Actin diluted 1:5000, HSC70 diluted 1:10000 in blocking solution) was added to probe the membrane while shaking overnight at 4 °C. The next day the membrane was washed five times with 1x PBST containing 2 PBS tablets (Life Technologies) per 500 ml MilliQ water and 0.05% Tween 20, each time for 5 min and incubated in 10 ml blocking solution with 1:20000 diluted secondary antibody for one hour at room temperature. This was followed by a five-times washing step with 1x PBST before proceeding to protein detection step (see the section below).

3.17.2 Enhanced Chemo Luminescence (ECL) protein detection

Finally, the membrane was developed and the immobilized proteins were visualized using the ECL detection method. In this method, the membrane was incubated for 5 min with an adequate volume of ECL solution to cover the membrane, which was prepared by mixing the enhancer and peroxide solutions 1:1, solutions being provided in the SuperSignal® West Pico Chemoluminescent Substrate Kit (Thermo Scientific). The developed membrane was placed immediately inside the Molecular Imager® Gel Doc™ XR System (FujiFilm, Tokyo, Japan) between two plastic foils with the protein side facing up, and analyzed by Image reader LAS 1000 Pro v2.6 (Fujifilm) software.

3.17.3 Stripping of membrane

This technique was used to detect a different protein on the same nitrocellulose membrane. First, the membrane was washed with 10 ml PBS for 15 min. The previously bound antibodies were next stripped from the membrane by incubation in 10 ml of Restore™ WesternBlot Stripping Buffer (Thermo Scientific) for 30 min. The membrane was then rinsed with PBS for another 15 min before blocking and immunoblotting with a different set of antibodies (for the immunodetection procedure, refer to Section 3.17.2).

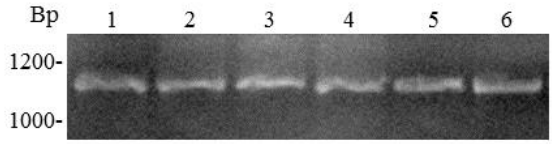
4.0 Results

For in vitro recombinant protein studies

4.1 Generation of mutant *STUB1* constructs by site-directed mutagenesis

Three mutations corresponding to p.K145Q (c.433A>C), p.M211I (c. 633G>A), and p.S236T (c. 707G>C) were introduced into WT *STUB1* cDNA in pETM-41-MBP/His vectors and the vectors were subsequently transformed into XL10-Gold Ultracompetent cells as described in Sections 3.1 and 3.2. Bacterial colonies were selected and examined for successful transformation by PCR amplification and agarose gel electrophoresis (Figure 4.1.1 A). In order to confirm the presence of correct mutations, DNA sequencing was performed for those colonies with correct DNA size (1100 bp for *STUB1*+ MBP/His tag) using primers that cover the mutated region. The DNA sequence of each mutation is shown in Figure 4.1.1 B.

A.



B.

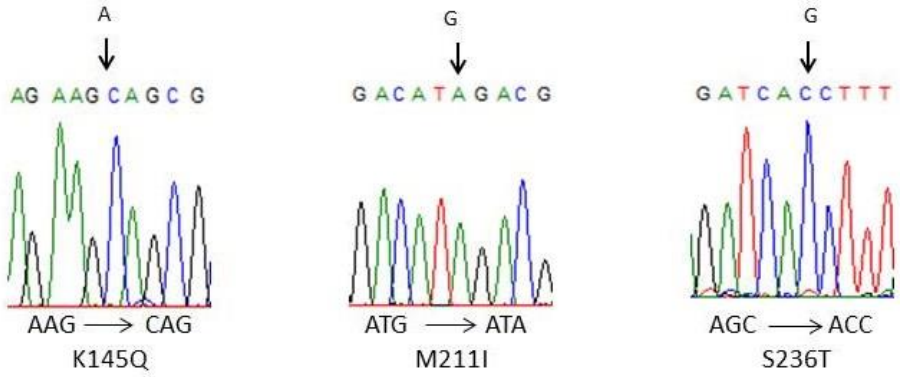


Figure 4.1.1 Generation of *STUB1* mutants by site-directed mutagenesis. The *STUB1* cDNA pETM-41-MBP/His vector was used for *in vitro* recombinant studies. A) Agarose gel showing the PCR product from individual bacterial colonies transformed with potential K145Q (lane 1, 2), M211I (lane 3, 4), and S236T (lane 5, 6) mutant cDNA plasmids using T7 promoter and T7 terminator primers. B) cDNA sequencing of mutant plasmids directly from bacterial colonies, confirming the introduction of K145Q, M211I and S236T mutations analyzed by FinchTV software.

All the mutant plasmids were purified from selected colonies (see Section 3.4) and DNA sequencing of the whole cDNA region verified that no additional mutations were present (data not shown) (for primer sequences and sequencing procedure, see Section 3.3). The plasmid vectors were used for further *in vitro* recombinant protein production together with E28K, N65S, T246M mutant vectors which had previously been prepared by other members of the research group [30].

4.2 Expression and purification of recombinant MBP-CHIP fusion proteins

For *in vitro* studies of WT and mutant CHIP ligases, the His-tagged MBP fusion proteins were expressed and purified from BL21 competent *E.coli* following the Ni-NTA purification procedure as described in Section 3.5. Successful protein expression was verified for each mutant by SDS-PAGE and Coomassie blue staining before and after IPTG induction as shown in Figure 4.2.1. The expressed MBP-CHIP proteins were seen as protein bands of ~72.5 kDa after 24 hours of IPTG induction. The fusion proteins were further cleaved by TEV protease in order to separate MBP from CHIP. To examine the optimal MBP/TEV ratio, three different ratios of TEV/MBP-CHIP (1:10, 1:20, and 1:50, by mass) were tested during cleavage periods of one and two hours (see Section 3.6). Results from the SDS-PAGE analysis showed a complete cleavage after two hours with the ratio of TEV protease to MBP-CHIP 1:10 and 1:20, both generating three protein bands of MBP (~42.5 kDa), CHIP (~35 kDa), and TEV (~29 kDa) on the gel (Figure 4.2.2), with only a tiny fraction left of uncleaved MBP-CHIP. Therefore, a TEV/MBP-CHIP ratio of 1:10 for a 2 hour incubation was used for the downstream studies.

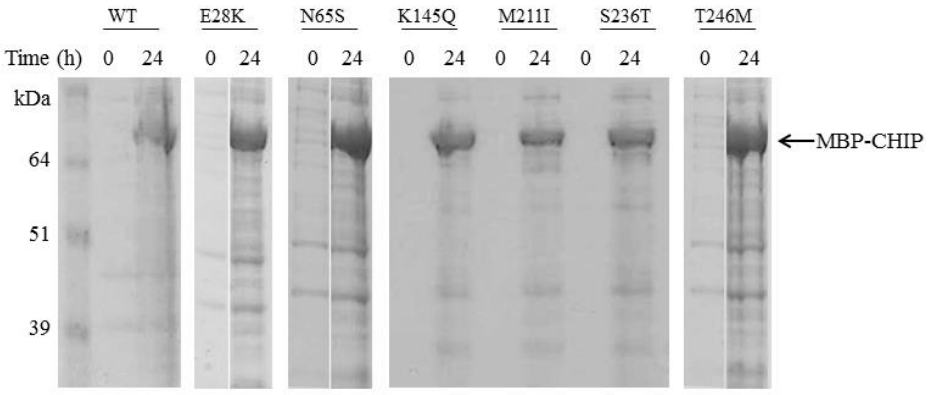


Figure 4.2.1 Expression of recombinant WT and mutant MBP-CHIP proteins. The expression of MBP-CHIP proteins of both WT and mutant forms at time 0 and 24 hours after induction of protein expression by 0.5 mM IPTG in BL21 *E.coli* cells, shown by SDS-PAGE and Coomassie blue staining.

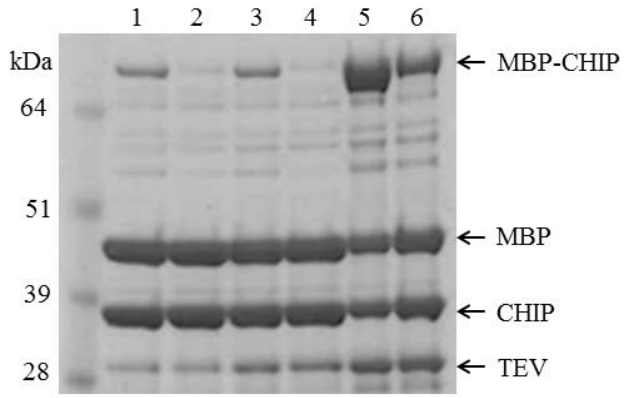


Figure 4.2.2 MBP-CHIP fusion protein cleavage by TEV protease. Coomassie blue stained SDS-PAGE gel loaded with 30 µg TEV-cleaved WT MBP-CHIP protein, using TEV protease to MBP-CHIP mass ratio of 1: 10 and an incubation time of one hour (lane 1), 1:10, two hours (lane 2), 1:20, one hour (lane 3), 1:20, two hours (lane 4), 1:50, one hour (lane 5), and 1:50, two hours (lane 6).

4.3 *In vitro* ubiquitination activity of E3 ubiquitin ligase CHIP for WT and mutant variants

As a co-chaperone with ubiquitin ligase activity, CHIP participates in the ubiquitination of unfolded or misfolded substrates bound to chaperones via its U-box domain, mediating their degradation within the ubiquitin-proteasome pathway. In addition to the substrates, CHIP itself, and the heat shock protein chaperones are also known to be ubiquitinated by the same mechanism. In order to examine whether the enzymatic activity of CHIP is altered for the mutants, the ubiquitination assay was performed on recombinant WT and mutant CHIP both as MBP-fusion and MBP-free (cleaved) proteins. As described in Section 3.7, the reaction was performed for one hour at 37˚C, using Hsc70 recombinant protein as the ubiquitination substrate. The negative control sample was prepared without adding ubiquitin to the reaction. Proteins were then analyzed for both self- and Hsc70- ubiquitination by Western blotting against CHIP and Hsc70, respectively.

Results are shown in Figure 4.3.1 A and B. A complete ubiquitination reaction was observed as additions of several 8.5 kDa ubiquitin molecules to Hsc70 protein (shown as protein bands with increasing molecular weight) for the WT and four of the CHIP mutants. However, N65S and T246M displayed impaired activity for both MBP-fusion and MBP-free forms (Figure 4.3.1, lanes 4 and 8). Hsc70 seemed to be mono-ubiquitinated by the N65S mutant, while no Hsc70-ubiquitination was detected for the T246M. Furthermore, T246M was found to be non-functional in terms of self-ubiquitination for both MBP-fusion and cleaved proteins, but the N65S mutant showed similar level of self-ubiquitination as the WT, and other mutants, in both protein forms.

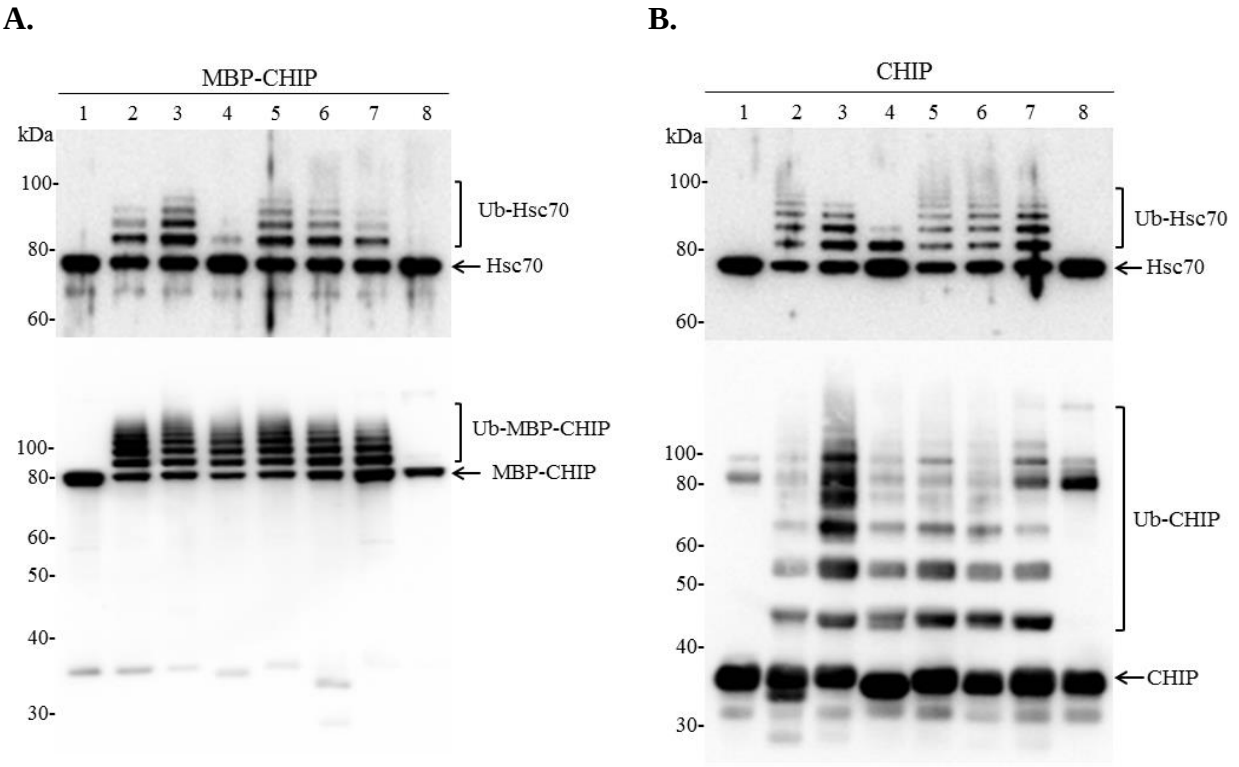


Figure 4.3.1 Ubiquitination activity of recombinant WT and mutant CHIP proteins. *In vitro* ubiquitination activity assay was performed using WT (lane 2), E28K (lane 3), N65S (lane4), K145Q (lane5), M211I (lane 6), S236T (lane 7), and T246M (lane8) as ubiquitin ligases and Hsc70 recombinant protein as substrate. Self- ubiquitination and Hsc 70-ubiquitination activities were explored on MBP-CHIP fusion proteins (A) and MBP- cleaved CHIP protein (B). The reaction was incubated for one hour at 37 °C, and samples were analyzed for both self- and Hsc70-ubiquitination by SDS-PAGE and Western blotting using antibodies against CHIP and Hsc70, respectively. As a negative control (lane1), the WT CHIP protein was used in a separate reaction set up without adding ubiquitin.

In summary, the N65S and T246M variants were mostly affected by the mutation in this context, and the ubiquitination activity of proteins was not found to be significantly different when comparing cleaved (Figure 4.3.1 A) and MBP-tagged CHIP (Figure 4.3.1 B).

4.4 Protein stability analysis by limited trypsin proteolysis

As a serine protease, trypsin targets peptide chains at the carboxyl sides of the amino acids lysine or arginine. Some of these residues are normally buried inside the protein structure, being protected from proteolytic cleavage. The rate of tryptic cleavage can identify conformational changes happening in the protein structure caused by certain mutations. Thus mutations resulting

in increased structural flexibility can accelerate trypsin access to cleavage sites, and thereby increase the degradation rate. Therefore, in order to elucidate the impaired ubiquitination activity observed in the N65S and T246M variants, and to further characterize the structure and dynamics of the mutants, limited proteolysis assay was performed using trypsin for digestion. MBP-cleaved CHIP proteins were subjected to trypsin cleavage during a time period of 30 min, and their susceptibility to proteolysis was analyzed at different time points by SDS-PAGE analysis and SYBRO Ruby staining (see Section 3.8).

