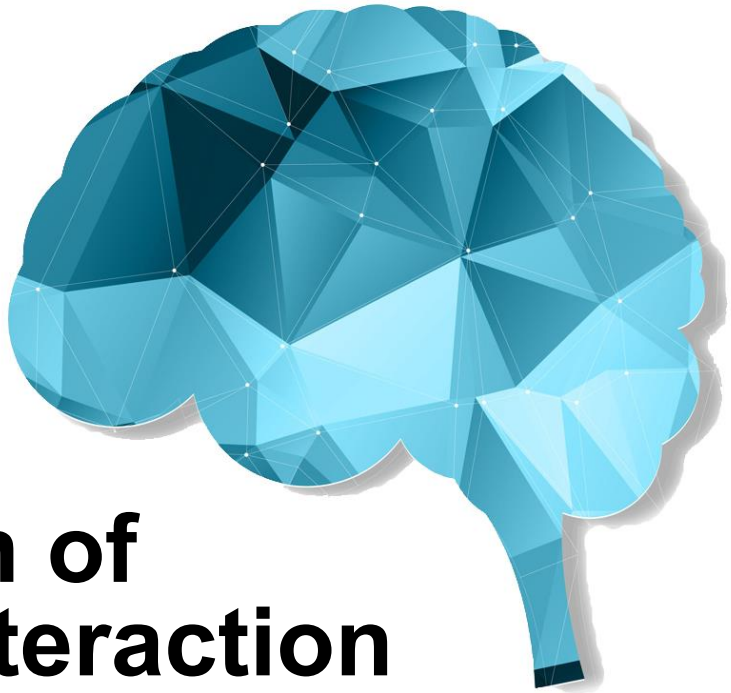

Identification of molecular interaction surfaces in Ataxin-3: targets for anti-aggregation strategies

Miguel Ângelo Cardoso de Azevedo Moreira
Mestrado em Bioquímica
Departamento de Química e Bioquímica
2016
Orientador
Sandra Macedo-Ribeiro, Professor, I3S – UP
Coorientador
Annalisa Pastore, Professor, IoPPN – King's College London

Todas as correções determinadas pelo júri, e só essas, foram efetuadas.

O Presidente do Júri,

Porto, _____/_____/_____

“We are just an advanced breed of monkeys on a minor planet of a very average star. But we can understand the Universe. That makes us something very special.”

S. Hawking

Acknowledgments

It is with genuine pleasure that I would like to express gratitude towards:

My supervisor Prof. Dr. Sandra Macedo-Ribeiro, for the opportunity to work on this project, the encouragement, guidance, advice and knowledge she has given to me throughout this year;

My co-supervisor Prof. Dr. Annalisa Pastore, for the opportunity of doing a short scientific mission in her lab to learn about the challenging, however, powerful, interesting and incredibly useful world of protein NMR, for the friendship, knowledge, guidance and advice provided throughout this experience;

Dr. Bruno Almeida, for the friendship, the excellent training, guidance, support, patience, and motivation he gave me during this project. Without him this year would not have been possible;

Dr. Alessandro Sicorello, for the companionship, training, guidance, support and especially for the patience he has given me while training me in NMR spectroscopy;

Dr. Zsuzsa Sárkány, for the training, guidance, help and discussions;

Everyone in Prof. Dr. Ribeiro's lab for the companionship, inestimable assistance and for making me feel welcome, especially Dr. Maria Oliveira and Joana Fraga;

Everyone in Prof. Dr. Pastore's lab for the outstanding support, friendship, and welcoming that made me feel almost at home when alone in such a big city as it is London, especially Dr. Elsa Zacco and Giulia Milordini;

My friends from the Master course, specially Inês Lopes, for all the support, friendship, dinners and moments we spent together. Without them, this would have been a dull journey;

My family, for all the unconditional love, the support, the patience and the efforts they have gone through to raise me and allow me to reach this stage;

My long-time and impeccable friends, specially Sérgio Moreira, Helder Berenguer and Liane Carvalho for all the support and friendship that molded me to be who I am today;

Finally, I would like to express my gratitude to the COST framework and all those involved that under the COST Action BM1405 funded the short-term scientific mission in Prof. Dr. Annalisa Pastore's laboratory.

Abstract

Machado-Joseph disease (MJD) is an autosomal dominantly inherited ataxia belonging to the group of polyQ-expansion disorders, for which there is no effective treatment. Similar to other polyglutamine-related diseases, polyQ tract expansion of the causative protein results in its self-assembly into amyloid fibrils and aggregation into neuronal inclusions. It is widely accepted that early amyloid protofibrils, rather than insoluble aggregates, are the neurotoxic species. Ataxin-3, the MJD causative protein, is a modular protein displaying ubiquitin hydrolase activity, composed of a globular domain (Josephin) followed by an extended tail comprising two ubiquitin-interacting motifs (UIMs), the expandable polyQ tract, and a C-terminal region that may contain a third UIM, depending on the isoform that is generated from the alternative splicing process that this protein is subjected to during its synthesis.

Self-assembly of taxin-3 occurs in two steps. The first step leads to the formation of protofibrils that can be monitored by Thioflavin-T binding. The second phase, which is dependent on the presence of the expanded polyQ sequence, generates long, straight and SDS-resistant mature fibrils. Several pieces of evidence underscore the central role of self-association of the globular Josephin domain on the initial aggregation step of both non-pathological and expanded ataxin-3. However, the modulatory function of the C-terminal domains of ataxin-3 on its aggregation kinetics remains largely unexplored. We have analyzed the influence of the different ataxin-3 domains on its first step of self-assembly by monitoring the early steps of ataxin-3 aggregation with a scaled-down Thioflavin-T binding assay.

Several hypotheses have been proposed to explain the mechanisms linking the expansion of the polyQ tract to MJD. Among them, the formation of toxic protein aggregates, the creation of toxic expanded polyQ based fragments and the toxic gain of function of ataxin-3 as a consequence of its polyQ expansion involving ataxin-3 aggregation. Ataxin-3 aggregation has been described to be modulated by native functional interactions. Protein expression and purification of known ataxin-3 interactors (VCP/p97 and PNKP) were carried out and the modulatory role of VCP/p97 on ataxin-3 aggregation was assessed. To map the interaction surfaces on ataxin-3 and to evaluate the molecular basis of the interaction between ataxin-3 and VCP/p97, NMR spectroscopy was used.

Key-words: Machado-Joseph disease, Ataxin-3, aggregation, VCP, protein-protein interaction, aggregation modulation.

Resumo

A doença de Machado-Joseph (MJD) é uma ataxia hereditária dominante, pertencente ao grupo das doenças causadas pela expansão de poliglutaminas, para as quais não há cura. À semelhança das outras doenças causadas pela expansão de poliglutaminas, na MJD a expansão de uma região de poliglutaminas na proteína causadora resulta na sua agregação em fibras de amiloide e subsequentemente a sua agregação em inclusões neuronais. É aceite que em vez dos agregados insolúveis, são as protofibras de amiloide, formadas inicialmente neste processo de agregação, as espécies neurotóxicas. A ataxina-3, a proteína causadora da MJD, é uma proteína modular que exibe atividade de hidrólise de ubiquitina e é composta por um domínio globular (domínio Josefina) seguido por uma cauda fléxivel que contém dois domínios de interação com ubiquitina (UIMs), a região extensível de poliglutaminas e uma região C terminal que pode conter um terceiro UIM dependendo da isoforma obtida por *splicing* alternativo.

O processo de agregação da ataxina-3 ocorre em dois passos. O primeiro origina protofibras de amiloide cuja formação pode ser monitorizada através da ligação a tioflavina-T. O segundo passo, dependente da expansão da região de poliglutaminas, leva à formação de fibras maduras de amiloide, longas, retas e resistentes a SDS. Várias evidências salientam o papel central do domínio Josefina no primeiro passo de agregação de ataxina-3 tanto para a forma não patológica como para a forma expandida da proteína. No entanto, o papel modulatório dos restantes domínios da ataxina-3 na sua cinética de agregação permanece maioritariamente ignorado. A influencia destes domínios no primeiro passo de agregação foi analisada através da monitorização do primeiro passo de agregação da ataxina-3 com ensaios miniaturizados de ligação à tioflavina-T.

Existem várias hipóteses propostas para explicar os mecanismos através dos quais a expansão da região de poliglutaminas causa a MJD. Entre estes, encontram-se a formação de agregados tóxicos, a criação de fragmentos tóxicos contendo a região de poliglutaminas e o ganho de toxicidade da ataxina-3 como consequência da sua agregação. A agregação da ataxina-3 foi descrita como sendo modulada por interações funcionais nativas. A expressão e purificação de interatores descritos da ataxina-3 (VCP/p97 e PNKP) foi conseguida e o papel modulatório da VCP/p97 na agregação da ataxina-3 foi analisado. Tendo em vista o mapeamento das superfícies desta interação

na ataxina-3 e para desvendar as bases moleculares da interação da ataxina-3 com a VCP/p97, recorreu-se a espectroscopia de ressonância magnética nuclear.

Palavras-chave: Doença de Machado-Joseph, Ataxina-3, agregação, VCP, interações proteína-proteína, modulação de agregação.

Table of contents

List of Tables

List of Figures

Supplementary figures

List of Abbreviations

- 13Q (refers to 13Q-Atx3) – Atx3 with a non-pathogenic number of glutamines (13) on its polyQ expandable tract
- ^{15}N – Nitrogen isotope (has a spin of ½ and thus is visible in NMR spectroscopy)
- ^{1}H – Hydrogen isotope (has a spin of ½ and thus is visible in NMR spectroscopy)
- 77Q (refers to 77Q-Atx3) – Atx3 with a pathogenic number of glutamines (77) on its polyQ expandable tract
- Atx3 – Ataxin-3 (protein)
- *ATXN3* – Ataxin-3 (gene)
- CEC – Cationic exchange chromatography
- cDNA – Complementary DNA
- D1 (refers to Atx3) – Atx3 truncated version encompassing the JD and UIM1/2
- D_2O – Deuterated water
- Da – Dalton (kDa – kilo-Dalton)
- DNAse - Deoxyribonuclease
- DRPLA – Dentatorubralpallidoluysian atrophy
- DTT – Dithiothreitol
- DUB – Deubiquitinating (enzyme)
- *E. coli* – *Escherichia coli*
- EDTA - Ethylenediaminetetraacetic acid
- HD – Huntington disease
- HSQC (NMR spectroscopy) – Heteronuclear single quantum coherence (NMR spectroscopy)
- IPTG – Isopropyl β-D-1-thiogalactopyranoside
- JD – Josephin domain

- LB – Luria Broth (media)
- A.u. – Absorbance units (mA.u. – miliA.u.)
- MAP2 – Microtubule-associated protein 2
- MBP – Maltose-binding protein
- MJD - Machado-Joseph disease
- MTOC – Microtubule organizing center
- NCI – Neuronal Cytoplasmic Inclusion
- NMR – Nuclear Magnetic Resonance
- O.D. – Optical Density
- VCP – Valosin-containing protein
- PMSF – Phenylmethanesulfonyl fluoride
- PNKP – Polynucleotide kinase 3'-phosphatase
- PolyQ – Polyglutamine (Glutamine one letter code is Q)
- PolyUb – PolyUbiquitin
- REM – Rapid Eye Movement (sleep phase)
- SBMA – Spinal and bulbar muscular atrophy
- SCA – Spinocerebellar ataxia
- SCA3 – Spinocerebellar ataxia type 3
- SEC – Size exclusion chromatography
- TEM – Transmission Electron Microscopy
- TEVp – Tobacco etch virus protease
- Ub – Ubiquitin
- UIM – Ubiquitin-interacting motif
- UPP – Ubiquitin-proteasome pathway
- VCP-N – N domain of VCP

- VCP-ND1 – VCP comprising N and D1 domains

- INTRODUCTION -

1. Introduction

1.1. Proteins

Proteins are macromolecules synthesized from the polymerization of a set of up to 20 different monomers, the amino acids (reviewed in [1]). The extent of their functions is so widespread that they are fundamental elements of life. These functions depend on each protein structure which is also dictated by its amino acid sequence. Proteins can provide mechanical support, generate movement, provide immune protection, perform catalysis, transport, and store other molecules, such as oxygen (if we consider, for example, the human hemoglobin), transmit nerve impulses and also control most of the molecular processes *via* biochemical pathways. In the grand scheme of things, proteins could be considered life's (as we know it) workhorses.

Being such an important element in life and able to perform such a comprehensive set of functions, its production and therefore its function is tightly regulated. For the majority of the proteins, this regulation leads to the folding and the adoption of a conformation named as the native state. The native state of a protein is a stable conformation (that can have some flexibility needed for function or regulation). This native state is characterized by a low free energy (resulting usually from the establishment of intramolecular hydrophobic interactions, thus reducing the free energy of these regions) and is a conformation in which the protein usually has its peak of performance for its biological function [2].

Protein structure is therefore obtained by a sequential process. The polymerization of a set of up to 20 different L-amino acids linked end to end originates a polypeptide linear chain described as the primary structure of a protein. This linear chain depending on its amino acids and the short-range interactions that they can establish between themselves – as a result of their individual side chains - can produce three-dimensional structures, either α-helixes or β-sheets. The further folding of the protein that is a consequence of the establishment of long-range interactions between amino acids provides the basis of the tertiary structure of proteins. However, proteins can also achieve a quaternary structure when different chains coordinate and interact among themselves through hydrogen, electrostatic and/or van der Walls interactions for the creation of a protein comprised by the interaction of various polypeptide chains.

1.2. Protein aggregation

Generally, the majority of proteins are stable systems that once folded correctly are not subject to abnormal aggregation. Although protein aggregation can have protective roles against pathological conditions (such as it happens when fibrinogen is processed into fibrin that subsequently aggregates to form blood clots thus preventing bleeding) it is also linked to the pathological mechanisms of several diseases including neurodegenerative diseases. Proteins have different propensities to aggregate and this usually occurs as a result of protein misfolding or unfolding (reviewed in [3]). Protein aggregation if allowed to happen, is a favored energetic state against unfolded states and even native states (Fig.1)

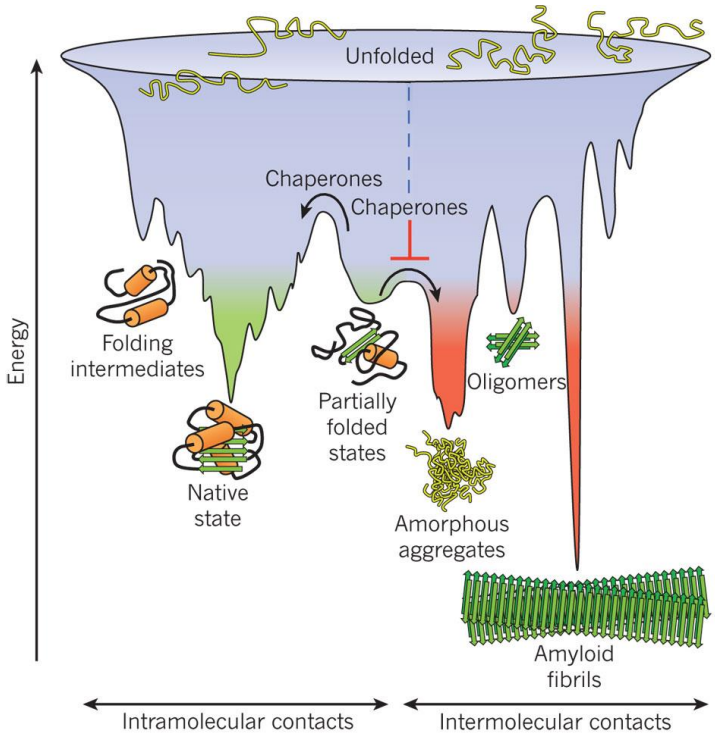

Fig. 1 – Schematic depiction of the free-energy surface that proteins experience when transitioning from an unfolded state to folded states and aggregation states. Usually, to achieve a correct folding chaperones catalyze the process to overcome free-energy barriers that obstruct the energetically favored native state and also prevent the formation of amorphous aggregates that usually may occur when several molecules fold simultaneously in the same compartment. When such thing happens and molecular chaperones do not act, proteins can experience a set of conditions that allows them to establish intermolecular contacts and overcome free-energy barriers that may result in the formation of aggregates. Adapted with permission from Hartl *et* al. 2011 [3].

Protein aggregation is a biochemical phenomenon characterized by the clumping of proteins together. It can occur both *in vitro* or *in vivo* and on the latter option both in an intracellular or extracellular manner. It depends on several factors such as the specific molecular environment experienced by the protein, its interaction with other proteins and even its concentration [4]. Protein aggregation can generate soluble or insoluble aggregates [5] that vary in their size, the number of monomers, structure, and shape [6].

