Università degli Studi di Palermo

AREA RICERCA E TRASFERIMENTO TECNOLOGICO
SETTORE DOTTORATI E CONTRATTI PER LA RICERCA
U. O. DOTTORATI DI RICERCA

Dottorato di Ricerca in Tecnologie e Scienze per la Salute dell'Uomo

Dipartimento di Scienze e Tecnologie Biologiche, Chimiche e Fisiche (STEBICEF)

S.S.D BIO/11

UNRAVELLING PATHOLOGICAL ATAXIN-3 AGGREGATION: NEW INSIGHTS FROM AND FOR *IN VITRO* STUDIES

IL DOTTORE

MARTINA SOLLAZZO

IL COORDINATORE

Prof. BRUNO GIUSEPPE PIGNATARO

IL TUTOR

Dott.ssa CATERINA ALFANO

IL CO TUTOR

Prof.ssa VALERIA VETRI

CICLO XXXVI

ANNO CONSEGUIMENTO TITOLO 2024

The research work here presented was funded by a three-year doctoral scholarship provided by Fondazione Ri.MED,

and the curricular period in an international institution was carried out at the European Synchrotron Radiation Facility (Grenoble, France) and financially supported also by an additional six-month scholarship received from the Erasmus+ Traineeship Program.

Abstract

Neurodegenerative diseases represent a severe global burden for millions of suffering patients and their families, as well as a large sanitary emergency for governments. Despite significant differences in clinical manifestation and prevalence, neurodegenerative diseases share a common pathological mechanism that correlates cellular toxicity and the increase of neuronal collapse with the progressive appearance of abnormal proteinaceous assemblies in brain tissue. Therefore, the blocking of the conversion of normally soluble peptides and proteins into oligomers or intractable amyloid deposits has emerged as a potential therapeutical strategy.

In this context, the focus of this PhD project centres on investigating the molecular mechanisms of the aggregation of Ataxin-3, a ubiquitin-hydrolase involved in the protein quality control system of the cell. Mutated Ataxin-3 typically accumulates in nuclear aggregates, mediating cytotoxicity and leading to Spinocerebellar Ataxia type 3, the most common among inherited ataxias, belonging to the group of polyglutamine neurodegenerative diseases. Although a large number of studies have approached to the pathogenesis of this disease, the details of the supramolecular assembly process of Ataxin-3 remains controversial, and a targeted and effective therapy for Spinocerebellar Ataxia type 3 is currently lacking.

Here, a multiplexed experimental approach based on molecular biology methods, nuclear magnetic resonance, fluorescent spectroscopy, and advanced microscopy techniques was used to investigate Ataxin-3 behaviour in different experimental conditions, focusing in particular on the effect of temperature and buffer composition on different steps of Ataxin-3 self-assembly. The investigation had the double aim of revealing new insights on Ataxin-3 aggregation, and setup a robust and reproducible aggregation assays for screening anti-aggregating compounds. The aim was successfully reached within the three years of the PhD fellowship.

Moreover, this project also aimed at characterizing Ataxin-3 in complex with polyubiquitin chains, one of its recognized physiological partners, and investigating the effect of these chains on Ataxin-3 aggregation. Indeed, evidence suggests an existing competition among proteins' normal function and their tendency to aggregate, thus encouraging the exploration of interactions between neurodegeneration-related proteins and their natural binders as a new promising therapeutic perspective. Accordingly, gold-standard Structural Biology techniques, such as SAXS and X-ray crystallography, were employed to try to solve the structure of Ataxin-3 in complex with tri-ubiquitin chains, to gain structural information on the binding surface involved into this crucial interaction. SAXS experiments and preliminary models obtained by rigid body approach opened the way to further investigation by X-ray crystallography, to obtain atomic details. A first crystal screening was

performed through access to the EMBL Grenoble's High Throughput Crystallization and Fragment Screening Facility. Crystal hints have been obtained and will be used for the next steps.

In parallel, the effect of polyubiquitin chains on Ataxin-3 aggregation was investigated by fluorescence spectroscopy and atomic force microscopy. Interesting clues have been obtained, but further optimization of the experimental setup is required to complete this part of the project.

Table of contents

1 Introduction

Neurodegenerative diseases have become a real burden for millions of suffering patients and their families, as well as a large sanitary emergency for governments.