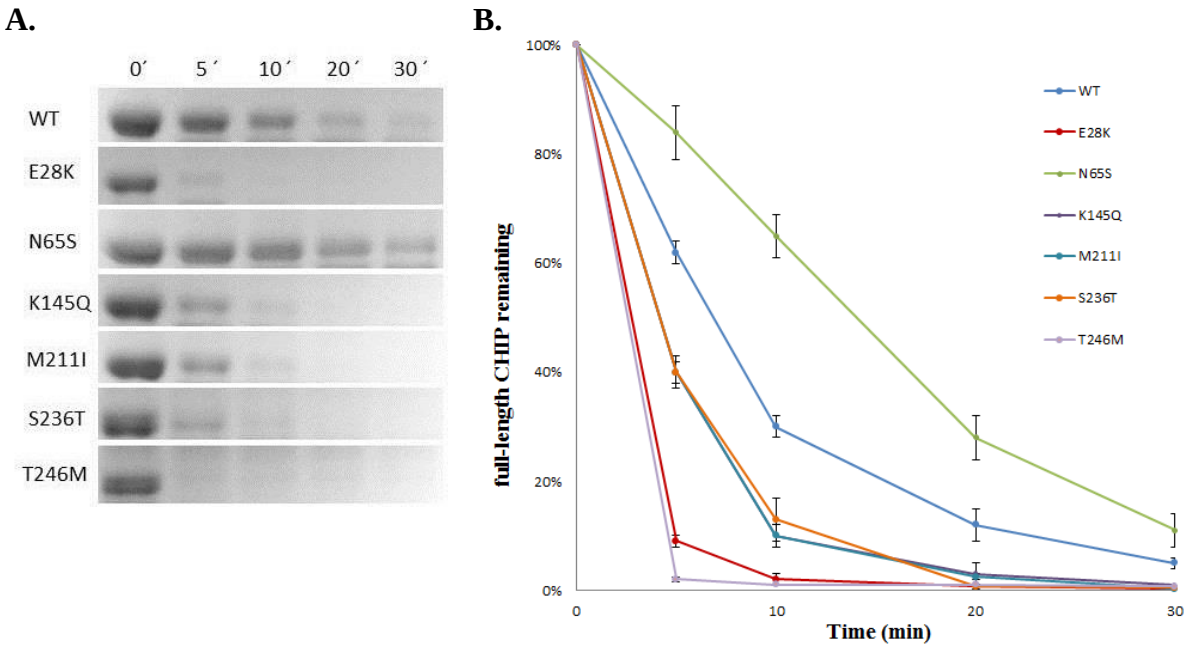


Figure 4.4.1 Limited proteolysis of WT and mutant CHIP proteins. The proteolytic susceptibility of MBP-free CHIP proteins was explored for WT and mutants using trypsin for protein digestion. A) Proteins were detected by SDS-PAGE and SYBRO Ruby staining after being subjected to proteolysis for various time periods (0-30 min). B) Full length proteins were quantified by Multi Gauge and ImageJ software, and the data for the average of three individual experiments were plotted against time. Each time point represents the mean of three readings ± SD (*n*=3).

Results for the limited proteolysis of the WT and mutant CHIP proteins are presented in Figure 4.4.1. Compared to the WT, increased rate of proteolysis was shown for all but one of the mutants as less detectable protein at time 5 min (and later), while N65S which was associated with a slower rate of trypsin cleavage. Among those mutants with increased susceptibility to proteolysis, the E28K and T246M variants were most affected, as indicated by a more rapid cleavage observed at 5 min for these mutants compared to other mutants including K145Q, M211I, and S236T.

Overall, the N65S was found to be the only mutant with increased structural stability against proteolysis.

4.5 Small-scale CHIP protein purification by amylose resin

Amylose resin generates a cross-linked matrix with high affinity for maltose-binding protein which can be used for the isolation of proteins fused to MBP tag. In order to test whether the purification of CHIP, free from MBP, can be achieved with this system, the TEV-cleaved WT MBP-CHIP was bound to amylose resin (see Section 3.9). In this way, CHIP is expected to remain in the flow through while MBP should bind to the amylose resin. Figure 4.5.1 shows results from the SDS-PAGE and Coomassie blue staining of the flow through after two washing steps. The presence of MBP (~42.5 kDa) together with CHIP (~35 kDa) and TEV (~31 kDa) in both lanes indicates that MBP, unexpectedly, did not bind to the amylose resins and thus protein purification was unsuccessful.

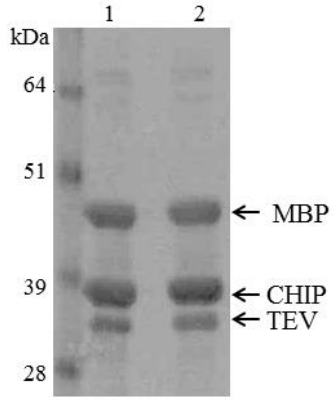


Figure 4.5.1 Purification of CHIP from MBP using amylose resin. ~400 µg of the TEV-cleaved WT MBP-CHIP was subjected to amylose resin purification by being passed through a gravity flow column. The collected flow through from the first (lane 1) and second (lane 2) washing steps were analyzed by SDS-PAGE and Coomassie blue staining.

4.6 Size-exclusion chromatography as a purification strategy

Protein purification can be achieved by using gel-filtration chromatography columns in which the medium, consisting of a porous matrix of spherical particles, is packed to form a chromatographic bed. Upon the injection of protein solution, larger molecules pass through the column at the same speed as the buffer flows, and appear as the first elution peak on the chromatogram after being monitored for UV light absorption at 280 nm. Smaller molecules with partial access to the matrix pores elute from the column later in order of decreasing size and create the next peaks of the chromatogram. The smallest molecules with full access to the pores appear in the final column eluate. Purification of CHIP from MBP was tested with this method,

using different experimental conditions as described in Section 3.10. Fractions were collected and analyzed by SDS-PAGE and Coomassie blue staining. Following the procedure, three separated peaks were expected to be seen on the chromatogram due to the elution of dimeric CHIP (~70 kDa), MBP (~42.5 kDa), and TEV protease (~28 kDa), in that order.

Figure 4.6.1 shows the results from purification of 200 µg WT TEV-cleaved MBP-CHIP protein after passing through the Superdex 200 Increase 10/300 GL chromatography column, using phosphate buffer (20 mM NaH_2PO_4, pH 7.4) as the mobile phase. As is shown (Figure 4.6.1 A), the first small elution peak (Peak I, fractions 14, 15) appeared on the chromatogram at approximately 14 ml, immediately followed by another large peak (Peak II, fractions 16-20) at 16 ml. The last peak (Peak III, fractions 28, 29) showed up at the end of the elution profile (21 ml) and was predicted to contain TEV protease.

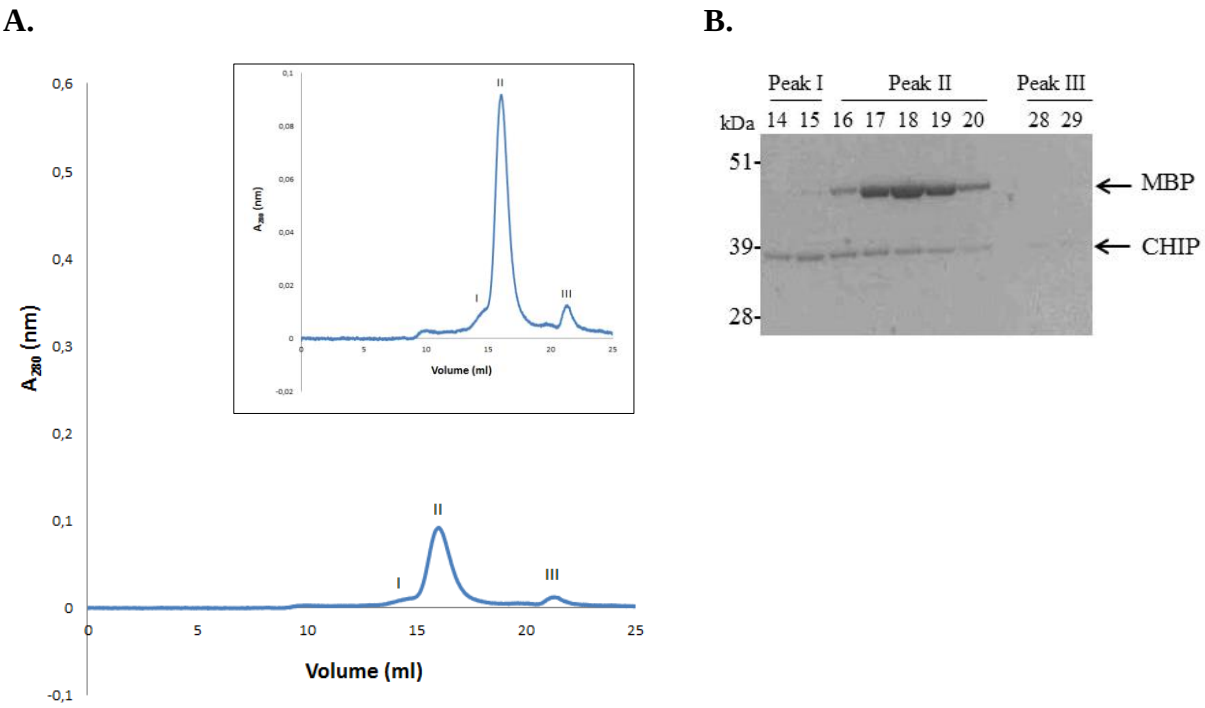


Figure 4.6.1 Separation of WT CHIP from MBP by size-exclusion chromatography. The experiment was performed on 200 µg TEV-cleaved WT MBP-CHIP at 0.5 ml/min with 50 ml phosphate buffer (pH 7.4) as the mobile phase while monitoring protein absorbance at 280 nm. A) The elution profile obtained from chromatography on a Superdex 200 10/300 GL column at 4 °C, showing three elution peaks at 14, 16, and 21 ml. B) SDS-PAGE analysis of the collected fractions associated with peaks I-III.

Results from SDS-PAGE analysis (Figure 4.6.1 B) indicated the presence of only CHIP (~35 kDa) in fractions 14 and 15, corresponding to the first small peak. Fractions of peak II contained both CHIP and cleaved MBP (~42.5 kDa), and no protein band with molecular weight corresponding to TEV protease (28 kDa) was detected in fractions from the last peak. However, weak protein bands corresponding to CHIP seemed to be presented in these lanes. These results suggest that CHIP was eluted continuously in nearly all fractions, and a complete purification was not achieved during chromatography. Besides, those fractions containing only CHIP were too few to obtain a proper concentration of the protein.

In the second attempt made to obtain pure CHIP protein (devoid of MBP), 2 mg of TEV-cleaved MBP-CHIP was subjected to gel-filtration chromatography, using a Superdex 200 Increase 10/300 GL column using a different buffer (50 mM NaH_2PO_4, 5 mM DTT, pH 7.2).

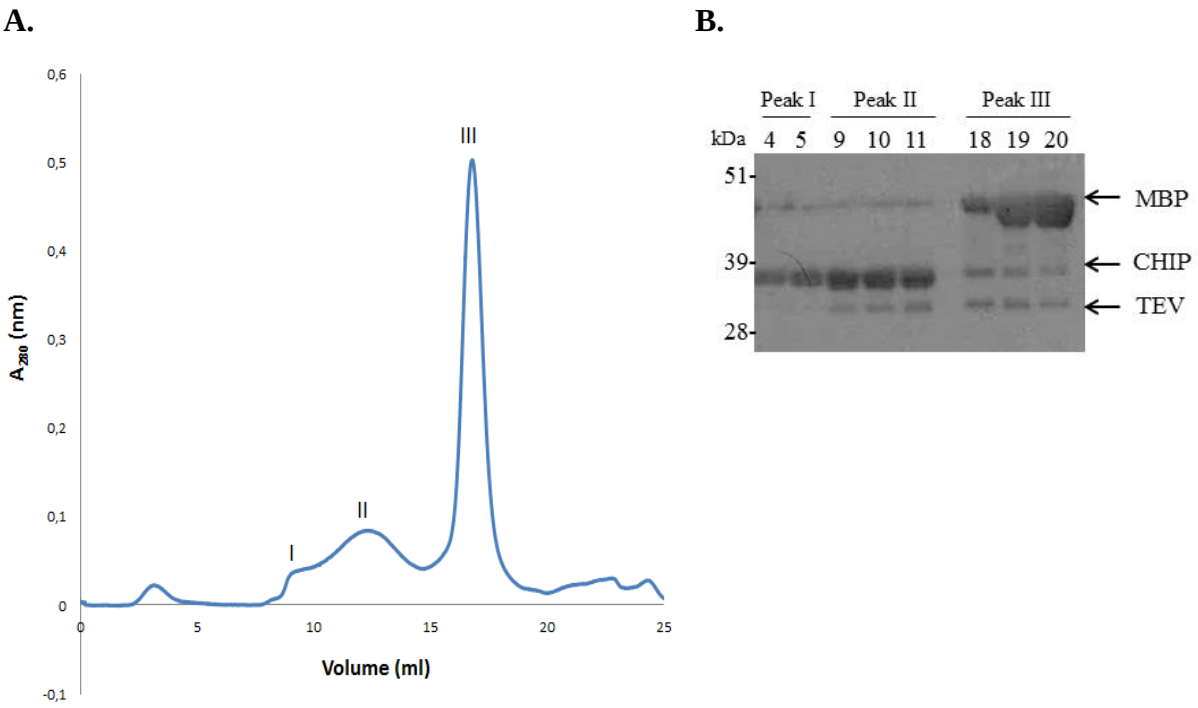


Figure 4.6.2 Separation of WT CHIP from MBP by size-exclusion chromatography. Two mg of TEV-cleaved WT MBP- CHIP was subjected to Superdex 200 10/300 GL chromatography column, flowing with the rate of 0.25 ml/min at 4 °C. A buffer containing 50 mM NaH_2PO_4, 5 mM DTT (pH 7.2) was used as the mobile phase while monitoring protein absorbance at 280 nm. A) Elution profile showing three peaks at 9, 11, and 15 ml. B) SDS-PAGE analysis of collected fractions associated with peaks I-III.

Results showed three connected peaks at approximately 9 ml (Peak I, fractions 4, 5), 11 ml (Peak II, fractions 9, 10, 11), and 15 ml (Peak III, fractions 18, 19, 20) (Figure 4.6.2 A). SDS-PAGE analysis of protein fractions detected CHIP and MBP as the major protein contents of the first two peaks and the last one, respectively (Figure 4.6.2 B). But all fractions seemed to contain other proteins, too. In fact, proteins were separated with a higher yield in this experiment than the previous one and the percentage of MBP was minor compared to the total yield of purified CHIP. However, the required concentration of CHIP was not yet achieved.

Using a longer chromatography column (Superdex 75 Increase 10/300 GL), the gel-filtration analysis was repeated once again for 2 mg of the cleaved WT MBP-CHIP. Columns with a longer packed bed provide more resolving power, and improve resolution in general. The column was equilibrated overnight with a buffer containing 100 mM Hepes, pH 8.5 mM DTT, 100 mM NaCl.