However, the cell has several cellular quality control mechanisms to prevent aggregation from occurring. These mechanisms include strategies that attempt to either refold, degrade or sequester misfolded proteins (reviewed in [7]). Crucial to these strategies lies a set of proteins called chaperones. Chaperones have several functions within this proteome homeostasis assessment and rectification (reviewed in [8]). During protein synthesis, they help to achieve the correct protein folding either by the protection of hydrophobic portions of the amino acid sequence from the media, which allows these portions to be folded correctly, or by actively folding the newly synthesized protein. They can promote the refolding of misfolded proteins, or if this is not possible, they can signal them to degradation *via* the ubiquitin-proteasome pathway (UPP). Another mechanism that helps balance and prevents the overloading of the chaperones and the proteasome system is the sequestration of these misfolded proteins onto specialized quality control compartments (reviewed in [9]). These also help the cell to cope with an overload of misfolded proteins that are not refolded or degraded thus preventing the toxic gain of function activity presented by soluble aggregation intermediates [10]. However, the drawbacks inherent to this mechanism include the sequestration of correctly folded proteins alongside misfolded proteins, the difficulty in clearing these inclusions and consequent non-recycling of the amino acids (reviewed in [9]).

1.3. Protein deposition diseases

Despite the existence of the abovementioned protein quality control mechanisms, protein aggregation is still a common occurrence, and it can lead to disease as it can originate higher order oligomers with a toxic gain of function [4] and the diminishment or complete loss of the original protein function (reviewed in [11]). Several disorders of this kind are known and usually referred to as protein deposition diseases and often arise due to cytotoxicity caused by intracellular or extracellular accumulation of conformationally altered monomers or oligomers (reviewed in [12]). An interesting subset of these protein deposition diseases is often referred as amyloidosis and arise from the

aggregation of a specific protein or peptide onto a fibrillary structure called amyloid fibril that presents a highly organized β-sheet structure (reviewed in [12]). These amyloid fibrils accumulate into intracellular inclusions, or extracellular plaques are present in several diseases such as Alzheimer's disease (AD), Huntington disease (HD), and the several spinocerebellar ataxias (SCA) among others.

1.4. Polyglutamine expansion disorders

From the several protein deposition disorders, a subset of nine can be identified and are named polyglutamine (polyQ) expansion disorders (Table 1). These group of nine neurodegenerative diseases has an average frequency of 1-10 cases per 100 000 people (reviewed in [13]) and includes Huntington's disease (HD), dentatorubralpallidoluysian atrophy (DRPLA), spinal and bulbar muscular atrophy (SBMA) and SCA 1/2/3/6/7/17 (reviewed in [14]).

Table 1 – PolyQ expansion disorders. A brief description of the several pathologies including the mutated genes, protein products, and their supposed functions along with the CAG repeat size for both the pathological and normal conditions and the neurological most affected regions. Adapted with permission from Matos *et al.* 2011 [15].

Disease name	Mutated gene	Protein product	Putative function	CAG repeat size		Most affected regions
				Normal	Pathogenic	
HD	HD	Huntingtin	Signaling, transport, transcription	6-34	36-121	Striatum, cerebral cortex
DRLPA	DRLPA	Atrophin1	Transcription	7-34	49-88	Cerebellum, cerebral cortex, basal ganglia, Luys body
SBMA	AR	Androgen receptor	Steroid hormone receptor	9-36	38-62	Anterior horn and bulbar neurons, dorsal root ganglia
SCA1	SCA1	Ataxin-1	Transcription	6-39	40-82	Cerebellar Purkinje cells, dentate nucleus, brainstem
SCA2	SCA2	Ataxin-2	RNA metabolism	15-24	32-200	Cerebellar Purkinje cells, brainstem, frontotemporal lobes
SCA3	ATXN3 /MJD /SCA3	Ataxin-3	Deubiquitination activity and transcription regulation	10-51	55-87	Cerebellar dentate neurons, basal ganglia, brainstem, spinal cord
SCA6	CACNA1A	CACNA1A	P/Q-type α1A calcium channel subunit	4-20	20-29	Cerebellar Purkinje cells, dentate nucleus, inferior olive
SCA7	SCA7	Ataxin-7	Transcription	4-35	37-306	Cerebellum, brainstem, macula, visual cortex
SCA17	SCA17	TBP	Transcription	25-42	47-62	Cerebellar Purkinje cells, inferior olive

PolyQ disorders are a group of hereditary diseases that develop due to the expansion of a polyQ region (reviewed in [16]) of an aggregation-prone causative protein (reviewed in [15]). These diseases lead to a selective and progressive loss of neuronal cells and ultimately to the manifestation of the disease as a consequence of the loss of cognitive, physical and/or motility functions (although the specific symptoms vary from disease to disease). These diseases exhibit a progressive worsening of the pathological features until the death of the patient which usually occurs in 15 to 20 years (reviewed in [17]). The formation of macromolecular protein aggregates or inclusion bodies within the nuclei or cytoplasm of the affected neuronal cells is also a common hallmark of these diseases (reviewed in [15, 18]).

Despite the common characteristics of these diseases an interesting fact arises. Apart from a polyQ expandable tract encoded by a CAG repeat region, no other similarities exist between the causative proteins of these diseases (Table 1). There is no identifiable functional or structural relation between the proteins, and the encoding genes share no homology outside the polyQ encoding region (reviewed in [14, 15]).

The aggregation of the disease-related proteins is a causative feature common to all these diseases. Such aggregation is potentiated by the expansion of a CAG repeat within the coding sequence beyond a determined and disease-specific threshold. Higher number of repeats beyond the disease-specific threshold have also been correlated with the phenomenon of anticipation: higher number of repeats increase disease severity and originate an earlier age of onset [17, 19-21]. Although the pathological mechanisms by which the polyQ expansion leads to disease are not entirely understood there are several hypotheses to explain this causal relationship. The polyQ expansion can: generate toxic polyQ-containing fragments after cleavage; alter the protein functionally and/or generate an aggregation-prone conformation with the subsequent formation of toxic aggregates; modulate transcription levels by facilitation of interactions between the causative protein and transcription factors; overload the aforementioned cellular quality control mechanisms and thus create a proteotoxic stress; as well as causing mitochondrial dysfunction (reviewed in [15]).

1.5. Spinocerebellar Ataxia type 3 – Machado-Joseph Disease

Machado-Joseph disease (MJD), also known as spinocerebellar ataxia type 3 (SCA3) is a rare progressive neurodegenerative pathology with no effective treatment.

The first description of this autosomal dominant hereditary condition occurred in 1972 among Portuguese emigrants in Massachusetts [22]. It is the most common inherited ataxia in the world (reviewed in [23]) making up to 20% of the ataxia cases in the USA studies and up to 50% or more in Portuguese, Japanese and German studies [24]. MJD has a global prevalence of 1.5 cases per 100000 people [25], but several clusters appear where these figures rise as high as 1 case per 143 citizens like Corvo Island, Azores, Portugal or 1 case per 158 people as it occurs in Flores Island, Azores, Portugal [26].

Being a polyQ expansion disease and therefore a neurodegenerative disorder, its toxicity mechanisms originate neuronal loss. The most affected regions are the cerebellar dentate neurons, basal ganglia, brainstem, and the spinal cord (reviewed in [15]). Pathologically this neuronal loss results in ataxia (impaired balance and coordination), muscular dystonia and spasticity, rigidity, tremors, bulging eyes, double vision and speech impairment. Subjects can also experience restless leg syndrome and/or REM sleep behavior disorder [15, 27-30].

Similar to the other polyQ expansion disorders, MJD is also progressive over time with an average life expectancy for the patient of 20 years after onset. The age of onset is scattered between 5 and 73 years [31], but the majority of manifestations occur during adulthood with a mean age at onset of 37.4 years [20]. This rather broad spectrum for the age of onset is explained by the above mentioned negative correlation between the size of the polyQ tract expansion and the age of onset observed for the polyQ expansion disorders [17, 20].

As covered before, polyQ expansion disorders share a common hallmark and MJD is not an exception. It exhibits the development of macromolecular protein aggregates or inclusion bodies within the nuclei or cytoplasm of the affected cells. Specifically, MJD affected cells display intranuclear inclusion bodies (Fig. 2) containing the causative protein ataxin-3, as well as ubiquitin (Ub), Ub-like proteins, heat-shock proteins (HSP), other polyQ containing proteins and transcription factors [32, 33]. Other inclusion structures, designated as neuronal cytoplasmic inclusions (NCI), are also described for MJD. These occur in the cytoplasm, are immunopositive for polyQ-expanded ataxin-3 but negative for ubiquitin [34] and consist of fine granules of 1.5 μm in diameter [21].

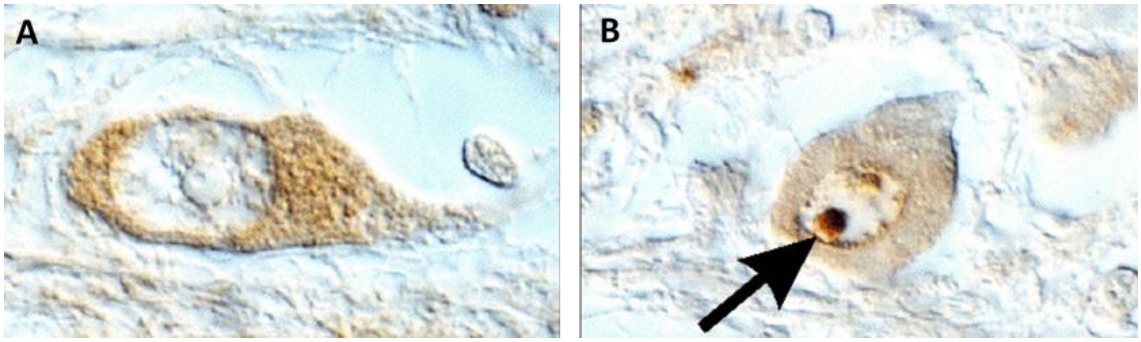

Figure 2 – Intranuclear inclusion bodies in MJD. A). Negative control for inclusion bodies immunoreactivity. B) The immunoreactivity of MJD inclusion bodies to the monoclonal antibody 1H9 directed against a UIM1 proximal region (between E214 and L233) of Atx3. The arrow indicates an inclusion body. Adapted with permission from Schmidt *et al*. 1998 [32].

1.6. Ataxin-3

MJD causative protein ataxin-3 is a protein encoded by a gene designated as ATXN3 located on the long arm of chromosome 14 (14q32.1) [35]. A year after it was mapped, it was cloned for the first time from individuals affected by MJD and found to possess an expandable CAG repeat (only interrupted by a lysine at the beginning of the repeat) [36], deeming MJD to join the other polyQ expansion disorders. The CAG trinucleotide repeat occurs in the coding region of this gene and thus is translated into a polyglutamine repeat in the gene product. Normal, non-pathological, conditions happen if the number of repeats stays between 10 and 51. The pathological condition has been associated with the expansion of the trinucleotide between 55 and 87 repeats [15, 37]. An intersection between pathological and non-pathological situations has also been identified when the number of repeats stays between 45 and 51 [15, 38].

The ~48Kb long ATXN3 gene possesses 11 exons, being the CAG repeat located in the 10th exon [21, 39] and is subject to alternative splicing as several transcripts and protein products have been identified. Although 20 different transcripts are thought to encode different isoforms of ataxin-3 (Atx3) [40], and Atx3 is expressed ubiquitously in various cell types and body tissues[15, 32, 39], unless stated otherwise, from hereon any reference to Atx3 will be to related to the brain-predominant isoform (also known as 3UIM Atx3 - UNIPROT identifier P54252-2) [32].

Atx3 containing a non-expanded polyQ tract (from hereon referred as non-expanded Atx3 or simply Atx3) has a molecular weight of ~42KDa and is non-globular modular protein. It encompasses a globular, 180 amino acid long, N-terminal catalytic domain

designated as Josephin domain (JD), followed by a flexible C-terminus comprising two ubiquitin-interacting motifs (UIMs), the polyQ stretch and a third UIM (splicing-dependent but present in the brain predominant isoform of Atx3) (Fig. 3a) [41].

Despite the many efforts to obtain a full-length structure of Atx3, only structures of isolated domains were determined. By NMR spectroscopy was unveiled that the JD is composed of two subdomains forming a semi-elongated L-structure consisting of a globular subdomain and a helical hairpin between which a catalytic cleft presents itself (Fig. 3b) [42]. Two ubiquitin binding sites are observable on the surface of JD located on adjacent but opposite surfaces being site one located near the catalytic cleft (Fig. 3c) [15, 21, 42]. The binding occurs through an induced-fit mechanism facilitated by the helical hairpin [43]. Also obtained by NMR-based spectroscopy was the structure of the two N-terminal UIMs complexed with ubiquitin. These domains assume α-helical structures linked by a short flexible region and liaise to bind ubiquitin more efficiently (Fig. 3d) [44].

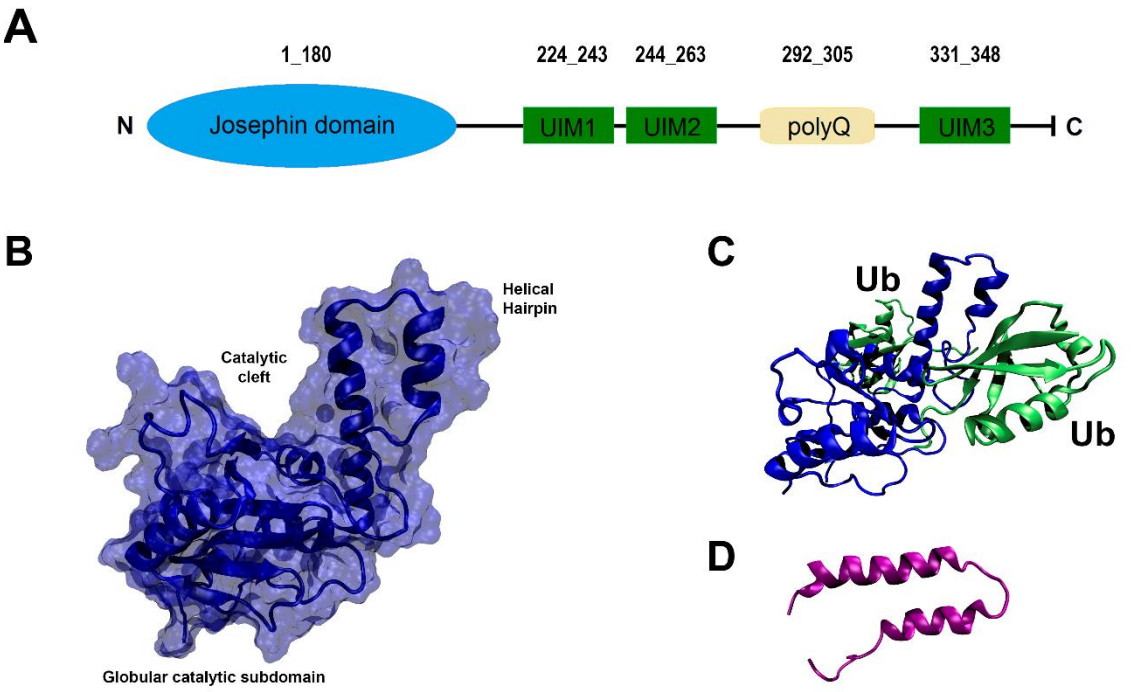

Figure 3 – Atx3 structure and domain architecture. A) Domain size and architecture of Atx3 (for the polyQ expandable region, a non-pathological expansion of 13 glutamines is considered). Atx3 is a modular protein composed of a globular catalytic Josephin domain, two UIMs followed by a polyQ expandable tract and a third isoform-dependent UIM on the C-terminus of the protein. B) The structure obtained by NMR spectroscopy of the globular catalytic Josephin domain (PDB entry 1YZB) evidencing the catalytic cleft as well as both subdomains: the globular subdomain and the helical hairpin. C) Binding sites for Ub on JD (PDB entry 2JRI). JD represented in blue and the two Ub molecules represented in green. D) The structure of the N-terminal UIMs of Atx3 obtained by NMR spectroscopy (PDB entry 2KLZ).

1.7. Ataxin-3 function

Although Atx3 first description occurred more than 20 years ago, its biological function remains a mystery. Its ubiquitin protease activity (the cleavage of Ub chains or peptidic bonds between Ub and other proteins) was originally foreseen due to bioinformatics analysis of Atx3 amino acidic sequence [45]. Confirmation arose by the use of model substrates and specific inhibitors for Ub-proteases which lead to the characterization of Atx3 and consequently other JD-containing proteins as deubiquitinating proteins (DUBs) [46]. Comparative analysis of the conservation of the catalytic amino acids of JD among Ub-proteases families evidenced that Atx3 belongs to the papain-like cysteine protease family [42, 47].

Despite the wide range of proposed biological functions for Atx3, focus will be given to those associated with protein homeostasis systems due to their direct relationship with protein aggregation.