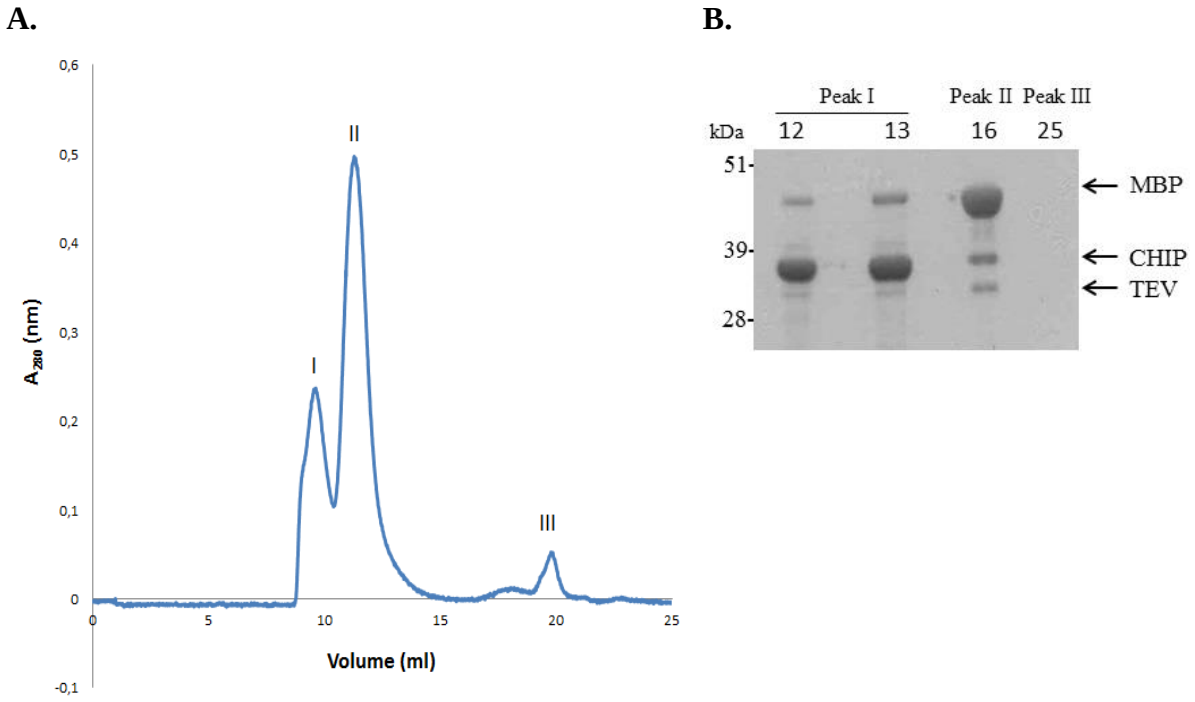


Figure 4.6.3 Separation of WT CHIP from MBP fusion protein by size-exclusion chromatography. Purification of 2 mg MBP-cleaved WT CHIP was tested on Superdex 75 10/300 GL column at 0.5 ml/min at 4 °C. A buffer containing 100 mM Hepes (pH 8), 5 mM DTT, and 100 mM NaCl was used as the mobile phase while monitoring protein absorbance at 280 nm. A) Elution profile showing three peaks at 9, 11, and 19 ml. B) SDS-PAGE analysis of collected fractions associated with peaks I-III.

Results from the protein elution profile (Figure 4.6.3 A) suggested an improved separation as can be seen by the appearance of two partially separated peaks at 9 ml (Peak I, fractions 12, 13) and 11 ml (Peak II, fraction 16) associated with CHIP and MBP proteins, respectively. This was followed by another peak at 19 ml (Peak III, fraction 25). SDS-PAGE analysis results (Figure 4.6.3 B) showed an increased amount of purified CHIP in fractions 12 and 13. However, the percentage of total MBP and TEV protease detected in these fractions was still substantial (~41% in fraction 12 and ~57% in fraction 13). No protein band was observed in fraction 25, corresponding to the last peak, probably because the amount of protein was too low to be detected by Coomassie staining.

Since a complete separation of CHIP from MBP could not be achieved neither with amylose resin nor with size-exclusion chromatography methods, it was decided that the remaining experiments for the *In vitro* protein studies should be performed on the MBP fusion proteins.

4.7 Oligomeric structure of the WT and mutant recombinant MBP-CHIP fusion proteins

As mentioned earlier, CHIP proteins are functional as dimers that are formed through interactions between their putative coiled-coil domains [36, 37]. To study the effects of mutations on oligomerization properties of CHIP protein, gel-filtration analysis was performed on recombinant MBP-CHIP fusion protein of both WT and mutant forms as described in Section 3.11. With this method, different conformational states of protein, including monomeric, dimeric, and oligomeric forms, can be separated according to molecular weight by passage through the size-exclusion chromatography column. Results from gel-filtration chromatography of WT MBP-CHIP indicated the presence of dimeric structures as the dominant oligomeric state of the protein, visible as a high peak located at approximately 12 ml (Figure 4.7.1 A, Peak II) on the chromatogram. The amount of other protein structures appeared to be less as can be seen by the smaller oligomeric and monomeric peaks at ~9 ml (Peak I) and ~15 ml (Peak III), respectively. SDS-PAGE analysis confirmed the presence of MBP-CHIP fusion proteins with protein bands of ~77.5 kDa in all fractions (Figure 4.7.1 B). SDS-PAGE though is not suitable for studying protein oligomerization as it is a denaturing protein electrophoresis technique.

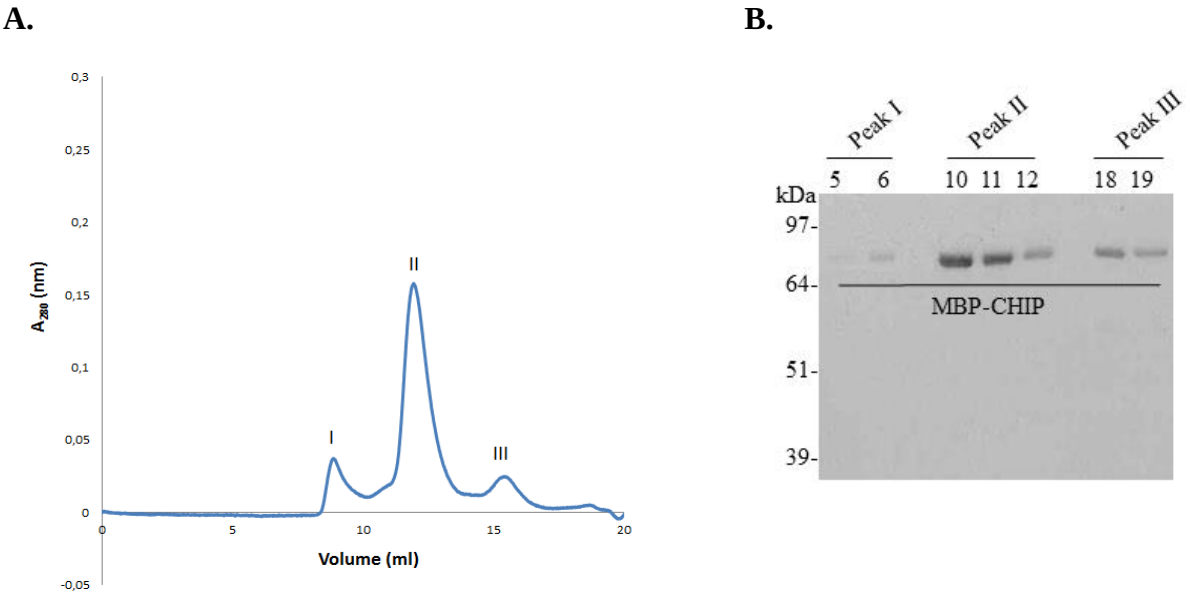
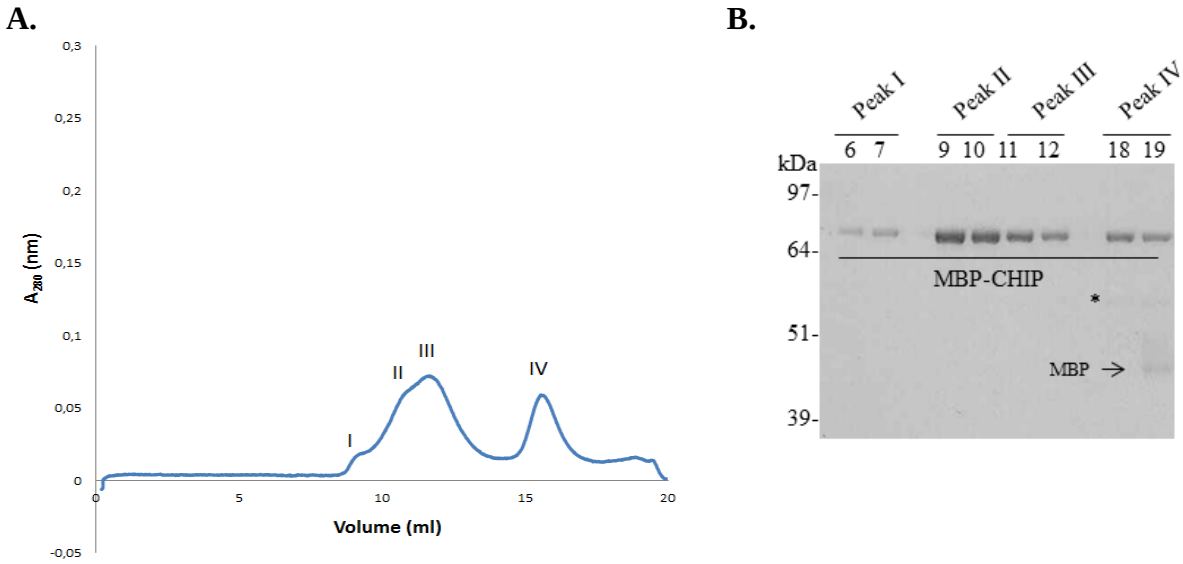
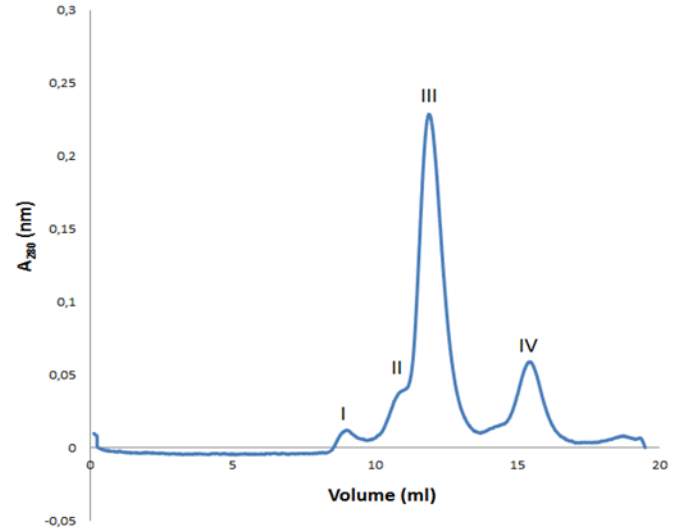


Figure 4.7.1 Size-exclusion chromatography of WT MBP-CHIP fusion protein. One mg of WT MBP fusion protein was subjected to size-exclusion chromatography on superdex 200 10/100 GL column at 0.3 ml/min at 4 °C. ~50 ml of phosphate buffer (20 mM NaH_2PO_4, pH 7.4) was used as the mobile phase while monitoring protein absorbance at 280 nm. A) Elution profile showing three peaks at 9, 12, and 15 ml. B) SDS-PAGE analysis of collected fractions showing the presence of MBP-CHIP fusion protein in Peak I (fractions 5, 6), II (fractions 10-12), and III (fractions 18, 19).

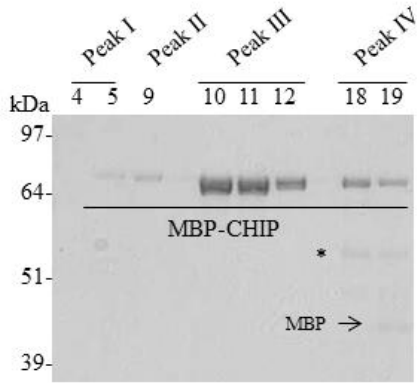
The process was repeated for all the mutants under the same experimental conditions. Figure 4.7.2 displays the elution profiles of mutant proteins as well as results from the SDS-PAGE analysis of fractions.



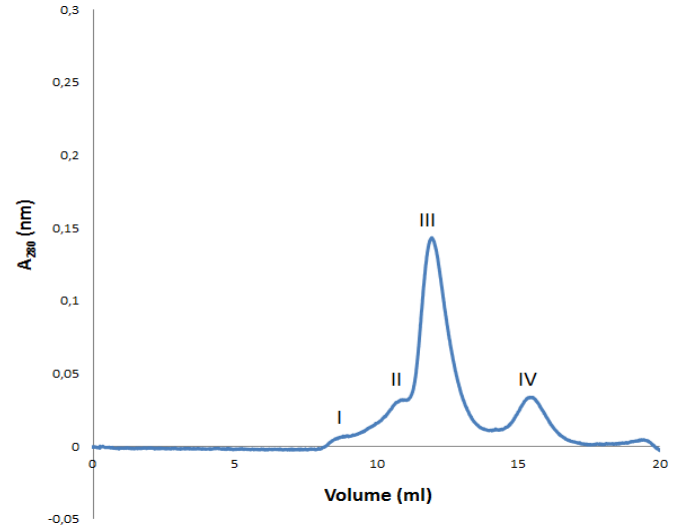
C.


0,3
0,25
0,2
0,15
0,1
0,05
0
-0,05
A280 (nm)
I
II
III
IV
0
5
10
15
20
Volume (ml)

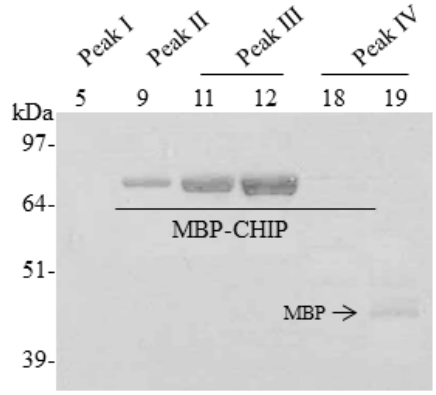
D.


Peak I
Peak II
Peak III
Peak IV
4 5 9 10 11 12 18 19
kDa
97-
64-
51-
39-
MBP-CHIP
*
MBP

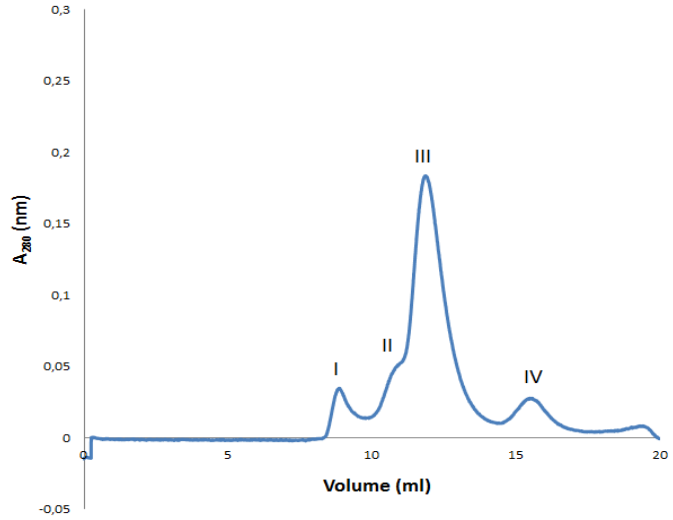
E.


0,3
0,25
0,2
0,15
0,1
0,05
0
-0,05
A280 (nm)
I
II
III
IV
0
5
10
15
20
Volume (ml)

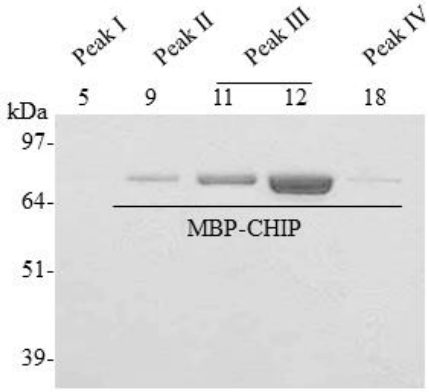
F.


Peak I
Peak II
Peak III
Peak IV
5 9 11 12 18 19
kDa
97-
64-
51-
39-
MBP-CHIP
MBP

G.


0,3
0,25
0,2
0,15
0,1
0,05
0
-0,05
A280 (nm)
I
II
III
IV
0
5
10
15
20
Volume (ml)

H.