The N-terminal UIMs of Atx3 play a crucial role in its activity since binding and DUB activity are inhibited when these are mutated. Mutation of the first UIM eliminates binding completely whereas mutation on the second UIM decreases binding affinity of polyUb chains [46]. Atx3 is able to bind both K48 and K63 Ub-linked chains although it displays a higher affinity to K63 Ub chains [48]. Since these chains are associated with cellular processes such as DNA repair, the NFkB pathway, Lewy body, aggresome formation and autophagy [49], and considering the editing function of Atx3 over these signaling chains, suggestions that Atx3 may play a part in these cellular processes can be made. Atx3 preference for polyUb chains of 4 or more monomers may also provide insights on its biological role since K48-linked Ub chains of four or more monomers are responsible for signaling proteins for proteasomal degradation through the UPP (a proteome quality control mechanism) [46-48]. In addition to this, it was described that inhibition of Atx3 catalytic activity leads to an accumulation of polyubiquitinated proteins in the nucleus similar to that occurring when the proteasome system in inhibited [50]. Considering the suggestions of Atx3 editing the size of polyUb chains instead of disassembling them [46, 48, 51], we may postulate that Atx3 interferes in the UPP, modulating the degradation of signaled proteins. Therefore, Atx3 may play a role in protein quality control.

Atx3 protein-protein interactions may also help unveil the biological functions of this Ub-protease. Interactions through the C-terminal region with VCP (Valosin-containing protein – also known as p97) [47, 52, 53] and through the Ub-binding site 2 of JD with Ub-like domains of HHR23A and HHR23B (the human homologs of the yeast DNA repair

protein Rad23) [42, 54] have been identified. Both VCP and HHR23A/B are known to be involved in the shuttling of manly ERAD-processed protein cargos. ERAD (endoplasmic reticulum-associated degradation) is a protein homeostasis system that signals (through polyubiquitination) misfolded proteins or unassembled components of the secretory pathway for proteasomal degradation. Despite showing a role in the ERAD pathway these findings in conjugation with the known interaction with the proteasome itself [55] (although possibly not in a direct manner [56]) further strengthen the hypothesis of Atx3 involvement in the UPP and therefore in protein homeostasis. Nevertheless, it is still not clear the role of Atx3 in both these systems since all these interactions may suggest different roles for Atx3: it may ensure the correct targeting of substrates for degradation by editing polyUb chains; it may enable the future subsequent disassembly of the polyUb chain (by other proteasome-associated DUBs) by trimming polyUb chains of a substrate; or it may recognize some of the proteasome substrates by performing as a transiently associated subunit [15, 53].

Atx3 described interaction with dynein and histone deacetylase 6 (HDAC6) [57, 58] (responsible for the transport of misfolded proteins to the microtubule organizing center - MTOC), as well as the interaction with cytoskeleton components tubulin and microtubule-associated protein 2 (MAP2) [59, 60] and the co-localization with aggresome and pre-aggresome particles [58] indicates also a role in another protein homeostasis mechanism. When the proteasome is overloaded and not capable of dealing with all the misfolded proteins, misfolded protein aggregates form near the MTOC, being these structures later degraded by lysosomes and called aggresomes [61]. Even though Atx3 co-localizes with these structures, it also appears to be important in the formation of the aggresomes themselves, by potentiating their formation [58, 62] leading to the hypothesis that Atx3 may stabilize carrier proteins or protect misfolded proteins during transport [15, 58].

Since we do not fully understand the involvement of Atx3 in these protein homeostasis systems and considering that these affect MJD, it is imperative to study and fully comprehend Atx3 functions and aggregation mechanisms to provide better targets for disease treatment.

Atx3 interactome suggests additional biological functions: interaction with histones, transcription repressors, transcription activators as well as direct influence on transcription levels implies a putative role in transcription regulation; interactions with dynein tubulin, MAP2, the morphological modifications and the loss of adhesion

observed in Atx3-silenced cell lines suggests a role in the organization of the cytoskeleton; and the impairment of the myoblast to muscle fiber transition in the absence of Atx3 indicates a role in myogenesis; (reviewed in [15]). Atx3 hypothesized biological functions are summarized in Fig.4.

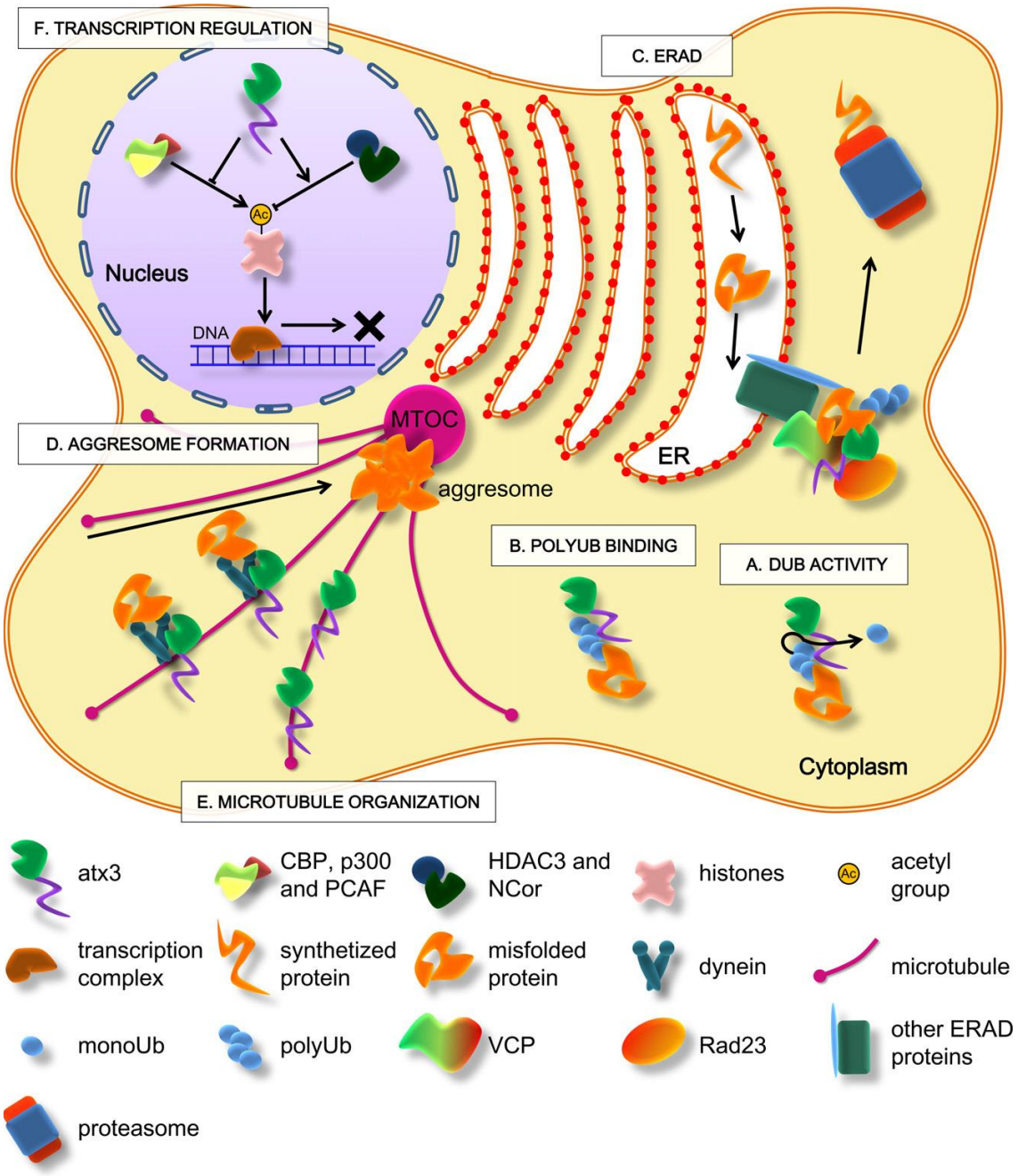

Figure 4 - Atx3 biological activity, summarized interactome and proposed biological roles. A) Atx3 Dub activity combined with its B) ability to bind polyUb chains as well as the C) described interactions with components of the ERAD and UPP protein homeostasis systems imply a biological role on these proteome quality control systems. D) Interaction with aggressome-associated proteins, the co-location with aggresome and pre-aggresome particles as well as interaction with

E) cytoskeletal proteins shows evidence of a biological role for Atx3 in aggresome formation and microtubule organization, respectively. F) Finally, interaction with transcription activators, repressors and histones denote a role in transcription regulation. Adapted with permission from Matos *et al.* 2011 [15].

1.8. Ataxin-3 aggregation

A common hallmark of protein deposition diseases is the formation of macromolecular protein aggregates or inclusion bodies within the nuclei or cytoplasm of the affected neuronal cells (reviewed in [15]). For MJD, as mentioned above, affected cells display intranuclear inclusion bodies containing Atx3 [32, 33] as well as other NCIs that occur in the cytoplasm and consist of fine granules of 1.5 µm in diameter [21].

Atx3 aggregation takes place *in vitro* through a two-step self-assembly process (Fig. 5). The first step starts by the self-association of the Josephin domain in a native or near-native conformational state and being the aggregation kinetics described by a sigmoid curve [63-65]. Aggregation occurs through an aggregation-prone region (residues 73-96) and although the molecular mechanisms remain a mystery, results from limited proteolysis suggest that the presence of the polyQ expanded tract affects and increases the time of exposure of this aggregation-prone region (either by lowering the energetic barrier for an aggregation-prone conformational state or by holding JD in this aggregation-prone conformational state) thereby leading to the observed faster assembly rate of the expanded Atx3 [63]. This association first originates Atx3 dimers (in both expanded and non-expanded Atx3) that act as a template for the assembly of further monomers. The addition of monomers to Atx3 oligomers occurs then in a linear manner. Ultimately this first step concludes with the formation of soluble amyloid non-mature fibrils [63] that have been shown to be cytotoxic [66]. The second phase of aggregation of Atx3 is polyQ expansion-dependent and therefore only occurs in the expanded Atx3. In this step the expanded polyQ regions, aligned due to the formation of the non-mature amyloid fibrils, undergo an independent polymerization through intermolecular interactions between the polyQ regions that ends in the formation of mature SDS-resistant amyloid fibrils, without the rearrangement of the protofibril core as discovered by limited proteolysis [63].

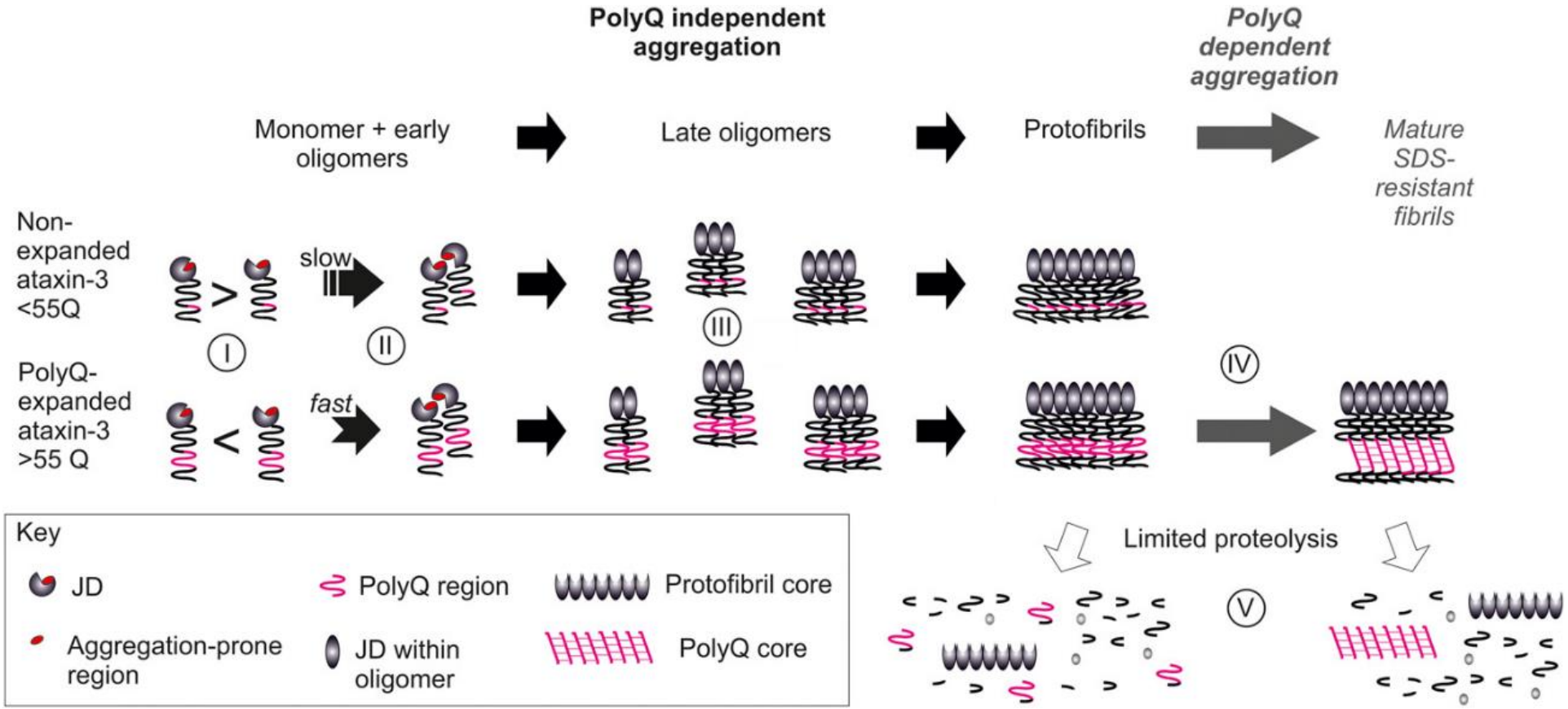

Figure 5 – Atx3 aggregation mechanisms. I) Both pathological and non-pathological variants of Atx3 undergo aggregation. The presence of the expanded polyQ tract in Atx3 leads to more frequent exposure an aggregation-prone region in JD, therefore, II) resulting in a faster dimerization of expanded Atx3 in comparison to non-expanded Atx3. III) The growth of the aggregation core by the addition of Atx3 monomers occurs in a linear manner in both variants. IV) In expanded Atx3 the protofibrils due to intermolecular polymerization of the polyQ regions mature onto SDS-resistant mature amyloid fibrils with V) Two independent aggregation cores: the protofibril core (common in expanded and non-expanded Atx3) and the polyQ core. Adapted with permission from Scarff *et al.* 2015 [63].

1.9. MJD therapy and the use of native interactors as modulators of Ataxin-3 aggregation

There are several proposed mechanisms for how Atx3 polyQ expansion leads to MJD. The most popular are the formation of toxic protein aggregates, the creation of toxic expanded polyQ based fragments as a result of proteolytic cleavage and the toxic gain of function of Atx3 as a consequence of its polyQ expansion (reviewed in [15]).

The current consensus on how to approach therapy for this pathology is to eliminate the causative expanded protein from the cells. Either by suppressing its production, increasing its degradation or by addressing the putative toxic roles of its proteolytic cleavage or aggregation (reviewed in [15]).

In this work, we will try to tackle Atx3 aggregation in order to modulate and delay its aggregation kinetics. A superimposition between the binding sites of JD for Ub and JD aggregation-prone region was described, thus establishing a link between Atx3 aggregation and function as well as a protective role of functional interactions against Atx3 aggregation, since the binding of Ub chains to JD prevents Atx3 aggregation [43, 64]. Moreover, native interactions between Ub and the UIMs of Atx3 have been shown to delay Atx3 aggregation [67]. Bearing in mind this information, targeting Atx3 aggregation with described interactors of this protein may lead to a blocking of the aggregation-prone sites of Atx3 and therefore, provide us with means to target and delay this aggregation *in vivo* (eventually allowing the design of anti-aggregation drugs).

To study the effect of interactors on Atx3 aggregation, two different proteins were selected: VCP and PNKP.

VCP is an essential and abundant AAA+ ATPase involved in a myriad of processes, with the most important functions being the disassembly of soluble and membrane protein complexes to aid the UPP in proteome quality control, the regulation of gene expression and mitotic fragmentation of endoplasmic reticulum and Golgi apparatus membranes. It forms a homohexamer, and each monomer comprises an N-terminal domain and two highly conserved D1 and D2 ATPase domains (Fig. 6a) [53]. It was previously shown that VCP interacts through an ND1 region (VCP encompassing the N-terminal domain and domain D1) with the Atx3 arginine/lysine-rich motif upstream the polyQ tract (Fig. 6b) [53, 68]. This interaction is modulated by a posttranslational modification on Atx3, which also decreases Atx3 aggregation [68].

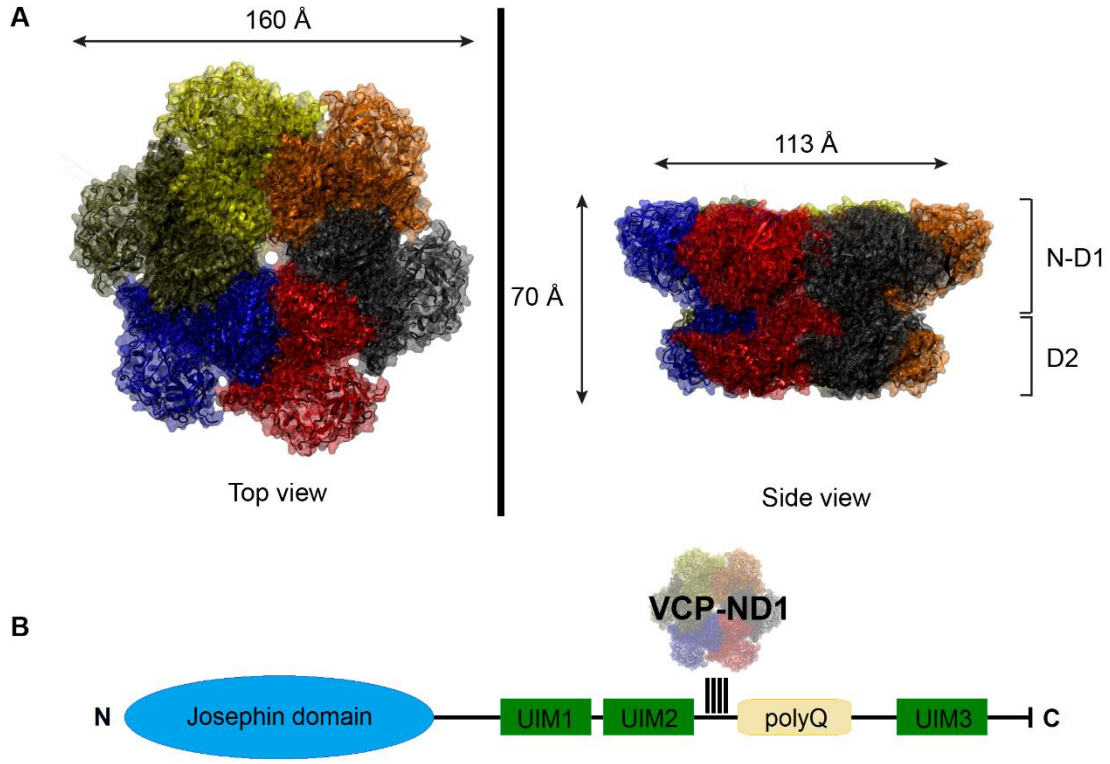

Figure 6 – VCP domains, structure, and interaction with Atx3. A) The structure of VCP, elucidating the location of N-D1 and D2 domains (top and side view) as well as their dimensions. B) Scheme of the modular structure of Atx3 displaying the interaction site of VCP.

PNKP (polynucleotide kinase 3'-phosphatase) is a quality control enzyme involved in the repair of DNA double-strand breaks or DNA single-strand breaks as it possesses both DNA 3'phosphatase and DNA 5' kinase activities [69]. PNKP is known to be associated with neurological diseases and an interaction with Atx3 was accidentally discovered [70].

This work aimed to evaluate the role of the several Atx3 domains on its aggregation kinetics and to perform a molecular characterization of Atx3 interactions with the abovementioned known interacting partners, as well as the determination of their impact on Atx3 aggregation. By a detailed characterization of the interactions between Atx3 and VCP (both ND1 domains and N domain alone) and Atx3 and PNKP, I expect to understand better how native functional interactions modulate Atx3 aggregation and to gather clues for translating this knowledge into possible therapeutic strategies.

- METHODS -

2. Methods

2.1. Expression plasmids

pDEST17 gateway vectors (Life Technologies) transformed with Atx3 cDNA were already available in the laboratory [63, 68, 71]. Encoding vectors for JD, JD+UIM1/2 (hereby designated as D1) as well as PolyQ non-expanded Atx3 (13Q) and polyQ expanded Atx3 (77Q) with 6xHis purification tags were available and used in this work.

pQE vectors encompassing 6xHis purification tag as well as the N and D1 domains of VCP (from here on designated as VCP-ND1) and the N domain of VCP (hereby designated as VCP-N) were kind gifts from Dr. Di Xia (National Cancer Institute, USA) and Dr. Alessandro Sicorello (NIMR, London, UK).

The expression plasmid containing the coding sequence of PNKP, a 6xHis purification tag linked to an MBP solubility tag was a gift from Dr. P. Sarkar (University of Texas Medical Branch, USA).

2.2. Protein expression and purification

All protein expression and purification steps were monitored by sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) using 12% polyacrylamide (wt/vol) gels.

2.2.1. Ataxin-3 proteins

Escherichia coli BL21 (DE3)-SI cells (Invitrogen Life Technologies) were transformed with the Atx3-encoding vectors (13Q-Atx3, 77Q-Atx3, JD, and D1) and the 6xHis-tagged Atx3 proteins were expressed. Cells were grown at 37ºC in Luria Broth (LB) media without NaCl and supplemented with 0.2% (w/v) glucose and 100mg/l of ampicillin. When an optical density (O.D.) of 0.6-0.8 (measured at a wavelength of 600nm) was reached, protein expression was induced by supplementing the media with NaCl to a final concentration of 0.3M and protein overexpression was allowed to occur for 3h at 30ºC. After protein overexpression, cells were harvested by centrifugation at 4000xg, resuspended in lysis buffer (20mM sodium phosphate, pH 7.5; 500mM NaCl; 20mM Imidazole) supplemented with lysozyme (50mg/l) and frozen at -20ºC. Before

purification, DNAse, Mg^{2+}, and PMSF were added to the protein extract to a final concentration of 1mg/mL, 25mM, and 5mM, respectively. The supernatant obtained after centrifugation at 30000xg was filtered through a low protein binding syringe filter (Millex, Merck). This protein extract was then loaded onto a Histrap chelating column (Amersham Biosciences) previously equilibrated with lysis buffer. Elution was performed with lysis buffer containing increasing concentrations of imidazole (50mM; 250mM; and 500mM). Selected fractions were filtered and loaded onto a HiPrep 26/60 Sephacryl S200 HR (GE Healthcare Life Sciences) previously equilibrated with filtration buffer (20mM sodium phosphate pH 7.5; 200mM NaCl; 5% Glycerol; 2mM EDTA; 1mM DTT) and subjected to a size exclusion chromatography (SEC). Elution occurred at a flow rate of 2ml/min at 6ºC and was monitored by light absorbance at 280nm. After analysis by SDS-PAGE, the purest fractions were concentrated on Amicon Ultra-15 concentrators (cutoff 10KDa, Amicon, Merck), immediately flash-frozen in liquid nitrogen and stored at -80ºC.

For the thioflavin-T binding aggregation assays, a further step of purification was performed. 50µl of a purified sample was filtered and loaded on a Superdex200 Increase 5/150 GL (GE Healthcare Life Sciences) previously equilibrated with the Thioflavin binding assay buffer (20mM sodium phosphate pH 7.5; 200mM NaCl; 1mM DTT) and eluted at a flow rate of 0.3ml/min.

2.2.2. VCP proteins

VCP-ND1 (the truncated version of VCP encompassing the domains N and D1 – a ~ 58kDa monomer that undergoes hexamerization under native conditions [72]) was expressed and purified as described [72], with slight modifications described below. Briefly, *E. coli* M15 pREP4 cells (Qiagen) transformed with a pQE vector containing the coding sequence of VCP-ND1 (under the control of a T5 promoter) were grown in LB media supplemented with 100mg/L ampicillin and 30mg/L kanamycin at 37°C. When cell culture reached an O.D. at 600nm comprehended between 0.6 and 0.8 the temperature was decreased to 25°C and protein overexpression was induced with a final concentration of 0.5mM of isopropyl β-D-1-thiogalactopyranoside (IPTG). Protein overexpression was allowed for 3-4 hrs. Cells were harvested by centrifugation at 4000xg, resuspended in 20mM sodium phosphate buffer, 200mM NaCl, 10mM Imidazole, pH 7.5 supplemented with 50mg/l of lysozyme and frozen at -20°C. Two protein purification steps were performed. A Ni^{2+} affinity chromatography on a Histrap chelating column (Amersham Biosciences) was carried out with two wash steps of 20mM

and 100mM Imidazole. VCP-ND1 was eluted from the column using 20mM sodium phosphate buffer, 200mM NaCl, 250mM imidazole, pH 7.5. The eluted fractions were then subjected to SEC on a Hiprep Sephacryl S300 HR (GE Healthcare Life Sciences) previously equilibrated with the gel filtration buffer (20mM sodium phosphate pH 7.5; 200mM NaCl; 1mM DTT) and eluted with a flow rate of 2ml/min.

Expression of the N-domain of VCP (~21.9kDa) was achieved using the same protocols as those used for the expression of VCP-ND1. VCP-N purification was performed in several steps. Initially, a Ni^{2+} affinity chromatography was performed using 20mM sodium phosphate, pH 7.5 and 500mM NaCl as a buffer. Proteins were eluted with increasing concentrations of imidazole (0mM; 50mM; 125mM; 250mM; and 500mM). Selected fractions were incubated with TEVp (Tobacco etch virus protease), in order to remove an MBP (maltose binding protein) tag fused with VCP-N to increase its solubility (total molecular weight of the expressed construct: ~65.3kDa), coupled with an overnight dialysis at 4ºC, against 20mM sodium phosphate pH 7.5; 200mM NaCl; 1mM DTT. After dialysis, the sample was subjected to a ¼ dilution with the Ni^{2+} affinity chromatography buffer and centrifuged for 10min at 16100xg. The supernatant was filtered through a low-binding protein syringe filter (Millex, Merck) and subjected to another Ni^{2+} affinity chromatography where the flow through was collected. The sample was concentrated and loaded onto a Superdex 75 10/300 GL or Hiprep Sephacryl S100 HR previously equilibrated with 20mM sodium phosphate pH 7.5, 200mM NaCl, 1mM DTT. After SEC, the eluted fractions were selected concentrated on Amicon Ultra-15 concentrators (cutoff 3KDa, Amicon, Merck), flash frozen in liquid nitrogen and stored at -80ºC. To further purify this VCP-N samples an MBP affinity chromatography was performed using an MBPTrap HR column (GE Healthcare Life Sciences) equilibrated with 20mM sodium phosphate pH 7.5, 200mM NaCl and 1mM DTT. After the MBP affinity chromatography, selected samples were concentrated on Amicon Ultra-15 concentrators (cutoff 3KDa, Amicon, Merck), flash frozen in liquid nitrogen and stored at -80ºC.

2.2.3. PNKP

The design of the protein expression and purification methodologies was based on previously described protocols [69] with some modifications. The PNKP-expressing plasmid was transformed into *E. coli* BL21 (DE3) cells (New England Biolabs, Ipswich, MA, USA). For expression of PNKP (~57.1 kDa), these cells were grown in LB media at 37ºC supplemented with 100mg/L of ampicillin until an O.D. between 0.6 and 0.8 at

600nm was reached. The temperature was then decreased to 30ºC and protein overexpression induced by supplementing the media with IPTG to a final concentration of 1mM. Protein overexpression was allowed to occur for 3h and cells were harvested and processed by the same methods used for the expression of Atx3 proteins using the following lysis buffer: 50mM sodium phosphate pH 7.5; 500mM NaCl; 10% glycerol; 1mM β-mercaptoethanol; 10mM Imidazole; and 1mM PMSF.

The protein was purified by three sequential chromatographic procedures: Ni^{2+} affinity chromatography, cation exchange chromatography, and size exclusion chromatography. The metal affinity chromatography was performed by equilibrating the Histrap chelating column with 50mM sodium phosphate pH 7.5, 500mM NaCl, 10% glycerol, 1mM β-mercaptoethanol, and eluting proteins with increasing concentrations of imidazole (0mM; 50mM; 100mM; 250mM; 300mM; and 400mM). Following buffer exchange to 50mM sodium phosphate pH 7.5, 100mM NaCl, 10% glycerol, 1mM DTT, the selected fractions were then submitted to a cation exchange chromatography (Mono S 5/50 GL, GE Healthcare Life Sciences) equilibrated with 50mM sodium phosphate pH 7.5; 10% glycerol; 1mM DTT. Protein was eluted with a NaCl gradient (0.1M to 1M). Eluted and selected fractions were applied to a Superdex 75 10/300 GL (GE Healthcare Life Sciences) previously equilibrated with 50mM sodium phosphate pH 7.5, 100mM NaCl, 10% glycerol and 1mM DTT. The purified protein was concentrated on Amicon Ultra-15 concentrators (cutoff 10KDa, Amicon, Merck), flash-frozen in liquid nitrogen and stored at -80ºC.

2.2.4. Expression and purification of labeled proteins for NMR spectroscopy

Protein expression and purification for NMR data collection were performed with minor modifications to those specified above.

Briefly, for ^{15}N-Atx3 13Q and ^{15}N-Josephin production, *E. coli* BL21 (DE3) cells (New England Biolabs, Ipswich, MA, USA) were transformed with the Atx3-encoding vectors and 6xHis-tagged proteins were expressed by incubating these cells in M9 minimal media supplemented with ^{15}N-ammonium sulfate at 37°C until an O.D. between 0.6 and 0.8 was reached. The temperature was then lowered to 25°C, and protein expression was induced by adding IPTG to a final concentration of 1mM. Protein overexpression

was allowed for 3-4 hrs. Cells were harvested by centrifugation at 4000xg and frozen at -20°C.

Before purification DNAse, lysozyme, and a protease inhibitor cocktail (Complete EDTA-free, Roche) were added to the protein extract, and the supernatant obtained after centrifugation at 30000xg containing the 6xHis-tagged proteins was purified by Ni^{2+} affinity chromatography and eluted with 20mM sodium phosphate buffer, 200mM NaCl, 300mM Imidazole, pH 7.4. The protein obtained was incubated at room temperature for one hour with TEVp for removal of the 6xHis-tag. The sample was loaded onto a Ni^{2+} affinity column, and the flow-through containing the cleaved protein was collected. SEC was then used as a final clean up step. The separation was performed in 20mM sodium phosphate buffer, 200mM NaCl, 1mM DTT, pH 7.4 using a Superdex 75 16/600 PG column (GE Healthcare Life Sciences). The purified protein was concentrated using 10kDa cutoff Vivaspin® 20 centrifugal concentrators (Sartorius Stedim Biotech GmbH, Goettingen, Germany).

2.3. Western blotting

Samples were separated on 12% polyacrylamide (wt/vol) SDS-PAGE gels and electrotransferred onto polyvinylidene fluoride membranes (Merck Millipore, Darmstadt, Germany). The membranes were blocked with TBS-Tween (20mM Tris–HCl, pH 7.5, 150mM NaCl, 0.1% Tween-20 (vol/vol)) with 5% non-fat dry milk (wt/vol). Antibodies were diluted in TBS-Tween with 1% non-fat dry milk (wt/vol). The membranes were incubated with the mouse monoclonal anti-His (1: 5.000; GenScript, Piscataway, NJ, USA) primary antibody for one hour at room temperature. This was followed by incubation with the goat polyclonal anti-mouse (1: 10.000; GenScript, Piscataway, NJ, USA) secondary antibody (one hour at room temperature). The membranes were scanned with a Molecular Imager GS800 calibrated densitometer (Bio-Rad, CA, USA).

2.4. Ataxin-3 aggregation kinetics

2.4.1. Thioflavin-T binding assays

Atx3 aggregation kinetics were evaluated in a miniaturized thioflavin-T amyloid binding assay [68, 71] using a FluoDia T70 (Photon Technology International, Edison, NJ, USA) microplate fluorimeter. Thioflavin-T fluorescence (emission at 480nm with

excitation at 440nm) was monitored at 37ºC using a solution containing 5µM protein and 30µM thioflavin-T in 20mM sodium phosphate pH 7.5, 200mM NaCl, 1mM DTT. Each well was covered with 20µl of paraffin oil to prevent evaporation.