Peak I
Peak II
Peak III
Peak IV
5 9 11 12 18
kDa
97-
64-
51-
39-
MBP-CHIP

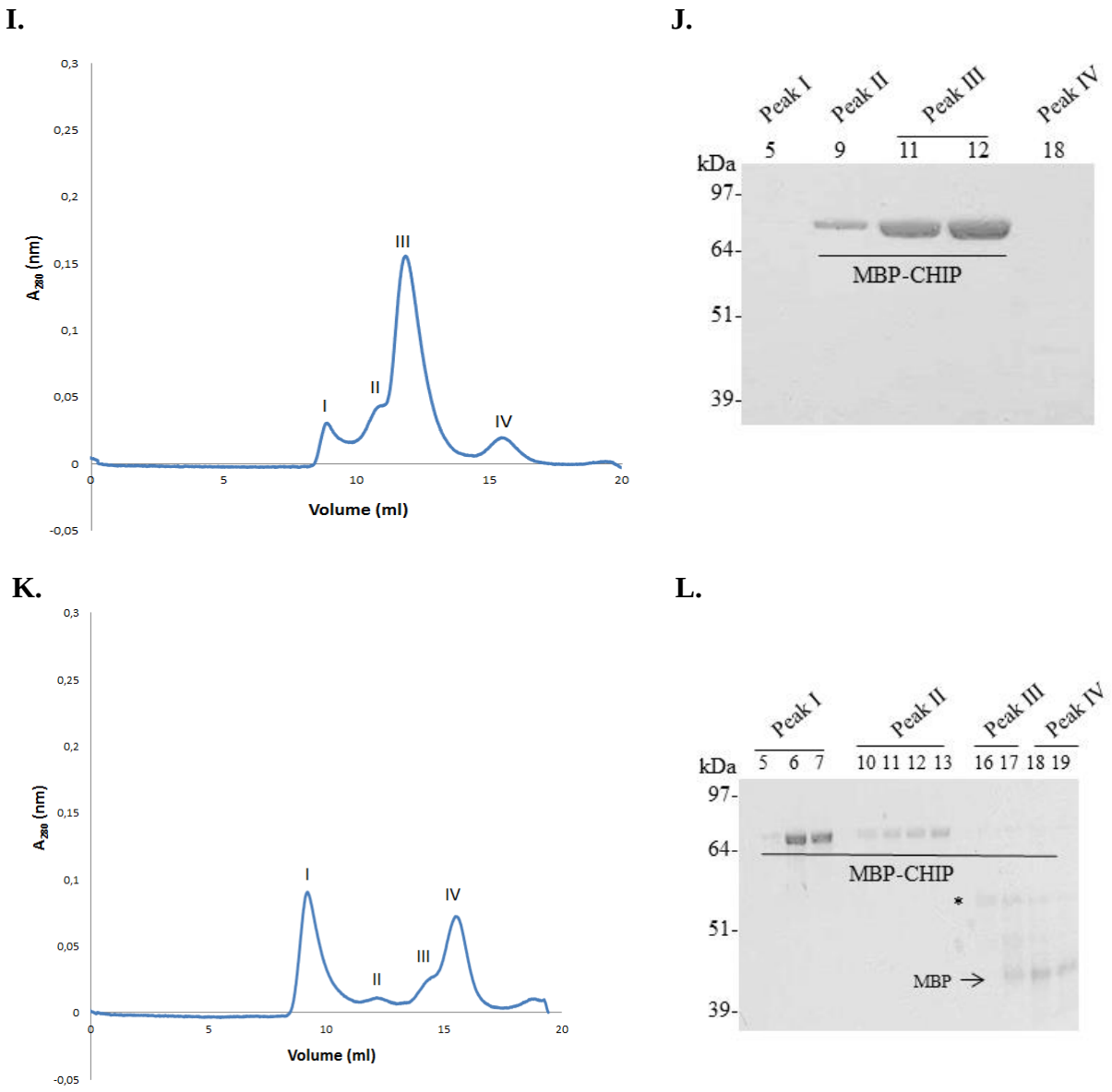


Figure 4.7.2 Size-exclusion chromatography of mutant recombinant MBP-CHIP fusion proteins. All the experiments were performed on one mg of MBP-CHIP at 0.3 ml/min at 4 °C, using Superdex 200 10/100 GL columns and ~50 ml of phosphate buffer (20 mM NaH_2PO_4, pH 7.4) as the mobile phase while monitoring protein absorbance at 280 nm. The elution profile and SDS-PAGE analysis of fractions corresponding to Peaks I-IV are shown for mutants E28K (A, B), N65S (C, D), K145Q (E, F), M211I (G, H), S236T (I, J), and T246M (K, L).

As shown for all the mutants, except for T246M, the oligomer and dimer peaks seemed to be connected with another peak at ~10 ml, which was speculated to be associated with trimeric structures (Figure 4.7.2 A, C, E, G, I, Peak II). The presence of MBP-CHIP fusion protein was confirmed by SDS-PAGE analysis and Coomassie blue staining in nearly all the fractions. In

addition, small amounts of cleaved MBP protein and one additional purified protein of ~55 kDa (marked with*) were detected together in fractions corresponding to the monomer peaks of E28K, N65S, and T246M mutants (Figure 4.7.2 B, D, L, Peak IV).

In the chromatogram of E28K, a smaller peak appeared at ~12 ml compared to what was seen for WT (Figure 4.7.2 A, Peak III, with UV absorbance of 0.07 nm vs. 0.16 nm), indicating the existence of a lower amount of dimeric structures for this mutant. In addition, trimeric structures seemed to be present at approximately the same level as the dimers as evidenced by a peak with UV absorbance of 0.06 nm, located at ~10 ml (Peak II). In contrast, for the N65S mutant, the dimer peak was higher compared to the WT (Figure 4.7.2 C, Peak III, UV absorbance of 0.23 nm vs. 0.16 nm). Results for K145Q, M211I, and S236T were similar to the WT (Figure 4.7.2 E-J). The K145Q mutant presented with a small oligomer peak at ~8 ml (Figure 4.7.2 E, Peak I). Moreover, the dimer peak was slightly smaller in this mutant as well as S236T compared to the WT (compare Peak III in Figure 4.7.1 A and 4.7.2 E, I). The presence of MBP-CHIP was not confirmed in the oligomer peaks of K145Q, M211I, and S236T as well as the monomer peak of S236T by SDS-PAGE and Coomassie blue staining (Figure 4.7.2 F, H, J, fractions corresponding to Peak I and IV) probably due to the small amount of protein. Finally, the T246M mutant was characterized by a remarkably increased amount of oligomeric structures, shown by a high oligomer peak at ~8 ml with UV absorbance of approximately 0.1 nm (Figure 4.7.2 K, Peak I,). This was continued by a very small dimer peak at ~11 ml (Peak II) and further by two connected peaks at ~13.5 ml (Peak III) and ~14.5 ml (Peak IV) in which small amounts of unidentified protein with ~55 kDa molecular weight and cleaved MBP were identified by SDS-PAGE analysis (Figure 4.7.2 L, fractions 16-19).

Altogether, these data suggested that the conformational structure of MBP-CHIP fusion protein is most prominently affected by mutations E28K and T246M, demonstrated by elevated trimer and oligomer peaks on their chromatograms. Moreover, N65S was identified to contain large amounts of dimeric structures, which was shown by a dimer peak higher than that of WT. Lastly, K145Q, M211I, and S236T displayed chromatographic patterns similar to the WT with only minor changes in the level of oligomeric and dimeric forms.

4.8 Oligomeric states of MBP-CHIP fusion proteins analyzed using native electrophoresis (native-PAGE)

Native polyacrylamide gel electrophoresis allows for a high-resolution analysis of the oligomeric state and molecular mass of native protein structure. This technique uses Coomassie G-250 as a charge-shift molecule which provides a net negative charge, and directs the electrophoretic mobility of proteins towards the anode while maintaining the proteins in their native conformation. The migration of 10 µg wild type and mutant MBP-CHIP proteins were determined on a native-PAGE gel as described in Section 3.15.3. Results are shown in Figure 4.8.1. The different migration patterns of CHIP protein band size (compared to 35 kDa observed in SDS-PAGE) is due to the fact that the mobility of proteins on a BN gel is influenced by both charge and size. Therefore, the molecular weight of proteins cannot be estimated with this method. Separation of five major protein bands in the size range of > 242 to >1048 kDa was observed for all the samples apart from T246M (Figure 4.8.1, lane 7) in which a higher-order oligomer band (≥1048 kDa) was the only visualized band, suggesting the formation of aggregates for this mutant.

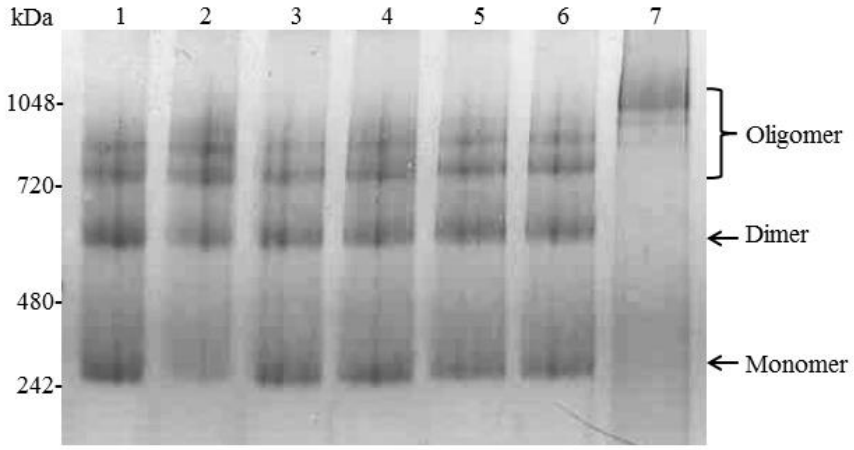


Figure 4.8.1 Oligomeric states of the WT and mutant MBP-CHIP by native gel electrophoresis. Ten µg of each WT (lane 1), E28K (lane 2), N65S (lane 3), K145Q (lane 4), M211I (lane 5), S236T (lane 6), and T246M (lane 7) MBP fusion protein was tested for migration on a 12% native-PAGE gel. Protein bands were visualized following Coomassie blue staining.

The E28K mutant (lane 2) showed lower amounts of monomers, as indicated by a faint band at ~242 kDa on the gel. In addition, the signal for dimers seemed to be slightly weaker for this mutant compared to the WT and other mutants. This indicates a small shift towards the lower-

order oligomers (720-1048 kDa) in the dimer-oligomer equilibrium of the protein. Overall, these results suggested that while the higher-order oligomers serve as the major conformational state of the T246M, other mutants displayed almost the same oligomeric patterns as the WT with small deviations observed for the E28K variant.

4.9 Secondary structures of WT and mutant MBP-CHIP fusion proteins measured by CD

Circular dichroism (CD) refers to the differential absorption of the left-handed and right-handed circularly polarized components of the plane polarized light, after passing through the subject. The CD spectrum is obtained when the dichroism is measured as a function of wavelength. For a protein sample, absorption in the far-UV region (180-250 nm) is related to the peptide bond. Therefore, different secondary structure composition of proteins leads to characteristic CD spectra in this region. Results from the far-UV CD spectra of both WT and mutant MBP fusion proteins showed two minima at 208 and 222 nm, which is typical of proteins with a high helical content (Figure 4.9.1) [68]. Different extents of CD signals ([Θ]) observed for the mutants indicate changes of the conformational properties of these proteins relative to WT.

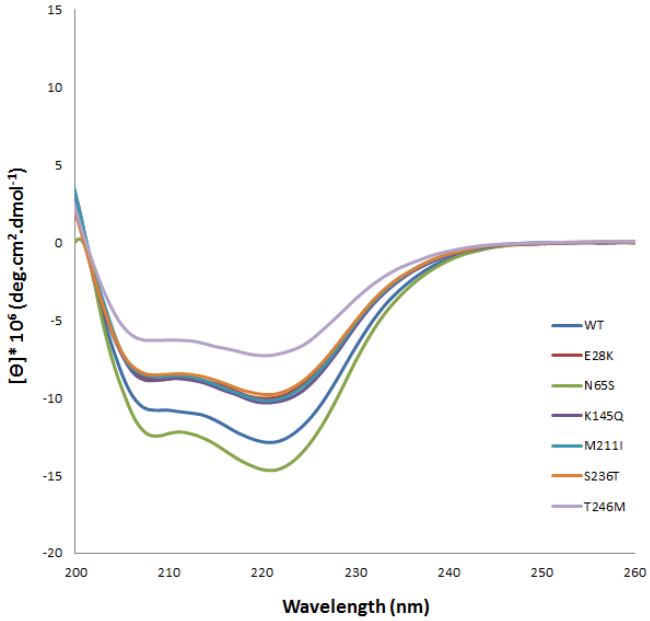


Figure 4.9.1 Far-UV CD spectra of WT and mutant MBP-CHIP proteins. CD spectra for WT (dark blue), E28K (red), N65S (green), K145Q (dark purple), M211I (light blue), S236T (orange), and T246M (light purple) are shown in units of molar ellipticity versus wavelength (nm). Protein samples were used at 6 µM concentration in a buffer containing 10 mM KH_2PO_4 (pH 7.4) and 100 mM NaF. The spectra were recorded in the range of 185-260 nm at 20 °C. The spectra shown were the average of four separate scans. Buffer scans were subtracted from the protein spectra, using the Spectra manager software. For a complete figure, see Appendix 7.1.

In order to provide an estimation of the secondary structure composition of WT and mutant CHIP proteins, the spectra corresponding to the CHIP portion were generated by subtraction of an "MBP-only" spectrum from the spectra of the MBP-CHIP fusion proteins as is shown for WT in Figure 4.9.2 A. The resulted CD spectra were further analyzed using CONTIN algorithm (Dichroweb software) [69, 70]. The distributions of α-helix and β-sheet are illustrated for each sample in Figure 4.9.2 B. Wild type protein was estimated to contain ~24% α-helices and ~10% β-sheet structures after being subtracted from MBP portion. Decreased percentages of α-helix (~22-23%) were calculated for those mutants locating at similar levels of ellipticity, but lower than the WT (E28K, K145Q, M211I, and S236T, Figure 4.9.1), and the largest loss of α-helicity (~11%) was observed for T246M mutant (Figure 4.9.1, light purple line). On the opposite side, N65S with ~30% α-helix and ~8% β-sheet content was identified as the only mutant with increased α-helical content compared to the WT (Figure 4.9.1, green line). Decreased α-helicity was reflected in an increased percentage of β-sheet in all cases.

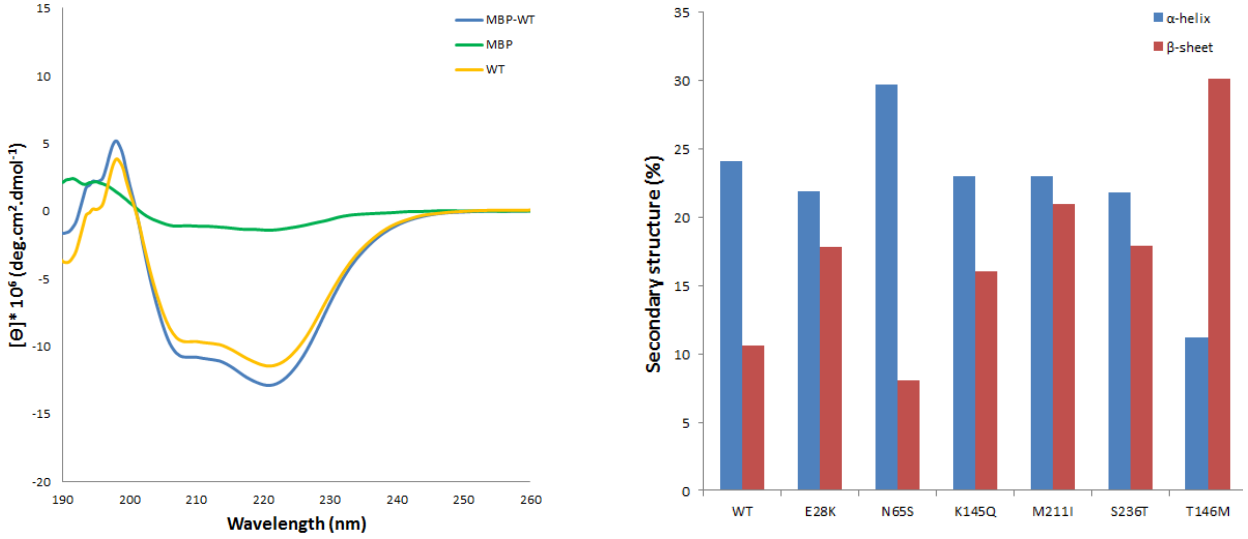


Figure 4.9.2 Secondary structure content calculation of WT and mutant CHIPs. Each of the WT and mutant CHIP spectra were subtracted from an 'MBP only' spectrum, and further analyzed for secondary structure content. A) Representative CD spectra for MBP-CHIP/ WT (blue), MBP (green) and the calculated WT CHIP (yellow) are shown. The WT CHIP spectrum was obtained by subtraction of the MBP spectrum from the MBP-CHIP/WT spectrum. B) Quantitative analysis of α-helix (blue) and β-sheet (red) percentage of WT and mutant CHIPs, performed on MBP-subtracted spectra by using CONTIN algorithm (Dichroweb software).

4.10 Conformational changes in CHIP structure as a function of temperature

The effects of mutations on the thermal stability of the CHIP protein were next examined through the investigation of changes in the molar ellipticity ([Θ]) of the WT and mutant MBP-CHIP fusion proteins over a temperature range of 20-90 °C, at 222 nm (Figure 4.10.1). The mid-point of the transitions (T_m) for denaturation of each sample is given in Table 4.10.1. Three transition temperatures in the range of ~41-45 °C, ~55-57 °C, and ~62-74 °C were monitored for denaturation of all protein samples excluding T246M which was appeared with two (T_m)s of 55.4 °C for the first transition and 66.8 °C for the second transition (Figure 4.10.1, light purple line and Table 4.10.1).

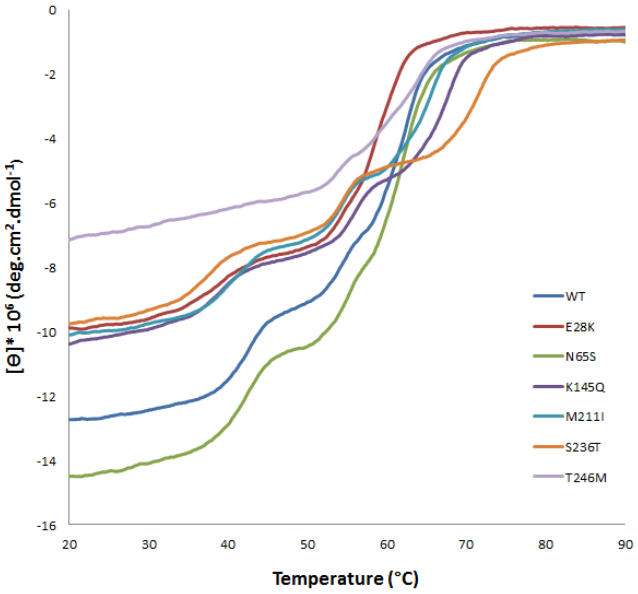


Figure 4.10.1 Thermal denaturation profiles of WT and mutant MBP-CHIP proteins. Thermal unfolding curves were obtained for WT (dark blue), E28K (red), N65S (green), K145Q (dark purple), M211I (light blue), S236T (orange), and T246M (light purple) by monitoring CD signal at 222 nm from 20 to 90 °C. Each curve was smoothed and analyzed for denaturation transitions by Spectra manager software. The curves are plotted in units of molar ellipticity versus temperature.

Table 4.10.1 Mid-point of transitions for thermal unfolding of MBP-CHIP proteins. Values are given by Spectra manager software in °C.

	WT	E28K	N65S	K145Q	M211I	S236T	T246M
T_m **(1)**	44.2	42.4	44.5	42.7	43.9	41.2	-
T_m **(2)**	55.8	55	56.2	57	56.2	56.2	55.4
T_m **(3)**	64.5	62.4	65.1	69.5	67.7	73.8	66.8

T246M was also shown as the only mutant with a broad, non-co-operative unfolding pattern, possibly indicating a loose, flexible tertiary protein structure. Other mutants presented unfolding

pathways similar to the WT and with more apparent transitions than that of T246M, which may point to loss of intermediate structures during their thermal denaturation [71].

In addition to the MBP-CHIP fusion proteins, the thermal denaturation of MBP protein itself was investigated by CD spectroscopy. As shown in Figure 4.10.2, a T_m of 56 °C was identified as the only unfolding transition for this protein, which nearly coincides with the second transition of MBP-fusion protein samples (compare with Table 4.10.1, T_m (2), small deviations (±1 °C) might be due to the binding of CHIP to MBP). Therefore, it is most likely that the first and third transitions on the thermal denaturation curves of MBP-CHIP proteins are associated with unfolding of CHIP protein itself (not MBP). It was further speculated that the first transition represents denaturation of the dimeric CHIP, while the third transition represents unfolding of the monomeric CHIP (see Section 5.2.4).

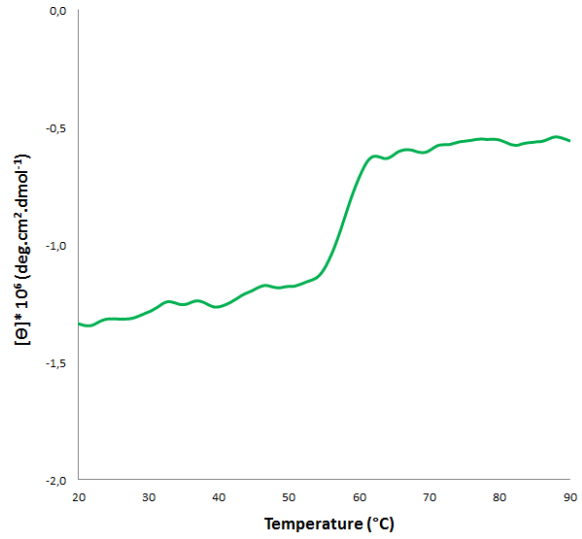


Figure 4.10.2 Thermal denaturation profile for MBP. CD signal was monitored at 222 nm for 6 µM of MBP in a buffer containing 10 mM KH_2PO_4 (pH 7.4) and 100 mM NaF over a temperature range of 20 to 90 °C with nitrogen flow rate of 5 l/min. Spectra manager software was used for smoothing the obtained curve as well as calculation of unfolding transition.

Considering the T_m of 44.2 °C as the dimeric unfolding transition of the WT CHIP, all of the mutants showed T_m values smaller than that of the WT (Table 4.10.1, T_m(1)), indicating decreased thermal stability of dimeric structures for these proteins. Hence, S236T with the dimeric T_m of 41.2 °C was regarded as the least stabilized protein in dimeric conformation (Figure 4.10.1, orange line and Table 4.10.1). Moreover, the absence of first transition on the thermal denaturation curve of T246M (light purple line) suggested a very small amount of dimeric structures for this protein.

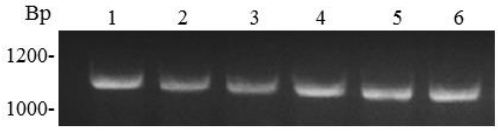
Thermal unfolding of mutant CHIP proteins seemed to occur more differently in monomeric phase than the dimeric phase, as can be seen by the melting transitions at a broader temperature range of ~62-74 °C than that of the dimeric transitions (Table 4.10.1, T_m(1) vs. T_m(3)). Compared to the WT with monomeric unfolding transition of 64.5 °C, E28K showed transition at lower temperature (62.4 °C), indicating a less conformational stability in monomeric state for this mutant. However, other mutants acquired more stabilize monomeric structures, displayed by T_m values larger than 64.5 °C. In this regard, S236T with T_m of 73.8 °C was presented as the protein with the highest thermal stability, followed sequentially by K145Q, M211I, T246M, and N65S.

For in cellulo protein studies

4.11 Generation of mutant *STUB1* constructs by site-directed mutagenesis

Mutations including p.K145Q (c.433A>C), p.M211I (c. 633G>A), and p.S236T (c. 707G>C) were introduced into WT *STUB1* cDNA in pcDNA3.1-V5/HisB plasmids, as described in Section 3.1 and 3.2.

A.



B.

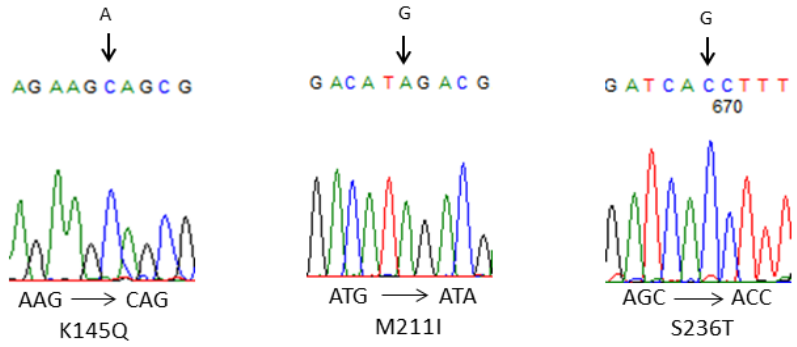


Figure 4.11.1 Generation of *STUB1* mutants by site-directed mutagenesis. The *STUB1* cDNA pcDNA 3.1-V5/HisB vector was used for *in cellulo* protein studies. A) Agarose gel showing the PCR amplification product from individual bacterial colonies transformed with potential K145Q (lane 1, 2), M211I (lane 3, 4), and S236T (lane 5, 6) mutant cDNA plasmids using T7 promoter and T7 terminator primers. B) cDNA sequencing of mutant plasmids directly from bacterial colonies, confirming the introduction of K145Q, M211I and S236T mutations analyzed by FinchTV software.

Successfully transformed colonies were selected and verified by PCR amplification and agarose gel electrophoresis (Figure 4.11.1 A). DNA sequencing was further performed on selected colonies in order to confirm the introduced mutations (Figure 4.11.1, B). Plasmids were purified, and verified for correct mutations once again with DNA sequencing of the whole *STUB1* cDNA insert (data not shown). Vectors containing E28K, N65S, and T246M mutations were prepared during previous studies performed by the group.

4.12 Effects of mutations on the exogenous CHIP protein expression level in HEK293 cells

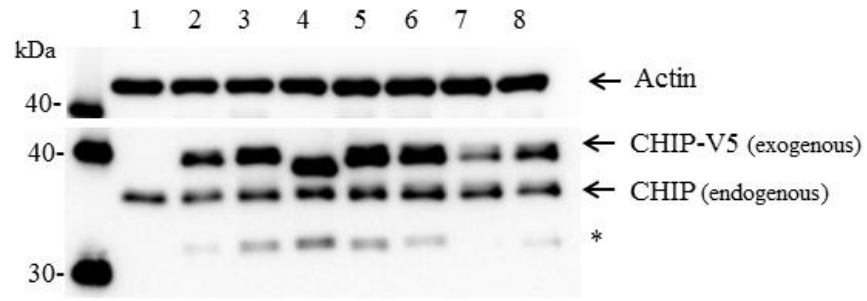
To investigate whether the expression level of exogenous CHIP protein was changed for the different mutant proteins, HEK293 cells were transiently transfected with the WT and mutant *STUB1* pcDNA3.1-V5/HisB plasmids. A pcDNA vector lacking the *STUB1* cDNA insert was used as a negative control. Cell lysates were analyzed by SDS-PAGE and immunoblotting using CHIP specific antibody. Equal amounts of proteins were loaded on the gel and actin bands were used for normalization during quantification. Results for the Western blotting analysis and protein expression levels are shown in Figure 4.12.1. The protein band corresponding to CHIP-V5/N65S (Figure 4.12.1 A, lane 4) migrates slightly faster on SDS-PAGE in comparison to the WT and other mutants probably because of protein conformational changes induced by this mutation.

The exogenous CHIP-V5 and endogenous CHIP proteins were identified at ~37 and ~35 kDa, respectively. Another band at the molecular weight of ~32 kDa was also detected on the membrane for all the proteins but not the negative control (Figure 4.12.1 A, marked with *), which might be due to the degradation of exogenous CHIP in cells. When normalized with actin, four mutants including E28K, N65S, K145Q, and M211I showed higher expression levels than the WT protein (Figure 4.12.1 A, lanes 3-6 and Figure 4.12.1 B). Among them, E28K with a relative protein level of ~122% had the most increased level of expression sequentially followed by K145Q, N65S, and M211I.

On the contrary, the two mutants S236T and T246M were found to be less expressed compared to the WT, shown by the protein levels of ~44% and ~87%, correspondingly. Taken together, these results indicate that *STUB1* mutations might affect the expression of exogenous CHIP protein *in-*

cellulo, by either an increase or decrease in expression level compared to WT. However, further experiments are needed before firm conclusions can be drawn.

A.



B.

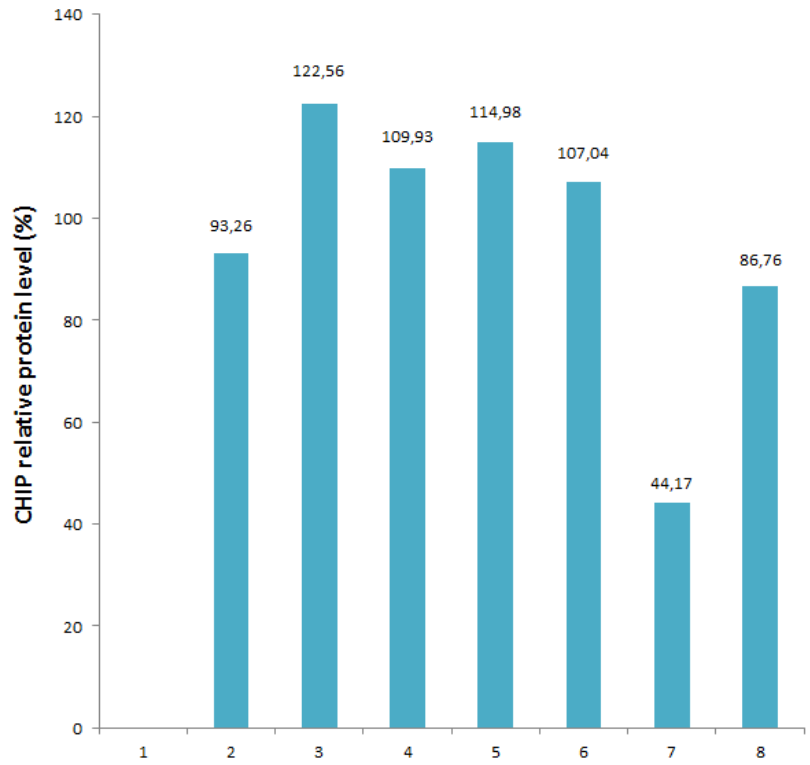


Figure 4.12.1 Expression of WT and mutant CHIP proteins in transiently transfected HEK293 cells. 10 µg of total lysate were analyzed for the level of exogenous CHIP-V5 protein expression after being transfected by the WT and mutant CHIP pcDNA 3.1-V5/HisB vectors. A) SDS-PAGE and immunoblotting analysis of cell lysates corresponding to empty vector (lane 1), WT (lane 2), E28K (lane 3), N65S (lane 4), K145Q (lane 5), M211I (lane 6), S236T (lane 7), and T246M (lane 8), using anti-CHIP antibody. Actin was used as an internal loading control. B) The level of CHIP-V5 proteins (normalized to actin) were quantified by MultiGauge software (n=1).

5. Discussion

Ataxia is a heterogeneous group of neurological disorders that involves degeneration or abnormal development of the cerebellum and spinal cord [51, 53]. Patients with ataxia are mainly characterized by walking, speech, and vision difficulties due to loss of coordination during muscle movements. Environmental factors as well as underlying medical conditions such as stroke, infection, or vitamin deficiencies are responsible for the acquired ataxia. Many ataxias, however, develop as genetic diseases with dominant, recessive, X-linked, or mitochondrial modes of inheritance. Inherited ataxias are estimated to include over 50 causative genes, and occur with a prevalence of 1-9 per 100,000 people [72].