2.4.2. Filter retardation assays

SDS-resistant fibers/aggregates formed in the thioflavin-T binding assays were detected after sample filtration. Briefly, 5µl samples from thioflavin-T assay endpoints were diluted 40-fold in TBS (50mM Tris-HCl, pH 7.5, 150mM NaCl) with 2% (wt/vol) SDS and boiled for 5 min. Samples were filtered through a cellulose acetate membrane (0.2 µm, Whatman), pre-equilibrated in TBS, using a Bio-Dot SF microfiltration apparatus (Bio-Rad). The membrane was then washed twice with TBS with 0.1% SDS (wt/vol) before being probed with anti-Atx3 antibodies using conventional western blotting protocols. Briefly, the membrane was blocked with TBS-Tween (20mM Tris–HCl, pH 7.5, 150mM NaCl, 0.1% Tween-20 (vol/vol)) containing 5% non-fat dry milk (wt/vol) and incubated for one hour at room temperature with the 1H9 mouse monoclonal anti-Atx3 antibody (1: 10.000; MAB5360, Merck Millipore, Darmstadt, Germany). The secondary goat polyclonal anti-mouse antibody (1: 10.000; GenScript, Piscataway, NJ, USA) was incubated for one hour at room temperature and detection was performed with ECL Plus reagent (GE Healthcare, Little Chalfont, UK). All antibodies were diluted in TBS-Tween with 0.5% non-fat dry milk (wt/vol). Scanning was performed on a Molecular Imager GS800 calibrated densitometer (Bio-Rad, CA, USA).

2.5. Nuclear Magnetic Resonance (NMR) data collection and analysis

The interaction between the Josephin domain of Atx3 and VCP was investigated using nuclear magnetic resonance (NMR) spectroscopy. A $^{1}H–^{15}N$ heteronuclear single quantum coherence (HSQC) titration experiment was performed on uniformly ^{15}N-labeled Josephin with and without an unlabeled truncated version of VCP encompassing the N and D1 domains. The spectra were acquired in 10mM HEPES, 150mM NaCl, 1mM DTT, 10% D_2O, pH 7.5 at 25°C on a Bruker spectrometer equipped with a cryoprobe and operating at a ^{1}H frequency of 600 MHz.

Sweep widths of 7812.5Hz and 2431.87Hz were used respectively for the direct (^{1}H) and indirect (^{15}N) dimension. For each spectrum, 704 increments in the indirect dimension were acquired with four scans per increment.

The spectrum of the Josephin domain was acquired using a sample of ^{15}N-labeled Josephin at a concentration of 0.2mM. To acquire the spectrum of Josephin in the presence of VCP-ND1, 115µl of Josephin at 0.4mM were mixed with 115µl of VCP-ND1 at 200µM to achieve a final concentration of 200µM and 100µM of Josephin and VCP-ND1, respectively (2:1 ratio).

The NMR spectra were processed with NMRPipe/NMRDraw [73] and analyzed with ccpNmr [74].

- RESULTS -

3. Results

3.1. Protein expression and purification protocols

Protein expression and purification was performed using the methods detailed in the methodology section. The protein constructs molecular weights are detailed in Table 2.

Table 2 – Purified proteins and respective molecular weight. List of the purified protein products including the several variants and respective molecular weights.

Protein	Variant		Molecular Weight (KDa)
Atx3	13Q	-	42.1
	77Q	-	51.6
	D1	-	32.5
	J1	-	22.1
VCP	N	-	65.3
		+MBP/6xHis-tag	21.9
	ND1	Monomeric	58.3
		Hexameric	350.1
PNKP	-	-	57.1

3.1.1. Purification of Ataxin-3 proteins

Purification of Atx3 proteins (13Q; 77Q; JD; and D1) was performed according to the previously optimized methods detailed in the methodology section [68, 71]. After a Ni^{2+} affinity chromatography, selected fractions were purified on a Hiprep Sephacryl 26/60 HR (GE Healthcare Life Sciences) gel filtration system. Typical chromatograms for this preparative SEC are detailed in the supplementary Fig. S1a.

After protein purification and for protein quality control, a sample of each purified Atx3 protein after concentration was analyzed by SDS-PAGE. Typical results are shown in Fig. 7.

Before protein aggregation assays, as a protein quality control procedure and also for removal of nucleation clusters of aggregation, a SEC was performed using a small volume Superdex 200 Increase 5/150 GL (GE Healthcare Life Sciences) – and typical chromatograms are presented in the supplementary Fig. S1b.

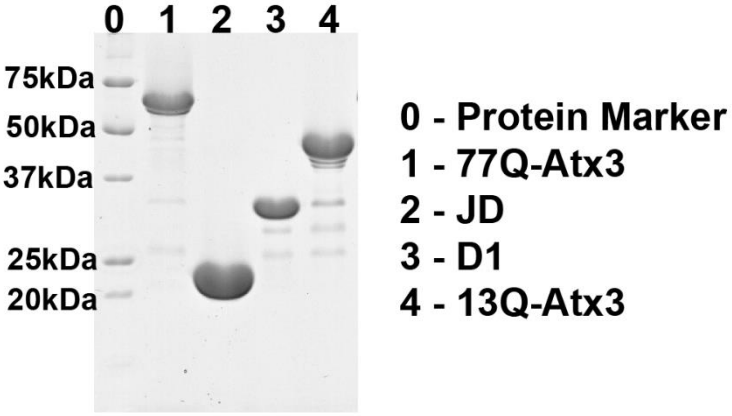

Figure 7 – Quality control of the purification of Atx3 proteins. Characteristic separation observed on a 12% polyacrylamide SDS-PAGE electrophoresis for the purified Atx3 proteins.

3.1.2. Purification of VCP-ND1

Expression and purification of VCP-ND1 were performed as detailed in the methods section and the typical gel filtration chromatograms are shown in the supplementary Fig. S2a. As a protein quality control procedure before aggregation assays, a sample of protein was analyzed by SEC using a Superdex 200 5/150 GL (GE Healthcare) gel filtration system. Typical results are displayed in the supplementary Fig. S2b. Typical results for the 12% polyacrylamide SDS-PAGE control of protein expression and purification are shown in Fig. 8.

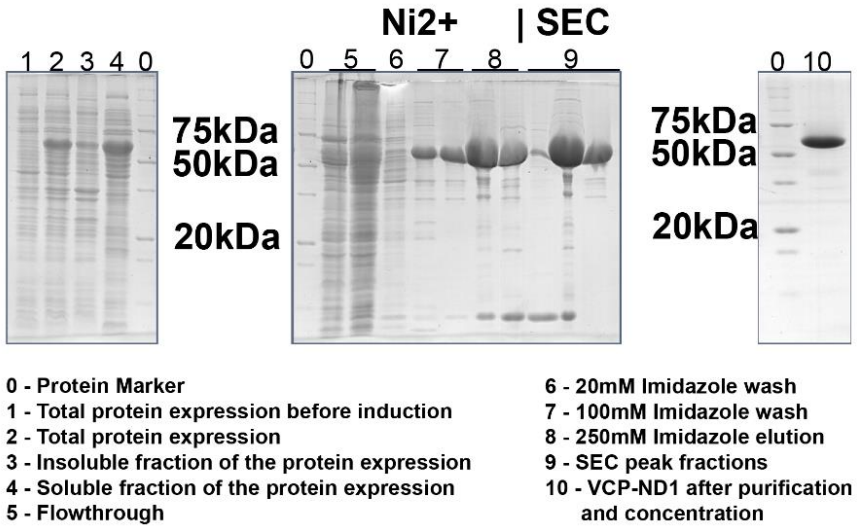

Figure 8 – Expression and purification of VCP-ND1 analyzed by SDS-PAGE. Samples were collected during protein expression and purification and electrophoretically separated on a 12% polyacrylamide SDS-PAGE gel.

3.1.3. Optimization of VCP-N expression and purification

Protein expression of the N-domain of VCP was performed according the methodologies specified and SDS-PAGE protein expression controls are shown in Fig. 9a.

A four-step protocol for the purification of VCP-N was established. Firstly, a Ni^{2+} affinity chromatography was performed (control SDS-PAGE on Fig. 9b) followed by a cleavage of the MBP solubility tag coupled with dialysis of the protein fractions eluted with 50mM Imidazole. A second Ni^{2+} affinity chromatography was performed to clean the MBP/6xHis-tag and fractions of eluted cleaved protein were analyzed by SDS-PAGE (Fig. 9c). Untagged VCP-N, was further purified by SEC on a Hiprep 26/60 Sephacryl S100 HR system - Fig. 9d (chromatogram is shown in the supplementary Fig. S3a). The eluted selected fractions were then concentrated, flash frozen and stored.

After purification by SEC, a clear ~40kDa contaminating protein could still be observed (Fig. 9d). We hypothesized that this band could correspond to the cleaved MBP/6xHis-tag (molecular weight is approximatively 40kDa). We performed a Western blot using a monoclonal anti-his mouse antibody, confirming that this 40kDa band was in fact recognized by the anti-his antibody (Fig. 9e). To further purify the VCP-N from this lingering MBP/6xHis-tag we performed an MPB affinity chromatography on an MBPTrap HP column – Fig. 9f (the respective chromatogram is shown in the supplementary Fig. S3b).

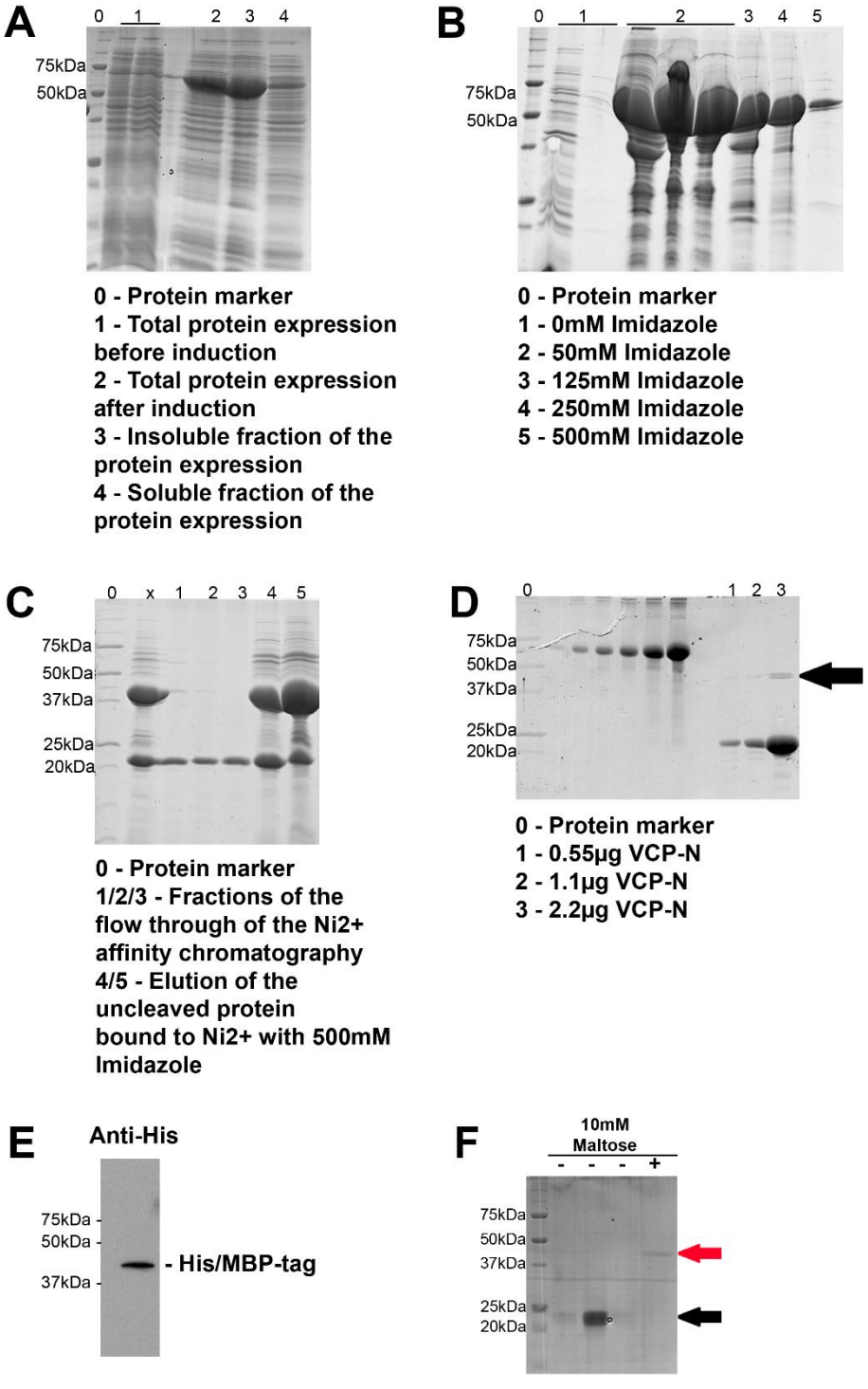

Figure 9 – Expression and purification of VCP-N. A) SDS-PAGE analysis of VCP-N expression. B) SDS-PAGE analysis of fractions eluted from the Ni^{2+} affinity chromatography. The fractions eluted with 50mM imidazole (fractions identified with 2) were selected and subjected to proteolytic cleavage of the MBP/6xHis-tag. C) After cleavage, the protein was further purified by a second Ni^{2+} affinity chromatography in which the cleaved MBP/6xHis-tag remained attached to the resin (eluted with 500mM imidazole in lanes 4 and 5), while tag-free VCP-N eluted in the flow-through (lanes 1, 2 and 3). D) The resulting VCP-N was further purified by SEC and the selected fractions subsequently analyzed on a 12% polyacrylamide SDS-PAGE gel to evaluate protein quality. Notice the ~40kDa (black arrow). E) Western blot shows that the ~40kDa is recognized by the anti-His antibody. F) Maltose affinity chromatography of the VCP-N sample was performed to remove the remaining MBP/6xHis-tag. Samples were collected during the chromatography and analyzed by electrophoresis. It is visible that the purification of VCP-N (first lanes and pointed with a black arrow) from the MBP/6xHis-tag (last lane and signaled with the red arrow) was successful.

3.1.4. Optimization of PNKP expression and purification

Expression of recombinant human PNKP (~57kDa) was performed according to the methodologies described before and a 12% polyacrylamide SDS-PAGE was used to assess protein expression (Fig. 10a). Soluble protein expression was enough for the downstream applications and therefore strategies to increase soluble expression were not applied.

For protein purification, a four-step protocol was devised. Due to the presence of the 6xHis-tag, a Ni^{2+} affinity chromatography was used as the first step of purification. Elution with 50mM and 100mM of imidazole proved to yield the highest amount of protein with the apparent molecular weight of PNKP (Fig. 10b). These fractions were selected and subjected to desalting to perform a subsequent cationic exchange chromatography (CEC). The cationic exchange of these samples was performed, and results are presented in Fig. 10c – 50mM Imidazole – and Fig. 10d – 100mM Imidazole. PNKP present in the 50mM and 100mM samples was eluted with 0.269mM NaCl and 0.276mM NaCl, respectively. Due to the similarities of the selected eluted fractions from both CECs, these selected samples (samples eluted from 10ml to 12ml - 3ml total) were mixed and further purified by SEC (Fig. 10e). The eluted fractions were analyzed and fractions 1 and 2 (Fig. 10e) were concentrated and frozen.