Early attempts to discover the genetic origin of ataxia diseases were made in 1993 while CAG repeat expansions were identified as the common causative mechanism for many dominantly inherited ataxias including spinocerebellar ataxias (SCAs) [73]. Later, with the development of next-generation technology, several new genes were found to cause ataxia, including *ATP2B3, KCND3, DNMT1,* and *TPP1* [74, 75]. Improved understanding of ataxia genes has led to identification of more disease pathways. For instance, many ataxia genes are involved in the ubiquitin-proteasome pathway. Mutations in the components of this pathway cause the accumulation of misfolded and damaged proteins inside the cell which, consequently, result in toxicity and cell death. An example is E3 ubiquitin ligase RNF170 which is found to be mutated in sensory ataxia. Another E3 ligase, parkin, interacts with deubiquitinating enzyme Ataxin-3. Mutated ataxin-3 results in more deubiquitinated parkin that accumulates and causes development of Parkinson disease (PD). Recently, mutations in the *STUB1* gene encoding CHIP E3 ligase have been found to be associated with ARCA and hypogonadism [72].

In the present study, six different *STUB1* mutations including E28K, N65S, K145Q, M211I, S236T, and T246M were characterized in terms of (1) ubiquitination activity, (2) structural/conformational properties, and (3) exogenous expression of the encoded CHIP protein using *in vitro/in cellulo* methods. These mutations were identified and reported by exome sequencing studies on individual families with SCAR 16 disease [30, 49, 57-63]. Some, including E28K, N65S, and T246M have been described previously [30, 61]. The selection of the three other mutations (K145Q, M211I, and S236T) was based on molecular analysis of the CHIP

protein structure by PyMol software, since the amino acid substitutions seem to be situated in crucial regions for CHIP dimerization or protein-protein interactions.

5.1 The effects of different *STUB1* mutations on the E3 ubiquitin ligase activity of CHIP

The *in vitro* ubiquitination activity assay showed reduced levels of Hsc70-ubiquitination activity for CHIP-N65S (Figure 4.3.1). This mutation is located in the TPR domain affecting a residue that was previously reported to be involved in CHIP substrate binding [76]. Considering the intact auto-ubiquitination activity of the N65S mutant, the mutation is therefore predicted to impact CHIP substrate binding affinity and reduces its ability to ubiquitinate Hsc70. Another mutation in the same domain (E28K) resulted in a similar ubiquitination activity as the WT for both Hsc70 substrate and itself (auto-ubiquitination), indicating that the substrate binding of CHIP was not affected by this mutation. These assumptions were further supported by the microscale thermophoresis (MST) analysis of the E28K and N65S mutants which was performed during MST demo day by using a Monolith NT.115 instrument (see Appendix 7.2). In this method, the biomolecular interactions can be measured precisely and within a short time by detecting changes in the hydrogen shell, charge or size of molecules while passing through microscopic temperature gradients [77, 78]. MST analysis of CHIP vs Hsp70 displayed no interaction for N65S, while a strong interaction (with a K_d of 9 nM) was observed for E28K. In addition, a study by Heimdal *et al.* on ubiquitination activity levels of N65S and E28K variants showed similar results for both Hsc70- and self-ubiquitination. However, the co-immunoprecipitation assays failed to firmly demonstrate the impaired binding between N65S and Hsc70 [30].

A complete loss of ubiquitination activity was observed for CHIP-T246M for both Hsc70 substrate and CHIP itself, indicating the deleterious effect of this mutation on CHIP's ubiquitin ligase activity. In fact, the T246M substitution (located at the U-box domain of CHIP) can affect its interactions with the ubiquitin-proteasome pathway compartments and result in abolished ubiquitination activity of the protein. This idea was also suggested by Shi *et al*., who discovered the T246M mutation and tested its ubiquitination activity as well as interactions with Hsp70 using both cell culture models and purified recombinant proteins [61]. They found an increased

interaction between T246M and Hsp70 compared to the WT. However, CHIP-T246M was not able to ubiquitinate Hsp70 as well as itself. Thus, the lack of chaperone activity for this mutant was suggested to be due to a defect in the U-box domain rather than the inability to bind to chaperone substrate.

Reduced substrate affinity of CHIP-N65S and the loss of function of CHIP-T246M are assumed to contribute to the development of ARCA disease through accumulation of misfolded client proteins and increased potential for toxic aggregation in cells of homozygous patients carrying these mutations. Patients affected by the T246M mutation are reported to develop hypogonadism together with ataxia (Gordon Holmes Syndrome) which was not found in cases carrying the other *STUB1* mutations [61]. Further studies are required to demonstrate whether the location of the mutation is critical for the development of hypogonadism.

Results from the ubiquitination activity of K145Q, M211I, and S236T mutants revealed similar levels of function as compared to the WT. Other studies reporting the activity of these mutations were few, and limited to bioinformatics predictions [58, 60, 62]. These three mutations together with E28K exist as compound heterozygous with either a non-sense or deletion mutation in the genome of related ARCA patients [30, 58, 60, 62]. Therefore, the development of ARCA in these families is likely to be more related to the effect of second alleles [30].

5.2 Structural and conformational characterization of WT and mutant CHIP proteins

5.2.1 Oligomerization studies discovered high aggregation propensity for CHIP-T246M

Size-exclusion chromatography of MBP-CHIP proteins displayed a high oligomeric peak on the elution profile of the mutant T246M which did not appear for the WT and the rest of the mutants, pointing to a conformational change in the structure of this mutant that leads to an induced tendency to form aggregates (Figure 4.7.2 K, L). A study on the native CHIP homologues from *Homo sapiens* (hCHIP) and *D.melanogaster* (dCHIP) indicated the necessity of dimerization for activity of hCHIP by performing ubiquitination assays on full-length hCHIP as well as deletion mutants lacking the putative coiled-coil domain [36]. Another study analyzing the crystal structure of Prp19 U-box suggested dimerization as a common architecture for the U-box and

RING-finger families of E3 ubiquitin ligases, which play essential roles in the stability and functionality of these enzymes [79]. Therefore, the complete loss of function observed in the *in vitro* ubiquitination activity assay for the CHIP-T246M mutant might be caused by both impaired ubiquitin ligase activity (as a result of impaired interaction with ubiquitination enzymes) and formation of oligomeric protein structures. In addition to T246M, E28K presented higher trimeric/tetrameric structures compared to the WT and other mutants (Figure 4.7.2 A, B). The level of structural modifications imposed by the mutation, however, might not be sufficiently severe to affect the *in vitro* ubiquitination activity as this appeared to be similar to that of the WT. Moreover, increased amount of dimeric structures were noticed for N65S through observation of a dimer peak higher than that of the WT (Figure 4.7.2 C, D). However, since the mutation is not located in the coiled-coil domain of CHIP, this finding cannot be considered as a direct effect of the location of mutation on protein dimerization.

These findings are in agreement with native gel electrophoresis results; a higher-order oligomer band was detected for T246M, indicating the formation of aggregates for this mutant, and a small shift towards the lower-order oligomeric structures (trimer/tetramer bands) was discovered for E28K (Figure 4.8.1). Small oligomerization changes occuring in the structure of N65S (and other CHIP mutants) were detected by gel-filtration analysis, while native-PAGE methodology was not able to show these changes possibly due to a lower resolution of this method *per se*. Native-PAGE is generally considered a reliable method for determination of the oligomeric states of proteins, but, special conditions applied to this technique such as the bound Coomassie blue dye can affect the folding of the proteins, and lead to a small decrease in the accuracy of the method [80]. Nevertheless, the rapidity and relative simplicity of this approach allows for its general use in the investigation of molecular mass and oligomeric states of proteins. Several studies have reported the analysis of membrane protein complexes [81, 82] as well as neurodegeneration-linked protein aggregates such as parkin and α-synuclein by using this technique [83, 84].

It is important to mention that these observations could be influenced by the presence of MBP attached to CHIP during the experiments. But, due to similar conditions applied to all chromatography procedures, the effect of MBP on CHIP oligomerization properties is considered to be the same for each of the protein samples, making the obtained results scientifically comparable.

5.2.2 Mutants T246M and N65S displayed remarkable differences in terms of secondary structure content

Crystallographic analysis of mammalian (mouse) CHIP indicates a high degree of α-helicity in the structure of this protein due to the presence of two domains consisting mainly of α-helices (TPR and coiled-coiled domains) [37]. Our far-UV spectroscopy data of the wild type CHIP displayed an α-helical content that corresponds well with the crystallography data (Figure 4.9.1). Reduced levels of elipticity were, however, observed in the spectra of the majority of mutants, suggesting a decreased level of α-helical structure content of these proteins compared to the WT. The largest loss of α-helicity was shown by the T246M mutant (by ~13%) with the mutation located in the second β hairpin of the U-box domain. Given the fact that methionine (the amino acid for which threonine is exchanged in the T246M mutation) has a high helix-forming propensity [85], it is not clear why this mutation results in such a substantial decrease in α-helicity. Nevertheless, it might be possible that the substitution of threonine as a polar amino acid by methionine as a non-polar amino acid disrupts the overall non-covalent polar interactions, from which the stabilizing energy of a fold comes from, and therefore causes some reductions in the level of α-helical structure content of the protein [86]. The loss of α-helix structures was associated with increased percentage of β-sheet in all cases. Mutants with increased β-sheet content have been proposed to show increased propensity for protein aggregation as a consequence of having exposed or extended β-sheet edges which are normally buried inside their native structure [71]. Accordingly, the aggregation behavior discovered by oligomerization studies of CHIP-T246M can be related to the significant increase in β-sheet content of this mutant (by ~20%). Additional studies on the structure and interactions of T246M are required to support this notion.

On the contrary, being located at the beginning of three antiparallel α-helices in the TPR domain of CHIP, the N65S mutation generated a protein structure with considerably increased secondary structure content of α-helices (by ~7%). However, this observation does not correlate with the 'helix breaker' role of serine as the exchanged amino acid in this mutation [87]. N65S was also identified with a larger amount of dimeric structures during gel-filtration analysis. But the relationship between increased α-helicity and dimerization still remains to be understood.

Although each of the MBP-CHIP protein spectra were subtracted from an "MBP only" spectrum

during data analysis (Figure 4.9.2), it is important to consider limitations associated with the use of MBP fusion proteins while monitoring structural characteristics. The overall protein folding and assembly may be different for MBP-CHIP fusion proteins compared to purified CHIP. In addition, mutations causing misfolding and aggregation (T246M, for example) may have a less severe effect in the presence of MBP as a protein with solubilizing activity [71]. In spite of these issues, two studies have reported the secondary structure of purified CHIP with a similar high degree of α-helicity as was found in this study [36, 41]. In addition, regarding the fact that these experiments were performed under the same conditions with the aim of comparing WT and mutant proteins, the findings are believed to be useful and scientifically reliable.

5.2.3 CHIP-N65S: the only mutant with increased structural stability against limited proteolysis

Limited proteolysis provides a general method to unravel molecular features as well as structure-function relationships of many proteins. Numerous studies have successfully uncovered protein folding pathways [88, 89] and identified optimal folding conditions for novel proteins [90-92] by using this approach. Protein conformational features predicted by limited proteolysis often correlate with results from other biophysical and spectroscopic methods such as circular dichroism and nuclear magnetic resonance (NMR) spectroscopy [93].

The sites of proteolysis across a protein surface are very few and restricted to areas characterized by enhanced backbone flexibility (segmental mobility). Therefore limited proteolysis of a globular protein occurs mostly at flexible loops, and regular secondary structures (such as helices) are not subjected to cleavage [93]. Studies on the proteolysis of a variety of proteins with known 3D structures supported this idea, indicating that cleavage mainly occurs at the site of loops, but never at α-helices [92, 94].

Taking this information into account, decreased susceptibility identified for CHIP-N65S against limited proteolysis is expected to be associated with a structure that contains a larger number of α-helices (compared to the WT) which is confirmed by far-UV CD spectra of the mutant (Figure 4.4.1 and 4.9.1). The induced stability and proteolytic resistance in the structure of N65S may also be the consequence of more dimeric states discovered during gel-filtration analysis of this protein. However, more investigation is required in order to clarify the relationship between these two findings.

Limited proteolysis monitors the process of protein aggregation via the presence of cross-β structures which are resistant to proteolytic degradation. However, the unfolded and (or) flexible states of protein made in the early stages of aggregation cause an enhanced susceptibility to proteolysis [93]. Considering the high levels of oligomerization observed during size-exclusion chromatography of CHIP-T246M, the initial formation of protein aggregates is suggested to make this mutant more susceptible than the other variants towards limited proteolysis.

Mutants including K145Q, M211I, and S236T displayed approximately the same extent of susceptibility towards proteolysis, being lower than that of T246M but still higher compared to the WT. These proteins are suggested to achieve some loose, flexible structures which make them more sensitive towards proteolytic cleavage. Likewise, results from far-UV spectroscopy of these mutants (together with E28K) showed similar levels of ellipticity, but lower than the WT spectrum, indicating that these mutants have lost the same amount of α-helical secondary structure. Therefore, a decreased number of α-helices in the structure of K145Q, M211I, and S236T mutants could be the reason behind their reduced stability against limited proteolysis. In contrast, CHIP-E28K presented approximately the same (slightly lower) level of susceptibility as T246M towards limited proteolysis. The high amount of trimeric/tetrameric structures identified for this protein during gel-filtration analysis could indicate that the high proteolytic susceptibility of E28K is due to the unfolded/flexible structures of the protein molecule that appear during the formation of lower-order oligomers.

5.2.4 Circular dichroism revealed new insights into the conformational dynamics and thermal stability of CHIP protein mutants

The thermodynamic stability of protein folding, i.e. the Gibbs free energy difference between folded and unfolded states (ΔG), is generally determined by thermal unfolding spectroscopy whereby a protein solution is heated at a constant rate, and changes in its conformation are reported as melting temperatures (T_m) of the unfolding response [95, 96]. Denaturation of small globular proteins generally follows a two-state mechanism involving a single unfolding transition (T_m), and two forms of fully native (N) and unfolded (U) proteins [97, 98]. However, many proteins have been recently observed to stabilize intermediates between the N and U states, and therefore show more than one transition during their unfolding profile [98, 99]. This behavior is

usually found in multi-domain/multimeric proteins where different domains/meres unfold independently and at different temperatures. Analysis of data obtained from such curves is often more complicated than that of a single-stage unfolding pattern [95].

In the case of MBP-CHIP, the three transitions detected during thermal denaturation of the WT protein are assumed to be associated with unfolding of (in order) dimeric CHIP, MBP, and monomeric CHIP (Figure 4.10.1). As a dimeric protein, CHIP structure is stabilized by the intermolecular forces which hold proteins together as a dimer as well as intramolecular forces that occur within the monomers [100]. Therefore, it is possible that the disruption of intermolecular interactions during the first transition temperature (~44 °C) of MBP-CHIP proteins generates both dissociated MBPs and monomeric CHIPs in the protein solution being examined. After that, increased temperatures result in unfolding of each MBP and CHIP through disruption of interactions within the proteins (intramolecular interactions). Following the examination of MBP (only) for thermal denaturation (Figure 4.10.2), and based on other CD studies done on the unfolding transition of this protein [71, 101], the second transition on the unfolding curve of MBP-CHIP (~56 °C) is thought to belong to the melting of MBP. Therefore, CHIP itself (monomer) is expected to unfold at ~64 °C corresponding to the last transition temperature.

The majority of mutants presented unfolding pathways with a similar level of cooperativity as the WT, indicating that the whole conformation and dynamics of the protein was not largely affected by the mutations. Differences, instead, appear to lie between the dimeric and monomeric unfolding transition temperatures determining the thermal stability of mutant protein structures. In the dimeric state, N65S was identified with a similar stability as the WT but higher than other mutants (T_m of 44.5 °C). The discovery of more dimers during gel-filtration analysis of this mutant can possibly explain this observation. Moreover, N65S was presented as the only mutant with an increased level of α-helicity and stability against limited proteolysis. Taken together, these findings indicate a stabilizing effect for N65S mutation, resulting in a more condensed CHIP protein structure.

However, as a monomer, the highest thermal stability was displayed by S236T (T_m of ~74 °C). This observation is not correlated with the loss of secondary structure content shown by the far-

UV spectrum of the protein. Interestingly, at the same time, this mutant was associated with the least stabilized dimeric structure (T_m of ~41 °C). Therefore, a dual effect was observed for the S236T mutation, resulting in two completely different levels of stability for dimeric and monomeric forms of the protein. This might indicate that S236T does not form stable dimers and can be functional as a monomer, although such a hypothesis is not supported by gel-filtration analysis results where S236T displayed a dimeric peak similar to the WT. Further investigations are required in order to elucidate these findings precisely.