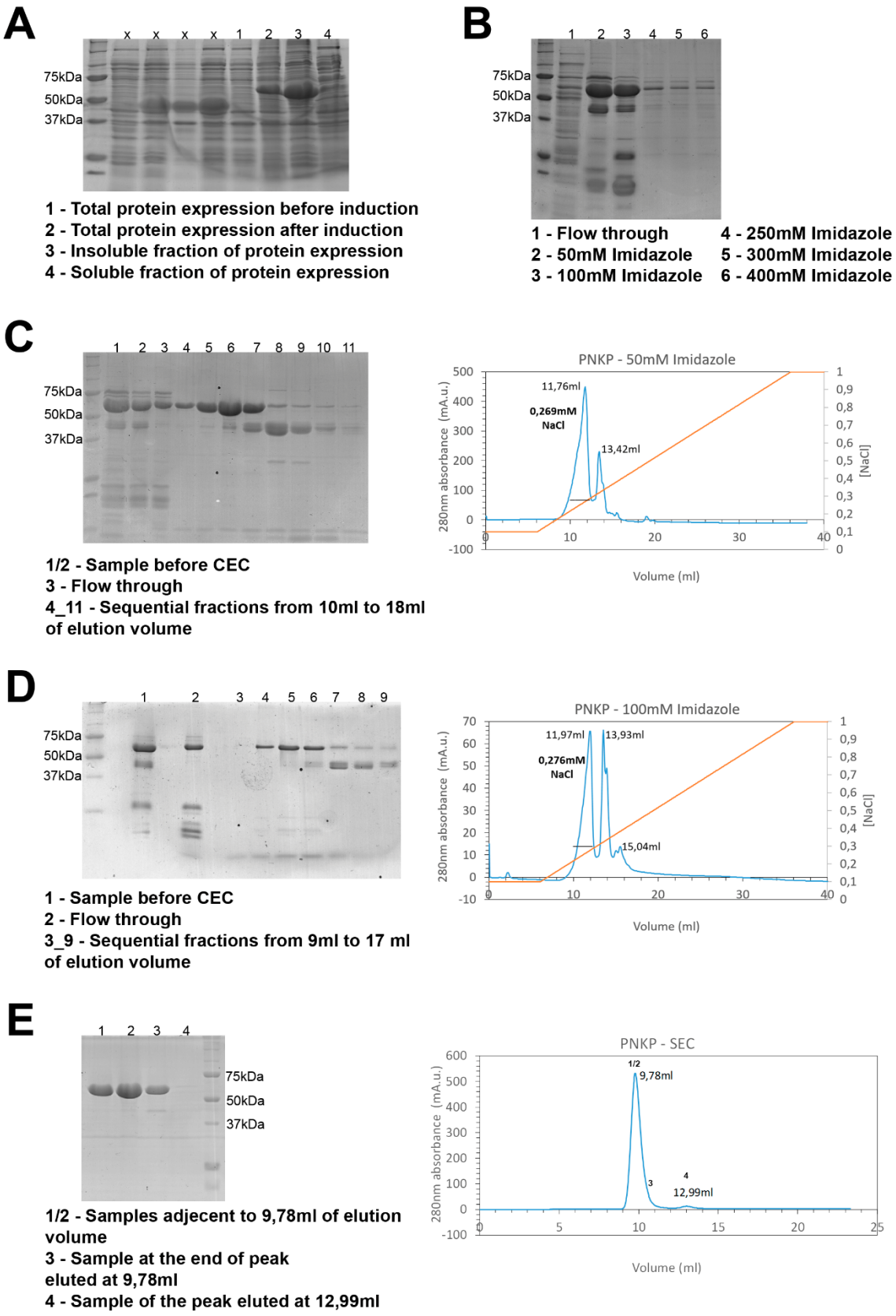

Figure 10 – Expression and purification of PNKP. A) PNKP expression profile evaluated on a 12% polyacrylamide SDS-PAGE. B) Ni^{2+} affinity chromatography of the PNKP expression protein extract. Samples eluted with 50mM Imidazole and 100mM Imidazole were selected and after desalting and applied to a CEC column. C) CEC of the desalted protein eluted

from the Ni^{2+} affinity chromatography with 50mM imidazole indicating the elution volumes of each peak and the concentration of NaCl of the gradient at the moment of elution of PNKP. The selected fractions for SEC are indicated with a line. D) CEC of the desalted protein eluted from the Ni^{2+} affinity chromatography with 100mM imidazole indicating the elution volumes of each peak and the concentration of NaCl of the gradient at the moment of elution of PNKP. The selected fractions for SEC are indicated with a line E) After the CECs the selected fractions (shown on each chromatogram) were further purified by SEC. The selected fractions corresponding to the top of the peak (marked with 1/2) were concentrated and frozen.

3.2. Atx3:VCP molecular basis of interaction

Previous data from the host laboratory has shown that Atx-3 and VCP-ND1 interact with an affinity in the nanomolar range (Kd=1.23nM) [68]. Other studies have also revealed that this interaction is mediated at least partly by a loop contiguous to the polyQ expandable tract of Atx3 [53].

To better map the interaction surfaces on Atx3, NMR spectroscopy was performed. NMR can be used to study protein-protein interactions at an atomic level. When two proteins interact, the chemical environment of the atoms at the interface of the complex is altered, leading to a shift in the resonances associated with these atoms. Selective observation of an individual component can be achieved by labeling only one protein with one isotope visible in NMR (e.g. ^{15}N) and leaving the other component invisible in NMR spectroscopy by the use of transparent isotopes (e.g. ^{14}N). Moreover, the formation of a protein-protein complex leads to an overall decrease in intensity of the spectrum since the average tumbling time of the protein complex is increased.

Previous unpublished data suggested the interaction between Atx3 and VCP may also occur through the JD. Therefore, we started by performing titrations of ^{15}N-labeled JD with non-labeled VCP-ND1 to try to identify chemical shifts on HSQC spectra of JD.

A comparison of the spectra of the Josephin domain without (Fig. 11a) and with VCP-ND1 (Fig. 11b) shows that the interaction between Atx3 and VCP-ND1 is not mediated by the Josephin domain of Atx3. The two spectra acquired overlap (Fig. 11c) showing no alteration of their intensity or chemical shifts of its peaks.

Due to these results, ongoing studies are occurring to characterize the interaction sites on the full-length Atx3 protein.

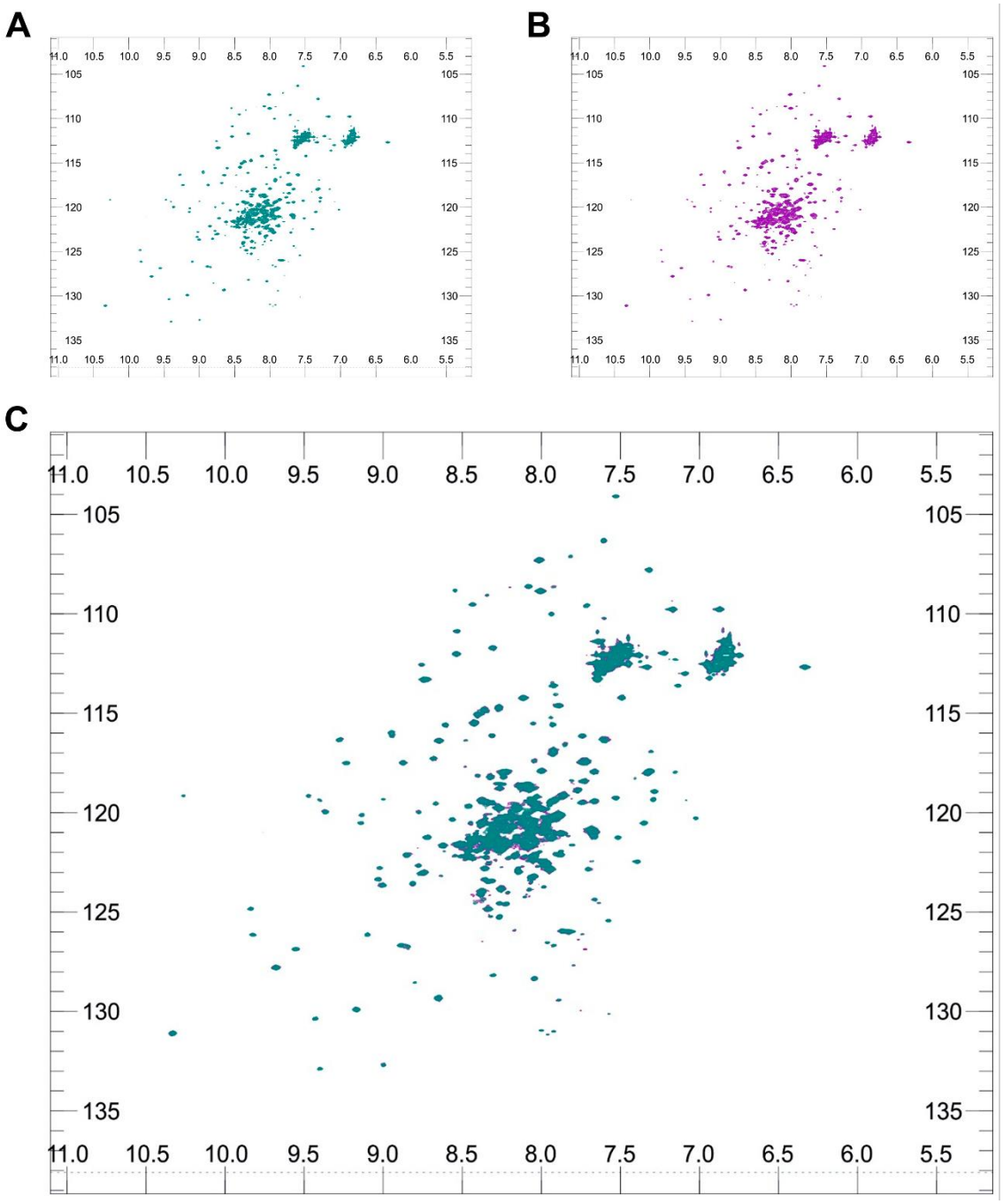

Figure 11 - Molecular basis of interaction between Atx3 and VCP. A) NMR HSQC spectrum of 0.2mM ^{15}N-Josephin. B) NMR HSQC spectrum of 0.2mM ^{15}N-Josephin titrated with 0.1mM VCP-ND1. C) Superimposition of the NMR HSQC spectrum of 0.2mM ^{15}N-Josephin with the NMR HSQC spectrum of the titration of 0.2mM ^{15}N-Josephin titrated with 0.1Mm VCP-ND1.

3.3. Ataxin-3 aggregation kinetics are modulated by its interaction partner VCP

3.3.1. Contribution of the different Ataxin-3 domains towards the modulation of Ataxin-3 aggregation kinetics

Thioflavin-T binding assays were performed to assess the influence of different domains (JD, D1, and non-expanded full-length Atx3 – Fig. 12a) of Atx3 in modulating its aggregation kinetics (Fig. 12). The thioflavin-T binding assay is based on the premise that thioflavin-T is a benzothiazole salt that binds to misfolded protein aggregates or amyloid fibrils and when bound, modifies its fluorescence spectrum, therefore being useful to quantify the formation of protein aggregates in real time [75, 76]. Specifically, thioflavin-T binding monitors the first step of aggregation and therefore the formation of amyloid protofibrils [77]. Since Atx3 aggregation is influenced greatly by the previous existence of nucleation points of aggregation [63], before thioflavin-T binding assays each Atx3 protein was re-purified by analytical SEC to remove any aggregation nucleation points (dimers, oligomers) that may have formed due to the freezing and thawing of the samples. Characteristic SEC chromatograms for these re-purification steps are shown in the supplementary Fig. S1b).

In the aggregation assay (shown in Fig. 12b), it is possible to see that the lag phase of aggregation curves of Atx3 and its truncated versions is similar. However, a higher fluorescence emission it is observable for 13Q-Atx3 and D1 when compared with the isolated JD. Despite the equal protein amount (5µM) of each reaction, JD is shown only to reach approximately half of aggregation potential of full-length Atx3. Furthermore, the construct lacking the polyQ-containing C-terminal region of Atx3 (D1 = JD+UIM1/2), has an aggregation behavior comparable to that of the non-expanded full-length protein, indicating that early aggregation steps might be mediated by other regions outside the JD domain, namely the UIM1/2 regions.

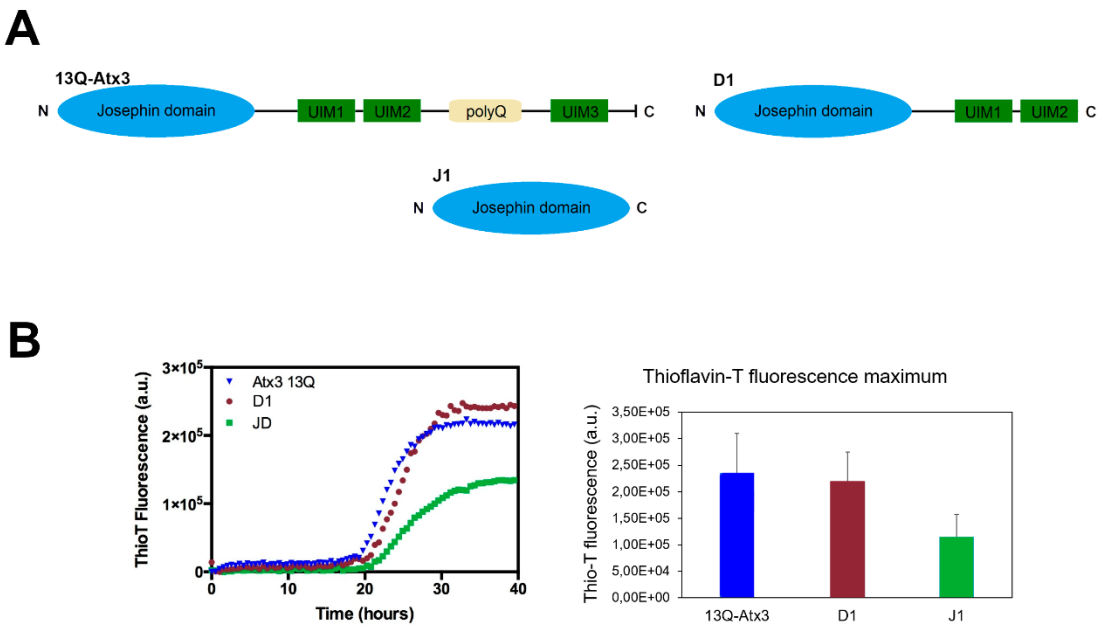

Figure 12 – The influence of Atx3 domains on its aggregation kinetics. A) Scheme of the several constructs of Atx3 used. B) Atx3 aggregation was assessed through thioflavin-T binding assays using 5μM of each protein with 30μM of thioflavin-T. On the representative assay (left panel) it is observable a similar aggregation kinetics for 13Q-Atx3 and D1 (blue triangles and brown circles). However, for JD (green squares) even though the lag phase is akin the slope and maximum fluorescence signal reached is diminished in comparison with the other variants of Atx3. On the right panel, the average of the maximum fluorescence achieved in three independent assays by each version of Atx3 is represented along with the standard deviation of these means.

3.3.2. Modulation of Ataxin-3 aggregation by VCP

Considering the hypothesized protective role of Atx3 interactions in modulating aggregation [43, 64] and the known interaction between Atx3 and VCP-ND1 [53], we tried to better characterize the interaction between those proteins and to assess the effect of this interaction on Atx3 aggregation. For this purpose, we used VCP-ND1 – an hexameric truncated version of VCP encompassing the domains N and D1 – and the domain N alone.

To study the effect of VCP on Atx3 aggregation we performed thioflavin-T binding assays (Fig. 13a). When 5μM 13Q-Atx3 was incubated in 5μM of VCP-N and 5μM of VCP-ND1 (in its hexameric form) a decrease on Atx3 aggregation was observed. The anti-aggregation effect is similar for 77Q-Atx3. Although addition of VCP-N to Atx3 aggregation assay contributes to a reduction the maximum fibril load (reflected by a decrease in the fluorescence maximum in the saturation phase), no significant changes are observed in the total duration of the lag phase. In contrast, the addition of VCP-ND1 completely decreases almost completely Atx3 aggregation both for expanded and non-

expanded protein, abolishing it for the non-expanded variant. The mean of the maximum of thioflavin-T fluorescence and the standard deviation registered for each incubation in three independent assays is shown in the third panel of Fig.13a.

Due to the observed modulation of the first phase of Atx3 aggregation kinetics caused by the incubation with VCP-ND1 and VCP-N, we analyzed the effect on the second phase of aggregation - the maturation of the amyloid fibrils in the pathogenic 77Q-Atx3. In order to do that, a filter retardation assay (also known as slot blot) was performed (Fig. 13b). As a negative control 13Q-Atx3 was used (since it does not undergo the second step of aggregation) and for a positive control, the mature SDS-resistant amyloid fibrils formed during then thioflavin-T assay by 77Q-Atx3 were used. 77Q-Atx3 incubation with VCP-N and with VCP-ND1 exhibits a diminished formation of mature SDS-resistant amyloid fibrils. Although slight, the reduction is noticeable when 77Q-Atx3 is incubated with VCP-N. When 77Q-Atx3 is incubated with VCP-ND1, the formation of mature SDS-resistant amyloid fibrils is completely abolished.