Denaturation of the T246M mutants was characterized by less cooperative melting transitions, which may point to a loose, flexible tertiary structure. Results from limited proteolysis support this idea, indicating a high level of flexibility in the structure of this mutant as a result of initial formation of aggregations. The large loss of secondary structure content observed during the far-UV spectroscopy of T246M can also explain the flexibility of its protein structure. In addition, the first transition corresponding to denaturation of dimeric intermediates was not found in the unfolding pathway of T256M. This observation correlates nicely with the gel-filtration analysis results where no peak associated with dimers was detected on the chromatogram of the T246M variant. In summary, the T246M mutation is likely to cause a high tendency for oligomerization in the CHIP protein structure which can be responsible for the lack of activity in corresponding mutants. However, on the way to the formation of aggregates, CHIP achieves a loose and flexible structure characterized by reduced stability toward proteolysis and a non-cooperative thermal unfolding profile.

5.3 Expression levels of exogenous CHIP under the effect of mutations

A pilot experiment was performed to study the effect of *STUB 1* mutations on the expression levels of exogenous CHIP in HEK293 cells. The results from this initial analysis indicated an increased level of expression of E28K, N65S, K145Q, and M211I mutants (Figure 4.12.1). These observations do not follow the *in vitro* analysis of CHIP mutants, where a stabilized protein structure was detected only for N65S, and other mutations were associated with reduced structural resistance and secondary structure content. Furthermore, the immunoblot analysis of fibroblasts showed decreased levels of steady-state CHIP in patients carrying E28K, N65S and M211I mutations [30, 58]. On the contrary, lower levels of exogenous S236T and T246M CHIPs

were detected compared to the WT in this study. *In vitro* analysis of T246M showed a high level of structural flexibility and aggregation for this mutant, while S236M mutation caused slighter changes in the structure of CHIP compared to T246M. No studies are available yet reporting the expression of these mutants in fibroblasts from patients.

These findings must be considered preliminary before they have been supported by further experiments. In fact, due to the lack of transcription control (which verifies that the same levels of RNA are produced for all the mutants), false results can be achieved as a result of differences in transfection efficiencies. Furthermore, more repetitions of the experiment are necessary to obtain statistically valid data, which unfortunately was not possible within the limited time-frame of this project.

5.4 Conclusion

This study through examining the activity and structural properties of the encoded CHIP protein mutants provided further evidence regarding the role of *STUB 1* mutations in the development of ARCA disease. The main results are summarized in Table 5.4.1. The ubiquitination activity of CHIP was impaired under the effect of mutations N65S and T246M while other mutants showed intact activities. Increased amounts of dimers as well as higher levels of secondary structures were discovered for CHIP-N65S, possibly resulting in the higher stability observed against limited proteoly°°sis and thermal unfolding. In contrast, T246M mutation generated a flexible protein structure with decreased α-helicity and a high tendency for aggregation. The CHIP-T246M mutant presented the lowest stability against limited proteolysis. No dimeric unfolding transition was detected during their thermal denaturation. However, as a monomer, thermal stability of these mutants was found to be greater than the WT. Other mutations including K145Q, M211I, and S236T modified CHIP structure and conformation to a similar degree. Decreased secondary structure content and proteolytic stability was observed for these variants. Changes were detected to a lesser extent than those caused by T246M mutation. In addition, proteins became less resistant in terms of thermal unfolding while existing as dimers. However, different results were obtained regarding the stability of monomers against thermal denaturation. Finally, the expression levels of exogenous CHIP were increased in HEK293 cells for all the mutants excluding S236T and T246M among which, S236T was associated with the lowest degree of expression (reduced by 50%).

Table 5.4.1 Summary of data achieved during this master project. Results for ubiquitination activity, structural/conformational characterizations, and exogenous expression levels of CHIP protein mutants are presented together with the WT. For each mutant, upwards and downwards pointing arrows indicate increased and decreased levels when compared to the WT, respectively. No dimeric unfolding was detected in the case of T246M.

	WT	E28K	N65S	K145Q	M211I	S236T	T246M
Ubiquitination activity	OK	OK	Impaired	OK	OK	OK	Impaired
Dominant oligomeric state	Dimer	trimer/ tetramer	Dimer	Dimer	Dimer	Dimer	Oligomer
Secondary structure content	—	↓	↑	↓	↓	↓	↓
Proteolytic stability	—	↓	↑	↓	↓	↓	↓
Thermal unfolding stability – Dimeric unfolding	—	↓	↑	↓	↓	↓	
Thermal unfolding stability – Monomeric unfolding	—	↓	↑	↑	↑	↑	↑
Exogenous expression level	—	↑	↑	↑	↑	↓	↓

5.5 Future perspectives

Future studies should focus on the *in cellulo* characterization of *STUB1* mutations using more relevant cell lines such as neuronal cell lines or induced Pluripotent Stem cells (iPSCs). Using this approach, major findings achieved during this study including high aggregation propensity of T246M and structural stability of N65S can be analyzed through different fluorescence techniques inside the cells. Moreover, study of interactions between CHIP and other co- As mentioned earlier, there are currently 18 *STUB1* mutations identified by exome sequencing in the genome of patients with ARCA disease. The number of these mutations is increasing rapidly, making the experimental measurement of the functional impact of each variant require much more time cost and often impractical. Therefore, massively parallel functional analysis of *STUB1*

mutations, which will provide a high-throughput prospective interpretation of all variants within a short time, will be of great future interest. The potential of this new approach has been recently demonstrated in a study where nearly 2000 variants of the breast cancer-associated gene *BRCA1* were characterized using this technique [102].

Finally, given the fact that protein aggregation is the most common phenotype of many complex diseases, results obtained from ARCA cases can be successfully used in the study of Alzheimer, Parkinson and many other protein aggregation diseases as well. For instance, recent studies indicated that CHIP cross-functions with Parkin E3 ligase as the main cause of hereditary Parkinson disease [103]. Thus, functional analysis of Parkin in association with *STUB1* mutations may in the future provide a better understanding of these diseases.

6. References

1. Gidalevitz, T., V. Prahlad, and R.I. Morimoto, *The stress of protein misfolding: from single cells to multicellular organisms.* Cold Spring Harb Perspect Biol, 2011. **3**(6).
2. Murata, S., T. Chiba, and K. Tanaka, *CHIP: a quality-control E3 ligase collaborating with molecular chaperones.* Int J Biochem Cell Biol, 2003. **35**(5): p. 572-8.
3. Marques, C., et al., *The triage of damaged proteins: degradation by the ubiquitin-proteasome pathway or repair by molecular chaperones.* FASEB J, 2006. **20**(6): p. 741-3.
4. Imai, J., et al., *Proteasomes and molecular chaperones: cellular machinery responsible for folding and destruction of unfolded proteins.* Cell Cycle, 2003. **2**(6): p. 585-90.
5. Schneider, C., et al., *Pharmacologic shifting of a balance between protein refolding and degradation mediated by Hsp90.* Proc Natl Acad Sci U S A, 1996. **93**(25): p. 14536-41.
6. Wickner, S., M.R. Maurizi, and S. Gottesman, *Posttranslational quality control: folding, refolding, and degrading proteins.* Science, 1999. **286**(5446): p. 1888-93.
7. Young, J.C., et al., *Pathways of chaperone-mediated protein folding in the cytosol.* Nat Rev Mol Cell Biol, 2004. **5**(10): p. 781-791.
8. Mayer, M.P. and B. Bukau, *Hsp70 chaperones: cellular functions and molecular mechanism.* Cell Mol Life Sci, 2005. **62**(6): p. 670-84.
9. Tutar, Y., Y. Song, and D.C. Masison, *Primate Chaperones Hsc70 (Constitutive) and Hsp70 (Induced) Differ Functionally in Supporting Growth and Prion Propagation in Saccharomyces cerevisiae.* Genetics, 2006. **172**(2): p. 851-861.
10. Kabani, M. and C.N. Martineau, *Multiple hsp70 isoforms in the eukaryotic cytosol: mere redundancy or functional specificity?* Curr Genomics, 2008. **9**(5): p. 338-248.
11. Boorstein, W.R., T. Ziegelhoffer, and E.A. Craig, *Molecular evolution of the HSP70 multigene family.* Journal of Molecular Evolution. **38**(1): p. 1-17.
12. Brocchieri, L., E. Conway Macario, and A.J. Macario, *hsp70 genes in the human genome: Conservation and differentiation patterns predict a wide array of overlapping and specialized functions.* BMC Evolutionary Biology, 2008. **8**(1): p. 1-20.
13. Daugaard, M., M. Rohde, and M. Jaattela, *The heat shock protein 70 family: Highly homologous proteins with overlapping and distinct functions.* FEBS Lett, 2007. **581**(19): p. 3702-10.
14. Zhao, R. and W.A. Houry, *Hsp90: a chaperone for protein folding and gene regulation.* Biochem Cell Biol, 2005. **83**(6): p. 703-10.
15. Li, J. and J. Buchner, *Structure, function and regulation of the hsp90 machinery.* Biomed J, 2013. **36**(3): p. 106-17.
16. Caplan, A.J., *What is a co-chaperone?* Cell Stress & Chaperones, 2003. **8**(2): p. 105-107.
17. Ballinger, C.A., et al., *Identification of CHIP, a novel tetratricopeptide repeat-containing protein that interacts with heat shock proteins and negatively regulates chaperone functions.* Mol Cell Biol, 1999. **19**(6): p. 4535-45.
18. Goldberg, A.L., *Protein degradation and protection against misfolded or damaged proteins.* Nature, 2003. **426**(6968): p. 895-9.
19. Pickart, C.M., *Mechanisms underlying ubiquitination.* Annu Rev Biochem, 2001. **70**: p. 503-33.
20. Koegl, M., et al., *A Novel Ubiquitination Factor, E4, Is Involved in Multiubiquitin Chain Assembly.* Cell, 1999. **96**(5): p. 635-644.
21. Hatakeyama, S. and K.I. Nakayama, *U-box proteins as a new family of ubiquitin ligases.* Biochem Biophys Res Commun, 2003. **302**(4): p. 635-45.
22. Aravind, L. and E.V. Koonin, *The U box is a modified RING finger - a common domain in ubiquitination.* Curr Biol, 2000. **10**(4): p. R132-4.

23. Welchman, R.L., C. Gordon, and R.J. Mayer, *Ubiquitin and ubiquitin-like proteins as multifunctional signals.* Nat Rev Mol Cell Biol, 2005. **6**(8): p. 599-609.
24. Bedford, L., et al., *Assembly, structure, and function of the 26S proteasome.* Trends in Cell Biology, 2010. **20**(7): p. 391-401.
25. Voges, D., P. Zwickl, and W. Baumeister, *The 26S proteasome: a molecular machine designed for controlled proteolysis.* Annu Rev Biochem, 1999. **68**: p. 1015-68.
26. Cyr, D.M., J. Hohfeld, and C. Patterson, *Protein quality control: U-box-containing E3 ubiquitin ligases join the fold.* Trends Biochem Sci, 2002. **27**(7): p. 368-75.
27. Hohfeld, J., D.M. Cyr, and C. Patterson, *From the cradle to the grave: molecular chaperones that may choose between folding and degradation.* EMBO Rep, 2001. **2**(10): p. 885-90.
28. Murata, S., et al., *CHIP is a chaperone-dependent E3 ligase that ubiquitylates unfolded protein.* EMBO Rep, 2001. **2**(12): p. 1133-8.
29. Rosser, M.F., et al., *Chaperone functions of the E3 ubiquitin ligase CHIP.* J Biol Chem, 2007. **282**(31): p. 22267-77.
30. Heimdal, K., et al., *STUB1 mutations in autosomal recessive ataxias - evidence for mutation-specific clinical heterogeneity.* Orphanet J Rare Dis, 2014. **9**: p. 146.
31. Qian, S.B., et al., *CHIP-mediated stress recovery by sequential ubiquitination of substrates and Hsp70.* Nature, 2006. **440**(7083): p. 551-5.
32. Jiang, J., et al., *CHIP is a U-box-dependent E3 ubiquitin ligase: identification of Hsc70 as a target for ubiquitylation.* J Biol Chem, 2001. **276**(46): p. 42938-44.
33. Pridgeon, J.W., et al., *Proteomic analysis reveals Hrs ubiquitin-interacting motif-mediated ubiquitin signaling in multiple cellular processes.* FEBS J, 2009. **276**(1): p. 118-31.
34. de Bie, P. and A. Ciechanover, *Ubiquitination of E3 ligases: self-regulation of the ubiquitin system via proteolytic and non-proteolytic mechanisms.* Cell Death Differ, 2011. **18**(9): p. 1393-402.
35. McDonough, H. and C. Patterson, *CHIP: a link between the chaperone and proteasome systems.* Cell Stress Chaperones, 2003. **8**(4): p. 303-8.
36. Nikolay, R., et al., *Dimerization of the human E3 ligase CHIP via a coiled-coil domain is essential for its activity.* J Biol Chem, 2004. **279**(4): p. 2673-8.
37. Zhang, M., et al., *Chaperoned ubiquitylation--crystal structures of the CHIP U box E3 ubiquitin ligase and a CHIP-Ubc13-Uev1a complex.* Mol Cell, 2005. **20**(4): p. 525-38.
38. Zhang, H., et al., *A bipartite interaction between Hsp70 and CHIP regulates ubiquitination of chaperoned client proteins.* Structure, 2015. **23**(3): p. 472-82.
39. Paul, I. and M.K. Ghosh, *The E3 ligase CHIP: insights into its structure and regulation.* Biomed Res Int, 2014. **2014**: p. 918183.
40. Smith, M.C., et al., *The E3 ubiquitin ligase CHIP and the molecular chaperone Hsc70 form a dynamic, tethered complex.* Biochemistry, 2013. **52**(32): p. 5354-64.
41. Graf, C., et al., *Insights into the conformational dynamics of the E3 ubiquitin ligase CHIP in complex with chaperones and E2 enzymes.* Biochemistry, 2010. **49**(10): p. 2121-9.
42. Waters, P.J., *Degradation of mutant proteins, underlying "loss of function" phenotypes, plays a major role in genetic disease.* Curr Issues Mol Biol, 2001. **3**(3): p. 57-65.
43. Bross, P., et al., *Protein misfolding and degradation in genetic diseases.* Hum Mutat, 1999. **14**(3): p. 186-98.
44. Alves-Rodrigues, A., L. Gregori, and M.E. Figueiredo-Pereira, *Ubiquitin, cellular inclusions and their role in neurodegeneration.* Trends Neurosci, 1998. **21**(12): p. 516-20.
45. Hohn, A., T. Jung, and T. Grune, *Pathophysiological importance of aggregated damaged proteins.* Free Radic Biol Med, 2014. **71**: p. 70-89.

46. Dickey, C.A., et al., *Brain CHIP: removing the culprits in neurodegenerative disease.* Trends Mol Med, 2007. **13**(1): p. 32-8.
47. Ross, C.A. and C.M. Pickart, *The ubiquitin-proteasome pathway in Parkinson's disease and other neurodegenerative diseases.* Trends Cell Biol, 2004. **14**(12): p. 703-11.
48. Tyedmers, J., A. Mogk, and B. Bukau, *Cellular strategies for controlling protein aggregation.* Nat Rev Mol Cell Biol, 2010. **11**(11): p. 777-88.
49. Ronnebaum, S.M., C. Patterson, and J.C. Schisler, *Emerging evidence of coding mutations in the ubiquitin–proteasome system associated with cerebellar ataxias.* Human Genome Variation, 2014. **1**: p. 14018.
50. Manto, M. and D. Marmolino, *Cerebellar ataxias.* Curr Opin Neurol, 2009. **22**(4): p. 419-29.
51. Vermeer, S., et al., *Autosomal recessive cerebellar ataxias: the current state of affairs.* J Med Genet, 2011. **48**(10): p. 651-9.
52. Hersheson, J., A. Haworth, and H. Houlden, *The inherited ataxias: genetic heterogeneity, mutation databases, and future directions in research and clinical diagnostics.* Hum Mutat, 2012. **33**(9): p. 1324-32.
53. Palau, F. and C. Espinos, *Autosomal recessive cerebellar ataxias.* Orphanet Journal of Rare Diseases, 2006. **1**(1): p. 47.
54. Anheim, M., C. Tranchant, and M. Koenig, *The Autosomal Recessive Cerebellar Ataxias.* New England Journal of Medicine, 2012. **366**(7): p. 636-646.
55. Bamshad, M.J., et al., *Exome sequencing as a tool for Mendelian disease gene discovery.* Nat Rev Genet, 2011. **12**(11): p. 745-755.
56. Johansson, S., et al., *Exome sequencing and genetic testing for MODY.* PLoS One, 2012. **7**(5): p. e38050.
57. Bettencourt, C., et al., *Clinical and Neuropathological Features of Spastic Ataxia in a Spanish Family with Novel Compound Heterozygous Mutations in STUB1.* Cerebellum, 2015. **14**(3): p. 378-81.
58. Casarejos, M.J., et al., *Trehalose improves human fibroblast deficits in a new CHIP-mutation related ataxia.* PLoS One, 2014. **9**(9): p. e106931.
59. Cordoba, M., et al., *Ataxia plus myoclonus in a 23-year-old patient due to STUB1 mutations.* Neurology, 2014. **83**(3): p. 287-8.
60. Depondt, C., et al., *Autosomal recessive cerebellar ataxia of adult onset due to STUB1 mutations.* Neurology, 2014. **82**(19): p. 1749-50.
61. Shi, C.H., et al., *Ataxia and hypogonadism caused by the loss of ubiquitin ligase activity of the U box protein CHIP.* Hum Mol Genet, 2014. **23**(4): p. 1013-24.
62. Shi, Y., et al., *Identification of CHIP as a novel causative gene for autosomal recessive cerebellar ataxia.* PLoS One, 2013. **8**(12): p. e81884.
63. Synofzik, M., et al., *Phenotype and frequency of STUB1 mutations: next-generation screenings in Caucasian ataxia and spastic paraplegia cohorts.* Orphanet J Rare Dis, 2014. **9**: p. 57.
64. NCBI, R.S. *Homo sapiens STIP1 homology and U-box containing protein 1 (STUB1), transcript variant 1, mRNA,*. [cited 2016 30 May]; Available from: http://www.ncbi.nlm.nih.gov/nuccore/NM_005861.3?report=GenBank.
65. Gill, S.C. and P.H. von Hippel, *Calculation of protein extinction coefficients from amino acid sequence data.* Anal Biochem, 1989. **182**(2): p. 319-26.
66. ExPASy Bioinformatic Resource Portal. *ProtParam tool*. [cited 2016 06 May]; Available from: http://web.expasy.org/protparam/.
67. Bohm, G., R. Muhr, and R. Jaenicke, *Quantitative analysis of protein far UV circular dichroism spectra by neural networks.* Protein Eng, 1992. **5**(3): p. 191-5.

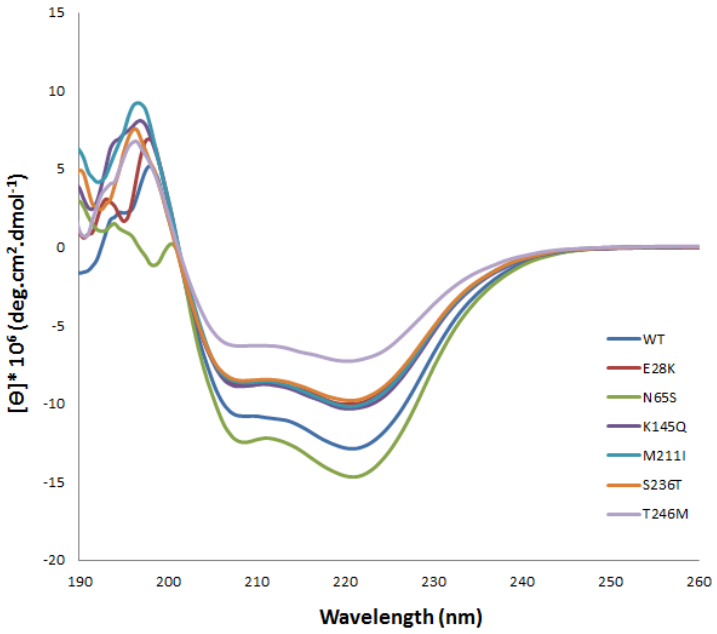
68. Greenfield, N.J., *Using circular dichroism spectra to estimate protein secondary structure.* Nature protocols, 2006. **1**(6): p. 2876-2890.
69. Whitmore, L. and B.A. Wallace, *DICHROWEB, an online server for protein secondary structure analyses from circular dichroism spectroscopic data.* Nucleic Acids Res, 2004. **32**(Web Server issue): p. W668-73.
70. Whitmore, L. and B.A. Wallace, *Protein secondary structure analyses from circular dichroism spectroscopy: methods and reference databases.* Biopolymers, 2008. **89**(5): p. 392-400.
71. Chakraborty, D., et al., *Structural characterization of the second intra-discal loop of the photoreceptor tetraspanin RDS.* FEBS J, 2013. **280**(1): p. 127-38.
72. Sandford, E. and M. Burmeister, *Genes and Genetic Testing in Hereditary Ataxias.* Genes, 2014. **5**(3): p. 586-603.
73. Orr, H.T., et al., *Expansion of an unstable trinucleotide CAG repeat in spinocerebellar ataxia type 1.* Nat Genet, 1993. **4**(3): p. 221-6.
74. Liu, Y.T., Y.C. Lee, and B.W. Soong, *What we have learned from the next-generation sequencing: Contributions to the genetic diagnoses and understanding of pathomechanisms of neurodegenerative diseases.* J Neurogenet, 2015. **29**(2-3): p. 103-12.
75. Coutelier, M., G. Stevanin, and A. Brice, *Genetic landscape remodelling in spinocerebellar ataxias: the influence of next-generation sequencing.* J Neurol, 2015. **262**(10): p. 2382-95.
76. Wang, L., et al., *Molecular mechanism of the negative regulation of Smad1/5 protein by carboxyl terminus of Hsc70-interacting protein (CHIP).* J Biol Chem, 2011. **286**(18): p. 15883-94.
77. Wienken, C.J., et al., *Protein-binding assays in biological liquids using microscale thermophoresis.* Nat Commun, 2010. **1**: p. 100.
78. Jerabek-Willemsen, M., et al., *Molecular interaction studies using microscale thermophoresis.* Assay Drug Dev Technol, 2011. **9**(4): p. 342-53.
79. Vander Kooi, C.W., et al., *The Prp19 U-box crystal structure suggests a common dimeric architecture for a class of oligomeric E3 ubiquitin ligases.* Biochemistry, 2006. **45**(1): p. 121-30.
80. Heuberger, E.H., et al., *Oligomeric state of membrane transport proteins analyzed with blue native electrophoresis and analytical ultracentrifugation.* J Mol Biol, 2002. **317**(4): p. 591-600.
81. Schagger, H., W.A. Cramer, and G. von Jagow, *Analysis of molecular masses and oligomeric states of protein complexes by blue native electrophoresis and isolation of membrane protein complexes by two-dimensional native electrophoresis.* Anal Biochem, 1994. **217**(2): p. 220-30.
82. Schägger, H. and G. von Jagow, *Blue native electrophoresis for isolation of membrane protein complexes in enzymatically active form.* Analytical Biochemistry, 1991. **199**(2): p. 223-231.
83. Van Humbeeck, C., et al., *Parkin occurs in a stable, non-covalent, approximately 110-kDa complex in brain.* Eur J Neurosci, 2008. **27**(2): p. 284-93.
84. Dimant, H., et al., *Direct visualization of CHIP-mediated degradation of alpha-synuclein in vivo: implications for PD therapeutics.* PLoS One, 2014. **9**(3): p. e92098.
85. Myers, J.K., C.N. Pace, and J.M. Scholtz, *Helix Propensities Are Identical in Proteins and Peptides.* Biochemistry, 1997. **36**(36): p. 10923-10929.
86. Gregort A Petsko, D.R., *Primers in Biology: Protein Structure and Function*. 1 ed. 2008: Oxfors University Press. 220.
87. Chou, P.Y. and G.D. Fasman, *Secondary structural prediction of proteins from their amino acid sequence.* Trends in Biochemical Sciences, 1977. **2**(6): p. 128-131.
88. Heiring, C. and Y.A. Muller, *Folding screening assayed by proteolysis: application to various cystine deletion mutants of vascular endothelial growth factor.* Protein Engineering, 2001. **14**(3): p. 183-188.

89. Teschner, W. and R. Rudolph, *A carboxypeptidase Y pulse method to study the accessibility of the C-terminal end during the refolding of ribonuclease A.* Biochemical Journal, 1989. **260**(2): p. 583-587.
90. Schwartz, T., et al., *Proteolytic Dissection of Zab, the Z-DNA-binding Domain of Human ADAR1.* Journal of Biological Chemistry, 1999. **274**(5): p. 2899-2906.
91. Nakagawa, N., et al., *Domain Structure of Thermus thermophilus UvrB Protein: SIMILARITY IN DOMAIN STRUCTURE TO A HELICASE.* Journal of Biological Chemistry, 1997. **272**(36): p. 22703-22713.
92. Fontana, A., et al., *Correlation between sites of limited proteolysis and segmental mobility in thermolysin.* Biochemistry, 1986. **25**(8): p. 1847-1851.
93. Fontana, A., et al., *Probing protein structure by limited proteolysis.* Acta Biochim Pol, 2004. **51**(2): p. 299-321.
94. Fontana, A., et al., *Limited Proteolysis in the Study of Protein Conformation*, in *Proteolytic Enzymes: Tools and Targets*, E.E. Sterchi and W. Stöcker, Editors. 1999, Springer Berlin Heidelberg: Berlin, Heidelberg. p. 253-280.
95. Grimsley, G.R., et al., *Measuring the conformational stability of a protein by NMR.* CSH Protoc, 2006. **2006**(1).
96. Benjwal, S., et al., *Monitoring protein aggregation during thermal unfolding in circular dichroism experiments.* Protein Science : A Publication of the Protein Society, 2006. **15**(3): p. 635-639.
97. Santoro, M.M. and D.W. Bolen, *Unfolding free energy changes determined by the linear extrapolation method. 1. Unfolding of phenylmethanesulfonyl .alpha.-chymotrypsin using different denaturants.* Biochemistry, 1988. **27**(21): p. 8063-8068.
98. Mayne, L. and S.W. Englander, *Two-state vs. multistate protein unfolding studied by optical melting and hydrogen exchange.* Protein Science : A Publication of the Protein Society, 2000. **9**(10): p. 1873-1877.
99. Englander, S.W., *Protein folding intermediates and pathways studied by hydrogen exchange.* Annu Rev Biophys Biomol Struct, 2000. **29**: p. 213-38.
100. Boonyaratanakornkit, B.B., C.B. Park, and D.S. Clark, *Pressure effects on intra- and intermolecular interactions within proteins.* Biochim Biophys Acta, 2002. **1595**(1-2): p. 235-49.
101. Ganesh, C., et al., *Thermodynamic Characterization of the Reversible, Two-State Unfolding of Maltose Binding Protein, a Large Two-Domain Protein.* Biochemistry, 1997. **36**(16): p. 5020-5028.
102. Starita, L.M., et al., *Massively Parallel Functional Analysis of BRCA1 RING Domain Variants.* Genetics, 2015. **200**(2): p. 413-22.
103. Kumar, P., et al., *Cross-functional E3 ligases Parkin and C-terminus Hsp70-interacting protein in neurodegenerative disorders.* J Neurochem, 2012. **120**(3): p. 350-70.

7. Appendix

7.1 Far-UV spectra of WT and mutant CHIPs in the wavelength range of 190-260 nm

Due to the buffer noise at wavelengths lower than 200 nm, results for the far-UV spectra were given in the range of 200-260 nm (Section 4.9). The original data is shown here:



7.2 MicroScale Thermophoresis (MST) analysis

The binding of Hsp70 to CHIP was quantified through measuring the changes in thermophoresis (directed movement of molecules along a temperature gradient) of Hsp70 as a fluorescent molecule at different ligand (CHIP) concentrations. Since the buffer is kept constant, alterations in the thermophoretic depletion or enrichment can only arise from changes in size, charge, or solvation entropy of the fluorescent molecule, indicating the presence of interaction between the molecule and ligand. Therefore, different traces can be observed as the thermophoretic movement of fluorescent Hsp70 changes upon binding to the non-fluorescent CHIP. For analysis, the change in thermophoresis was expressed as the change in normalized fluorescent (ΔF_{norm}), which is defined as F_{hot}/F_{cold} (average fluorescent values between hot and cold areas), and further plotted against different concentrations of ligand.

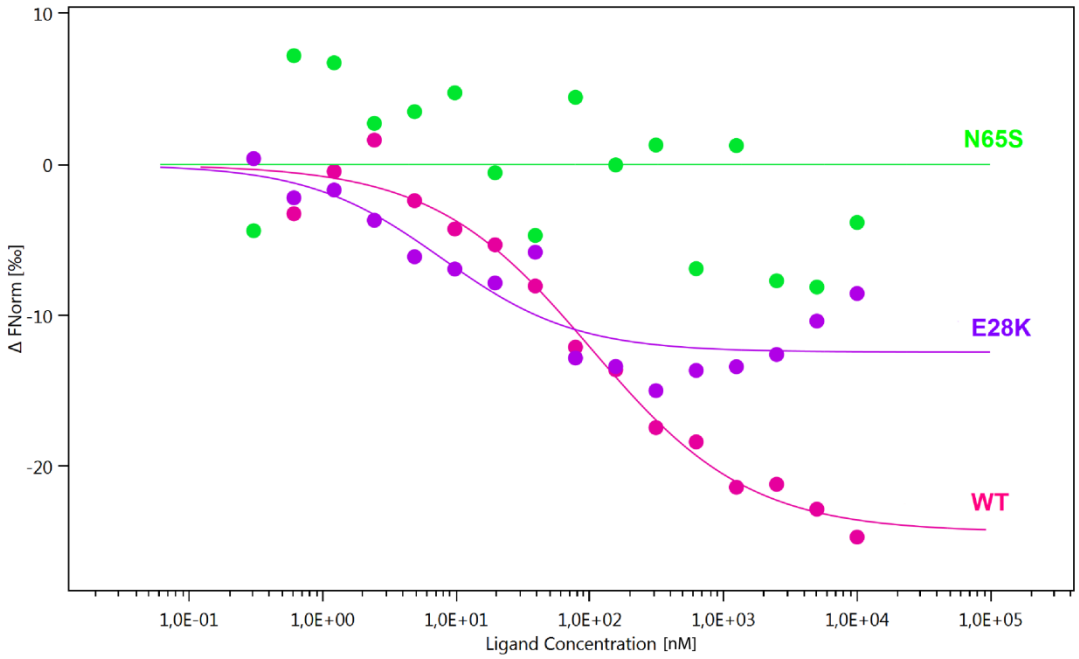


Figure 7.2.1 MST experiment of NT-647-Hsp70 VS. CHIP. A fluorescent label (NT-647) was covalently attached to the Hsp70 protein. The concentration of NT-647 labeled Hsp70 was kept constant while the concentration of CHIP is varied between 0.3 nM to 10 µM. The assay was performed in MST buffer containing 50 mM Tris-HCl pH 7.4, 150 mM NaCl, 10 mM $MgCl_2$, and 0.05% Tween. After a short incubation, the samples were loaded into MST NT.115 standard glass capillaries and the analysis was performed using the Monolith NT.115 instrument. Concentrations on the X-axis are plotted in nM. A *Kd* of 73 nM and 9 nM was determined for the WT CHIP and CHIP-E28K, respectively. No binding was observed between Hsp70 and CHIP-N65S mutant.