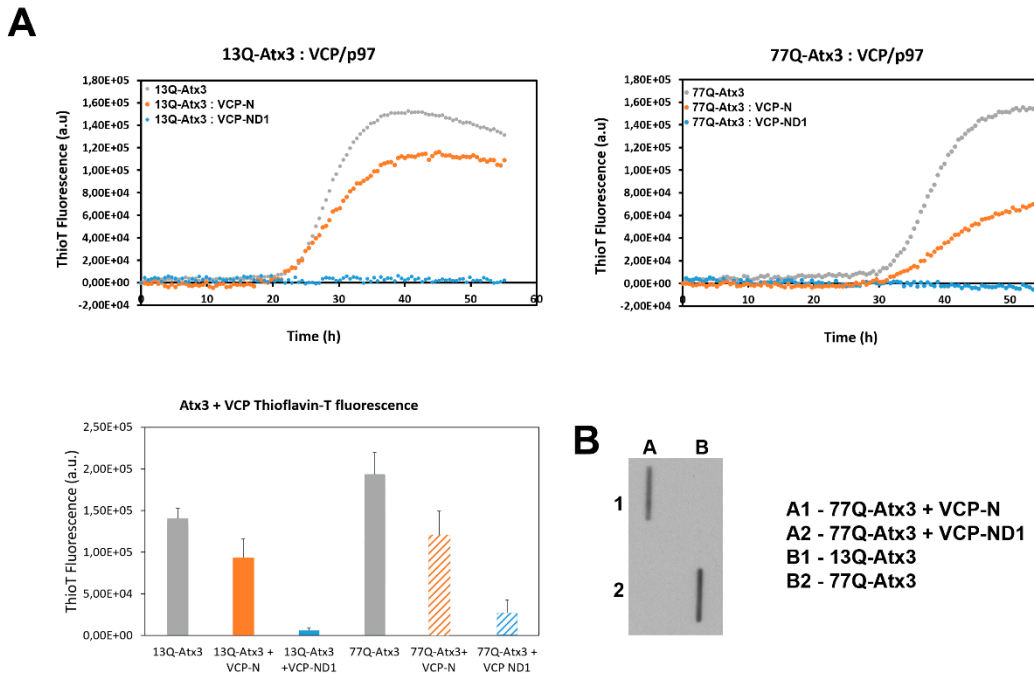

Figure 13 – Modulation of Atx3 aggregation by incubation with VCP proteins. A) The effects of VCP-N and VCP-ND1 on the kinetics of the first phase of aggregation of 13Q-Atx3 and 77Q-Atx3 were monitored by thioflavin-T binding assays. On the top left panel is observable that the equimolar incubation of 13Q-Atx3 + VCP-N and 13Q-Atx3 + VCP-ND1 reduces 13Q-Atx3 aggregation. Even though the lag phase is not affected, the maximum of aggregation is reduced by the incubation with VCP-N and greatly reduced in the presence of VCP-ND1. On the top right, the equimolar incubation of 77Q-Atx3 + VCP-N and 77Q-Atx3 + VCP-ND1 although not affecting the kinetics lag phase reduces 77Q-Atx3 aggregation. A noteworthy decrease on 77Q-Atx3 aggregation is observable with VCP-ND1. The basal signal obtained from the incubation of 5µM of VCP-N and 5µM VCP-ND1 in the thioflavin-T assay was subtracted to the overall signal of

their incubations with 13Q-Atx3 and 77Q-Atx3. On the bottom left panel, the mean of the maximum fluorescence observed in three independent assays is shown as well as its standard deviation. B) Filter retardation assay. On A2 (the 13Q-Atx3 negative control) no signal is observed. B2 corresponds to the 77Q-Atx3 positive control. A reduced amount of mature fibrils is observed on A1 – the interaction between 77Q-Atx3 and VCP-N. On A2 no signal is observable as there was no formation of mature 77Q-Atx3 amyloid fibrils.

- DISCUSSION -

4. Discussion

All proteins were successfully expressed and purified, and protocols were established for the expression and purification of VCP-N and PNKP.

4.1. Although central, Josephin domain is not the sole contributor for Ataxin-3 first step of aggregation

Towards deepening the knowledge of the effect of the other N-terminal domains on the first step of aggregation of Atx3 (independent of the polyQ expansion) we studied the effect of the N-terminal domains on Atx3 aggregation kinetics by thioflavin-T binding assays. We did not observe any differences between the lag phases of the kinetics of aggregation of the several domains. Nevertheless, differences in the maximum thioflavin-T fluorescence signal were observed uncovering the differential impact of Atx3 domains on the kinetics of its first aggregation step. Although a significant aggregation was observed on the JD *per se* (underlying the already known central role on Atx3 first step of aggregation), the maximum of aggregation displayed fell short of the maximum displayed by 13Q-Atx3. However, when D1 aggregation is observed, an aggregation profile similar to that of 13Q-Atx3 is found. This information suggests that in addition to JD the UIM1/2 also plays a role on Atx3 first step of aggregation. Therefore, macromolecular interactors targeting regions outside the JD domain might also have an impact on the early steps of Atx3 aggregation.

4.2. Atx3:VCP interaction may block Ataxin-3 aggregation mechanisms

Contrary to previous unpublished data available in the group, the interaction between Atx3 and VCP does not occur through the Josephin domain. Nonetheless, these results do not refute previous publications [53, 68]. The HSQC NMR spectra of the titrations reveal no chemical shifts of any peak of JD nor diminishment of their intensity, and thus, there is no detectable interaction between JD and VCP-ND1. This result, therefore, invalidates future attempts at mapping the interaction of Atx3 with VCP-ND1 through this domain. Subsequent studies of these yet non-unveiled interaction surfaces cannot also be addressed through this domain, therefore, forcing such studies to be done with the

full-length protein. This discrepancy between the previous data and this new data may have occurred due to the acquiring of the previous chemical shift data before the clear assignment of the Atx3 NMR spectra.

In the future, NMR spectra of the titration between 13Q-Atx3: VCP-ND1 shall be acquired as well as the 13Q-Atx3 HSQC NMR spectra shall be assigned in order to completely map this interaction to understand it better and therefore devise strategies to study it and to design strategies to mimic the observed effect of VCP-ND1 on Atx3 aggregation kinetics. Eventually, this data could also be used to design eventual therapeutic approaches.

For the aggregation assays of Atx3, the N-domain of VCP was also used. This was due to three reasons:

1. Most VCP interactions are mediated by the N domain of VCP [78-80];
2. The substantial size of VCP-ND1 in its hexameric form (~350kDa) may create an effect of macromolecular crowding and therefore affect protein stability and subsequently aggregation [81] hence adulterating results.
3. Since we also wanted to precisely map the interaction surfaces of the interaction between Atx3 and VCP (which may be useful to develop smaller molecules that could act as anti-aggregation molecules) the use of VCP-ND1 can be a problem. A known problem of protein NMR is the reduced, broad signal and therefore the low signal-to-noise ratio of NMR spectra involving large proteins (such as the hexameric VCP-ND1 - ~350kDa) due to the slower tumbling rate of these proteins and therefore the faster transverse and longitudinal relaxations [82].

Considering the abovementioned reasons, if we could register the same interaction between Atx3:VCP-ND1 and Atx3:VCP-N, it would be advantageous for the described scenarios due to the non-hexamerization and size of the N-domain *per se*.

On the thioflavin-T binding aggregation assays, we observed that equimolar incubation of hexameric VCP-ND1 with Atx3 (either containing or lacking the non-pathological polyQ expansion) decreases Atx3 aggregation to levels close to zero in the time frame of the assay (~50h). Although quite promising, these observed results could occur due to the phenomenon of macromolecular crowding as it is described that can affect protein stability and subsequently aggregation [81]. However, a contradicting point to this hypothesis is the results obtained from the incubation with VCP-N. Although not as good as those of VCP-ND1, the observed effect of VCP-N on Atx3 aggregation

provides us with some information. If the small molecular weight of this domain (~20kDa) and the fact that it is not capable of hexamerization (unlike VCP-ND1) is considered, VCP-N provides hints against the hypothesis that the observed effect of VCP-ND1 over Atx3 aggregation is due to macromolecular crowding. Nevertheless, it also shows us that alone it is capable of reducing Atx3 aggregation.

Nonetheless, to prove this hypothesis, further experiments have to be done. This hypothesis can be tested by using macromolecular crowding agents in these assays and evaluate their effect on Atx3 aggregation. Ideally, these crowding agents should be proteins of similar size to that of hexameric VCP-ND1 (~350kDa) that are not Atx3 interactors. Another option is to use Atx3 mutants with the VCP interaction loop [53] mutated and therefore not capable of binding to VCP.

To evaluate the effects of VCP interaction with Atx3 on its second step of aggregation (the formation of mature SDS-resistant amyloid fibrils), samples collected at the end-point of Thioflavin-T aggregation assay were used to perform filter retardation assays. The observable lack of aggregation of 77Q-Atx3 incubated with VCP-ND1 on the first step of aggregation is translated to the filter retardation assay where no detectable SDS-resistant mature amyloid fibrils are observed. This lack of mature amyloid fibrils is explained by the almost complete abolishment of 77Q-Atx3 first step of aggregation (as monitored in the thioflavin-T binding assay) and therefore by the lack of soluble amyloid protofibrils available and capable of undergoing the second step of aggregation and therefore the maturation of mature SDS-resistant amyloid fibrils.

Although preliminary, these results point towards the idea that VCP may be protective against aggregation and may also provide a further link between functional interactions and their protective role against aggregation. If these results are confirmed to be consequence of an interaction between VCP and Atx3 they can also lead to the hypothesis that Atx3 aggregation can be suppressed or delayed in the first step of aggregation, preventing the formation of the toxic Atx3 oligomers [16, 71, 83] and therefore preventing another pathological mechanism as well as reducing cellular stress which if possible may be translated into therapeutic approaches and treatments for the disease.

- CONCLUSION / FUTURE PERSPECTIVES -

5. Conclusion / Future Perspectives

Even though many of the results obtained during this year are preliminary, they are quite interesting and shall be further investigated.

Protocols for the expression and purification of several proteins were devised and successful, therefore will be used in future studies.

The assessment of the effects of VCP on Atx3 aggregation revealed a possible modulatory role of VCP over Atx3 aggregation. In the future, this hypothesis will be tested according to the strategies devised and explained in the discussion section.

Although a possible methodology for the assessment of the effects of incubation with VCP-ND1 and VCP-N on Atx3 aggregation, the visualization of the protein fibrils by TEM (Transmission Electron Microscopy) to evaluate their morphology and eventual modifications of said morphology, was unsuccessful probably due to a subpar sample preparation (paraffin oil contaminated the TEM grids). In future experiments, this analysis of the morphological differences on the aggregated fibrils in function of their incubation partners will be accomplished.

In the future, the unveiling of the molecular interaction surfaces of the interaction between Atx3 and VCP by NMR spectroscopy on Atx3 shall also be assessed.

Finally, even though the PNKP interaction with Atx3 was not yet investigated, future biophysical experiments will approach this matter to try to characterize this interaction. Subsequently, if the interaction is proven to exist and if reasonable, the effects of it on Atx3 aggregation shall be assessed by thioflavin-T binding assays.

- BIBLIOGRAPHY -

6. Bibliography

1. Lehninger, A.L., D.L. Nelson, and M.M. Cox, *Lehninger principles of biochemistry*. 6th ed. 2013, New York: W.H. Freeman.

2. Sali, A., E. Shakhnovich, and M. Karplus, *How does a protein fold?* Nature, 1994. **369**(6477): p. 248-51.

3. Hartl, F.U., A. Bracher, and M. Hayer-Hartl, *Molecular chaperones in protein folding and proteostasis.* Nature, 2011. **475**(7356): p. 324-32.

4. Olzscha, H., et al., *Amyloid-like aggregates sequester numerous metastable proteins with essential cellular functions.* Cell, 2011. **144**(1): p. 67-78.

5. Philo, J.S., *Is any measurement method optimal for all aggregate sizes and types?* AAPS J, 2006. **8**(3): p. E564-71.

6. Wang, W., *Protein aggregation and its inhibition in biopharmaceutics.* Int J Pharm, 2005. **289**(1-2): p. 1-30.

7. Powers, E.T., et al., *Biological and chemical approaches to diseases of proteostasis deficiency.* Annu Rev Biochem, 2009. **78**: p. 959-91.

8. Hartl, F.U. and M. Hayer-Hartl, *Molecular chaperones in the cytosol: from nascent chain to folded protein.* Science, 2002. **295**(5561): p. 1852-8.

9. Chen, B., et al., *Cellular strategies of protein quality control.* Cold Spring Harb Perspect Biol, 2011. **3**(8): p. a004374.

10. Kaganovich, D., R. Kopito, and J. Frydman, *Misfolded proteins partition between two distinct quality control compartments.* Nature, 2008. **454**(7208): p. 1088-95.

11. Manning, M.C., et al., *Stability of protein pharmaceuticals: an update.* Pharm Res, 2010. **27**(4): p. 544-75.

12. Stefani, M. and C.M. Dobson, *Protein aggregation and aggregate toxicity: new insights into protein folding, misfolding diseases and biological evolution.* J Mol Med (Berl), 2003. **81**(11): p. 678-99.

13. Margulis, B.A., et al., *Pharmacological protein targets in polyglutamine diseases: mutant polypeptides and their interactors.* FEBS Lett, 2013. **587**(13): p. 1997-2007.

14. Gatchel, J.R. and H.Y. Zoghbi, *Diseases of unstable repeat expansion: mechanisms and common principles.* Nat Rev Genet, 2005. **6**(10): p. 743-55.

15. Matos, C.A., S. de Macedo-Ribeiro, and A.L. Carvalho, *Polyglutamine diseases: the special case of ataxin-3 and Machado-Joseph disease.* Prog Neurobiol, 2011. **95**(1): p. 26-48.

16. Williams, A.J. and H.L. Paulson, *Polyglutamine neurodegeneration: protein misfolding revisited.* Trends Neurosci, 2008. **31**(10): p. 521-8.

17. Fan, H.C., et al., *Polyglutamine (PolyQ) diseases: genetics to treatments.* Cell Transplant, 2014. **23**(4-5): p. 441-58.

18. Zoghbi, H.Y. and H.T. Orr, *Glutamine repeats and neurodegeneration.* Annu Rev Neurosci, 2000. **23**: p. 217-47.

19. Orr, H.T. and H.Y. Zoghbi, *Trinucleotide repeat disorders.* Annu Rev Neurosci, 2007. **30**: p. 575-621.

20. Maciel, P., et al., *Correlation between CAG repeat length and clinical features in Machado-Joseph disease.* Am J Hum Genet, 1995. **57**(1): p. 54-61.

21. Costa M. do, C. and H.L. Paulson, *Toward understanding Machado-Joseph disease.* Prog Neurobiol, 2012. **97**(2): p. 239-57.

22. Nakano, K.K., D.M. Dawson, and A. Spence, *Machado disease. A hereditary ataxia in Portuguese emigrants to Massachusetts.* Neurology, 1972. **22**(1): p. 49-55.

23. Schols, L., et al., *Autosomal dominant cerebellar ataxias: clinical features, genetics, and pathogenesis.* Lancet Neurol, 2004. **3**(5): p. 291-304.

24. Paulson, H., *Machado-Joseph disease/spinocerebellar ataxia type 3.* Handb Clin Neurol, 2012. **103**: p. 437-49.

25. Wakap, S.S.N., *«Prevalence of rare diseases: Bibliographic data », Orphanet Report Series, Rare Diseases collection, March 2016, Number 1 : Diseases listed in alphabetical order* *http://www.orpha.net/orphacom/cahiers/docs/GB/Prevalence_of_rare_diseases_by_diseases.pdf*. 2016.

26. Araújo, M.A.d., et al., *Trends in the Epidemiology of Spinocerebellar Ataxia Type 3/Machado-Joseph Disease in the Azores Islands, Portugal.* JSM Brain Science, 2016. **1**(1: (1001)).

27. Riess, O., et al., *SCA3: neurological features, pathogenesis and animal models.* Cerebellum, 2008. **7**(2): p. 125-37.

28. Rosenberg, R.N., *Machado-Joseph disease: an autosomal dominant motor system degeneration.* Mov Disord, 1992. **7**(3): p. 193-203.

29. Soong, B., et al., *Machado-Joseph disease: clinical, molecular, and metabolic characterization in Chinese kindreds.* Ann Neurol, 1997. **41**(4): p. 446-52.

30. Sudarsky, L. and P. Coutinho, *Machado-Joseph disease.* Clin Neurosci, 1995. **3**(1): p. 17-22.

31. Sequeiros, J. and P. Coutinho, *Epidemiology and clinical aspects of Machado-Joseph disease.* Adv Neurol, 1993. **61**: p. 139-53.

32. Schmidt, T., et al., *An isoform of ataxin-3 accumulates in the nucleus of neuronal cells in affected brain regions of SCA3 patients.* Brain Pathol, 1998. **8**(4): p. 669-79.

33. Costa M do, C. and H.L. Paulson, *Toward understanding Machado-Joseph disease.* Prog Neurobiol, 2012. **97**(2): p. 239-57.

34. Yamada, M., et al., *CAG repeat disorder models and human neuropathology: similarities and differences.* Acta Neuropathol, 2008. **115**(1): p. 71-86.

35. Takiyama, Y., et al., *The gene for Machado-Joseph disease maps to human chromosome 14q.* Nat Genet, 1993. **4**(3): p. 300-4.

36. Kawaguchi, Y., et al., *CAG expansions in a novel gene for Machado-Joseph disease at chromosome 14q32.1.* Nat Genet, 1994. **8**(3): p. 221-8.

37. Cummings, C.J. and H.Y. Zoghbi, *Fourteen and counting: unraveling trinucleotide repeat diseases.* Hum Mol Genet, 2000. **9**(6): p. 909-16.

38. Maciel, P., et al., *Improvement in the molecular diagnosis of Machado-Joseph disease.* Arch Neurol, 2001. **58**(11): p. 1821-7.

39. Ichikawa, Y., et al., *The genomic structure and expression of MJD, the Machado-Joseph disease gene.* J Hum Genet, 2001. **46**(7): p. 413-22.

40. Bettencourt, C. and M. Lima, *Machado-Joseph Disease: from first descriptions to new perspectives.* Orphanet J Rare Dis, 2011. **6**: p. 35.

41. Masino, L., et al., *Domain architecture of the polyglutamine protein ataxin-3: a globular domain followed by a flexible tail.* FEBS Lett, 2003. **549**(1-3): p. 21-5.

42. Nicastro, G., et al., *The solution structure of the Josephin domain of ataxin-3: structural determinants for molecular recognition.* Proc Natl Acad Sci U S A, 2005. **102**(30): p. 10493-8.

43. Nicastro, G., et al., *Josephin domain of ataxin-3 contains two distinct ubiquitin-binding sites.* Biopolymers, 2009. **91**(12): p. 1203-14.

44. Song, A.X., et al., *Structural transformation of the tandem ubiquitin-interacting motifs in ataxin-3 and their cooperative interactions with ubiquitin chains.* PLoS One, 2010. **5**(10): p. e13202.

45. Scheel, H., S. Tomiuk, and K. Hofmann, *Elucidation of ataxin-3 and ataxin-7 function by integrative bioinformatics.* Hum Mol Genet, 2003. **12**(21): p. 2845-52.

46. Burnett, B., F. Li, and R.N. Pittman, *The polyglutamine neurodegenerative protein ataxin-3 binds polyubiquitylated proteins and has ubiquitin protease activity.* Hum Mol Genet, 2003. **12**(23): p. 3195-205.

47. Mao, Y., et al., *Deubiquitinating function of ataxin-3: insights from the solution structure of the Josephin domain.* Proc Natl Acad Sci U S A, 2005. **102**(36): p. 12700-5.

48. Winborn, B.J., et al., *The deubiquitinating enzyme ataxin-3, a polyglutamine disease protein, edits Lys63 linkages in mixed linkage ubiquitin chains.* J Biol Chem, 2008. **283**(39): p. 26436-43.

49. Erpapazoglou, Z., O. Walker, and R. Haguenauer-Tsapis, *Versatile roles of k63-linked ubiquitin chains in trafficking.* Cells, 2014. **3**(4): p. 1027-88.

50. Berke, S.J., et al., *Defining the role of ubiquitin-interacting motifs in the polyglutamine disease protein, ataxin-3.* J Biol Chem, 2005. **280**(36): p. 32026-34.

51. Nicastro, G., et al., *Understanding the role of the Josephin domain in the PolyUb binding and cleavage properties of ataxin-3.* PLoS One, 2010. **5**(8): p. e12430.

52. Hirabayashi, M., et al., *VCP/p97 in abnormal protein aggregates, cytoplasmic vacuoles, and cell death, phenotypes relevant to neurodegeneration.* Cell Death Differ, 2001. **8**(10): p. 977-84.

53. Boeddrich, A., et al., *An arginine/lysine-rich motif is crucial for VCP/p97-mediated modulation of ataxin-3 fibrillogenesis.* EMBO J, 2006. **25**(7): p. 1547-58.

54. Wang, G., et al., *Ataxin-3, the MJD1 gene product, interacts with the two human homologs of yeast DNA repair protein RAD23, HHR23A and HHR23B.* Hum Mol Genet, 2000. **9**(12): p. 1795-803.

55. Doss-Pepe, E.W., et al., *Ataxin-3 interactions with rad23 and valosin-containing protein and its associations with ubiquitin chains and the proteasome are consistent with a role in ubiquitin-mediated proteolysis.* Mol Cell Biol, 2003. **23**(18): p. 6469-83.

56. Todi, S.V., et al., *Cellular turnover of the polyglutamine disease protein ataxin-3 is regulated by its catalytic activity.* J Biol Chem, 2007. **282**(40): p. 29348-58.

57. Ouyang, H., et al., *Protein aggregates are recruited to aggresome by histone deacetylase 6 via unanchored ubiquitin C termini.* J Biol Chem, 2012. **287**(4): p. 2317-27.

58. Burnett, B.G. and R.N. Pittman, *The polyglutamine neurodegenerative protein ataxin 3 regulates aggresome formation.* Proc Natl Acad Sci U S A, 2005. **102**(12): p. 4330-5.

59. Rodrigues, A.J., et al., *Absence of ataxin-3 leads to cytoskeletal disorganization and increased cell death.* Biochim Biophys Acta, 2010. **1803**(10): p. 1154-63.

60. Mazzucchelli, S., et al., *Proteomic and biochemical analyses unveil tight interaction of ataxin-3 with tubulin.* Int J Biochem Cell Biol, 2009. **41**(12): p. 2485-92.

61. Johnston, J.A., C.L. Ward, and R.R. Kopito, *Aggresomes: a cellular response to misfolded proteins.* J Cell Biol, 1998. **143**(7): p. 1883-98.

62. Todi, S.V., et al., *Activity and cellular functions of the deubiquitinating enzyme and polyglutamine disease protein ataxin-3 are regulated by ubiquitination at lysine 117.* J Biol Chem, 2010. **285**(50): p. 39303-13.

63. Scarff, C.A., et al., *Examination of Ataxin-3 (atx-3) Aggregation by Structural Mass Spectrometry Techniques: A Rationale for Expedited Aggregation upon Polyglutamine (polyQ) Expansion.* Mol Cell Proteomics, 2015. **14**(5): p. 1241-53.

64. Masino, L., et al., *The Josephin domain determines the morphological and mechanical properties of ataxin-3 fibrils.* Biophys J, 2011. **100**(8): p. 2033-42.

65. Almeida, B., et al., *Trinucleotide repeats: a structural perspective.* Front Neurol, 2013. **4**: p. 76.

66. Invernizzi, G., et al., *The relationship between aggregation and toxicity of polyglutamine-containing ataxin-3 in the intracellular environment of Escherichia coli.* PLoS One, 2012. **7**(12): p. e51890.

67. Masino, L., et al., *Functional interactions as a survival strategy against abnormal aggregation.* FASEB J, 2011. **25**(1): p. 45-54.

68. Almeida, B., et al., *SUMOylation of the brain-predominant Ataxin-3 isoform modulates its interaction with p97.* Biochim Biophys Acta, 2015. **1852**(9): p. 1950-9.

69. Reynolds, J.J., et al., *Impact of PNKP mutations associated with microcephaly, seizures and developmental delay on enzyme activity and DNA strand break repair.* Nucleic Acids Res, 2012. **40**(14): p. 6608-19.

70. Chatterjee, A., et al., *The role of the mammalian DNA end-processing enzyme polynucleotide kinase 3'-phosphatase in spinocerebellar ataxia type 3 pathogenesis.* PLoS Genet, 2015. **11**(1): p. e1004749.

71. Gales, L., et al., *Towards a structural understanding of the fibrillization pathway in Machado-Joseph's disease: trapping early oligomers of non-expanded ataxin-3.* J Mol Biol, 2005. **353**(3): p. 642-54.

72. Kang, W. and J.K. Yang, *Crystallization and preliminary X-ray crystallographic analysis of the hexameric human p97/VCP ND1 fragment in complex with the UBX domain of human FAF1.* Acta Crystallogr Sect F Struct Biol Cryst Commun, 2011. **67**(Pt 10): p. 1199-202.

73. Delaglio, F., et al., *NMRPipe: a multidimensional spectral processing system based on UNIX pipes.* J Biomol NMR, 1995. **6**(3): p. 277-93.

74. Vranken, W.F., et al., *The CCPN data model for NMR spectroscopy: development of a software pipeline.* Proteins, 2005. **59**(4): p. 687-96.

75. Hawe, A., M. Sutter, and W. Jiskoot, *Extrinsic fluorescent dyes as tools for protein characterization.* Pharm Res, 2008. **25**(7): p. 1487-99.

76. Bolder, S.G., et al., *Thioflavin T and birefringence assays to determine the conversion of proteins into fibrils.* Langmuir, 2007. **23**(8): p. 4144-7.

77. Ellisdon, A.M., B. Thomas, and S.P. Bottomley, *The two-stage pathway of ataxin-3 fibrillogenesis involves a polyglutamine-independent step.* J Biol Chem, 2006. **281**(25): p. 16888-96.

78. Song, E.J., et al., *Human Fas-associated factor 1, interacting with ubiquitinated proteins and valosin-containing protein, is involved in the ubiquitin-proteasome pathway.* Mol Cell Biol, 2005. **25**(6): p. 2511-24.

79. Ye, Y., H.H. Meyer, and T.A. Rapoport, *The AAA ATPase Cdc48/p97 and its partners transport proteins from the ER into the cytosol.* Nature, 2001. **414**(6864): p. 652-6.

80. Kondo, H., et al., *p47 is a cofactor for p97-mediated membrane fusion.* Nature, 1997. **388**(6637): p. 75-8.

81. Sanfelice, D., et al., *The effect of crowding and confinement: a comparison of Yfh1 stability in different environments.* Phys Biol, 2013. **10**(4): p. 045002.

82. Frueh, D.P., *Practical aspects of NMR signal assignment in larger and challenging proteins.* Prog Nucl Magn Reson Spectrosc, 2014. **78**: p. 47-75.

83. Shao, J. and M.I. Diamond, *Polyglutamine diseases: emerging concepts in pathogenesis and therapy.* Hum Mol Genet, 2007. **16 Spec No. 2**: p. R115-23.

- APPENDICES -

7. Appendices

7.1. Appendix I

7.1.1. Supplementary Figures

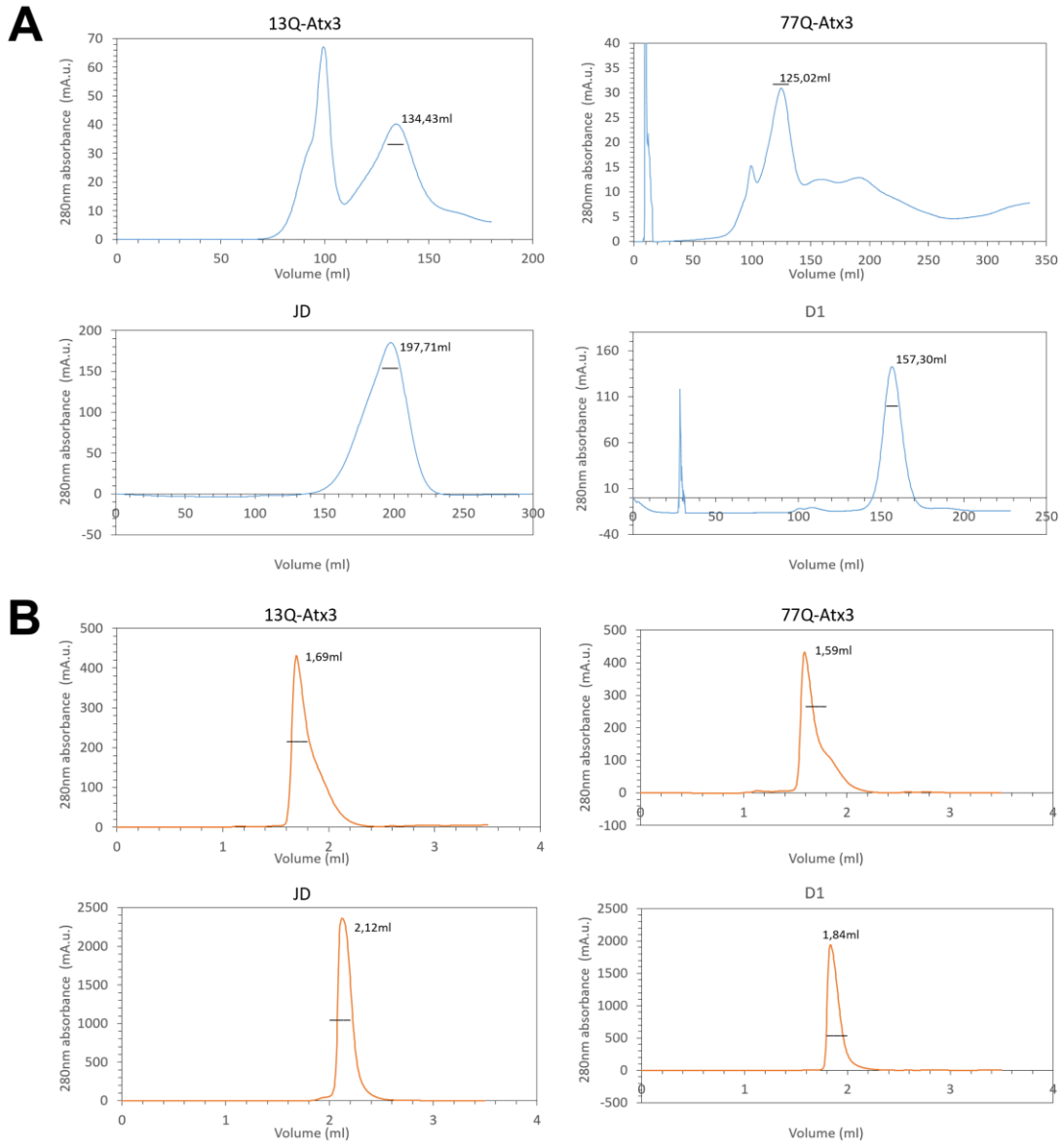

Figure S1 – Purification of Atx3 proteins. A) Typical SEC chromatograms for protein purification on a Hiprep 26/60 Sephacryl S200 HR gel filtration system of 13Q-Atx3, 77Q-Atx3, JD, and D1. Elution volume (in ml) plotted against 280nm absorbance (in mA.u.). Selected fractions represented by a line over the eluted peaks. B) Typical SEC chromatograms for protein quality control and final clean-up step of protein aggregates before protein aggregation assays on a Superdex 200 5/150 GL gel filtration system of 13Q-atx3, 77Q-atx3, JD, and D1. Elution volume (in ml) plotted against 280nm absorbance (in mA.u. – miliAbsorbance units). Selected fractions represented by a line over the eluted peaks.

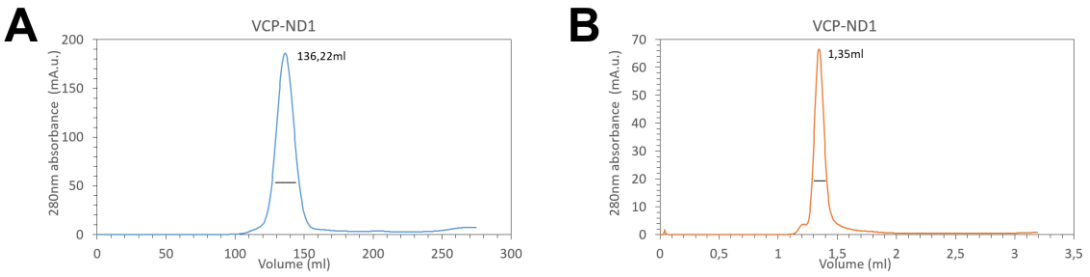

Figure S2 – Purification of VCP-ND1. A) Characteristic chromatogram of the preparative SEC of VCP-ND1 on a Hiprep 26/60 Sephacryl S300 HR gel filtration system. Selected eluted fractions are evidenced with a line. B) Chromatogram of a characteristic analytical SEC of purified VCP-ND1 on a Superdex 200 5/150 GL gel filtration system. Selected eluted fractions are indicated with a line.

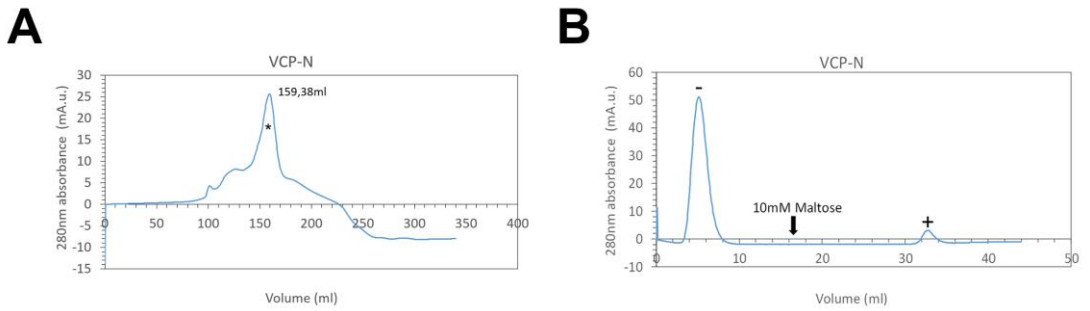

Figure S3 – Expression and purification of VCP-N. A) Chromatogram of the preparative SEC performed in order to further purify VCP-N after the MBP/6xHis-tag cleavage and cleaning (selected fractions are marked with a *). B) Chromatogram of the Maltose affinity chromatography performed on the VCP-N sample to remove the lingering MBP/6xHis-tag. The arrow asserts the elution point were maltose was added and therefore after which the MBP/6xHis-tag was eluted. The selected fractions (- and +) were analyzed by electrophoresis (Fig. 9e).