Notwithstanding large differences in clinical manifestation and prevalence, neurodegenerative diseases share a common pathological mechanism that correlates cellular toxicity and the increase of neuronal collapse with the progressive appearance of abnormal proteinaceous assemblies in brain tissue.[1] As a consequence, the blocking of the conversion of normally soluble peptides and proteins into oligomers or intractable amyloid deposits has emerged as a potential therapeutical strategy.[2]

In this context, the focus of this PhD project centres on investigating the molecular mechanisms of the aggregation of Ataxin-3, a ubiquitin-hydrolase involved in the protein quality control system of the cell. Mutated Ataxin-3 typically accumulates in nuclear aggregates, mediating cytotoxicity and leading to Spinocerebellar Ataxia type 3 (SCA3), the most common among inherited ataxias, belonging to the group of polyglutamine neurodegenerative diseases.[3,4] Although a large number of studies have approached the pathogenesis of this disease, the details of the supramolecular assembly process of Ataxin-3 remain controversial, and a targeted and effective therapy for SCA3 is currently lacking.

Here, a multiplexed experimental approach based on molecular biology methods, nuclear magnetic resonance, fluorescence spectroscopy, and atomic force microscopy, was used to observe Ataxin-3 behaviour in different experimental conditions, gaining new insights into the molecular mechanisms leading to pathological Ataxin-3 aggregation.

Moreover, this project also aimed at characterizing Ataxin-3 in complex with polyubiquitin chains, one of its recognized physiological partners,[5–7] and investigating the effect of these chains on Ataxin-3 aggregation. Indeed, evidence suggests an existing competition among proteins' normal function and their tendency to aggregate, thus suggesting that exploring interactions between neurodegeneration-related proteins and their natural binders might offer promising therapeutic perspectives.[8–10] Accordingly, gold-standard Structural Biology techniques, such as SAXS and X-ray crystallography, were employed to try to solve the structure of Ataxin-3 in complex with tri-ubiquitin chains, such to gain insights into crucial interactions that could protect against aberrant aggregation.

The project was funded by a doctoral scholarship provided by Ri.MED Foundation, and was conducted at the Ri.MED Structural Biology and Biophysics Unit (based in Palermo, Italy) under the supervision of Dr Caterina Alfano (Head of the Unit). A period of six months was spent at the European Synchrotron Radiation Facility in Grenoble (France), for which extra financial support was received from the Erasmus+ Traineeship program.

1.1 Neurodegenerative diseases

1.1.1 Overview on Neurodegenerative diseases

Neurodegenerative diseases (NDDs), which include but are not limited to Alzheimer's disease (AD), Parkinson's disease (PD), Huntington's disease (HD), and amyotrophic lateral sclerosis (ALS), constitute a heterogeneous group of debilitating and largely intractable conditions that progressively compromise the structure and function of the nervous system. This class of disorders shares a common thread of relentless neuronal degeneration, manifesting in a spectrum of severe clinical symptoms, encompassing cognitive decline, motor dysfunction, and, in some cases, a combination of both.

The pathogenesis of NDDs involves the accumulation of abnormal protein aggregates in neuronal cells, the selective loss of neurons, and the consequent disruption of neural circuitry. Neurodegeneration is usually sporadic and progressive but the genetic background, familial mutations, and environmental factors may contribute to early onset of the disease.[11–15] This generally begins with dysfunction in a discrete area, that progressively involves larger regions, leading to distinct pathological features. The molecular processes behind the pathophysiology of NDDs are markedly intricate, and there exists a substantial amount of evidence indicating that these disorders result from multifactorial conditions, including abnormal protein dynamics involving faulty protein degradation and aggregation, oxidative stress, and consequent formation of free radicals, compromised bioenergetics and mitochondrial dysfunction, and exposure to metal toxicity and pesticides.[16] The prevalence of neurodegenerative disorders is intricately tied to the ageing process, with age standing out as the primary risk factor for their onset. As global demographics shift towards an ageing population, thanks to the progress of biomedicine that promotes a lengthening in life expectancy, the impact of these disorders on public health and societal well-being becomes increasingly pronounced.

NDDs can be broadly defined by the interested anatomical areas and clinical presentations, such as resting tremors, lack of movement control and rigidity (i.e. neuronal loss localized in the motor pathway), or progressive loss of memory, of task performance, and speech (i.e. when the cognitive pathway is affected).[2,17–20] The severity of these symptoms, and the age-related rising prevalence, place a growing burden on healthcare systems, caregivers, and the millions of suffering patients, highlighting the urgency of unravelling the complex mechanisms that underlie neurodegeneration. Despite the extensive efforts, many aspects of neurodegenerative pathologies remain elusive. This lack of clarity hampers the development of targeted therapeutic interventions, emphasizing the need for further exploration.

1.1.2 Neurodegenerative diseases as proteinopathies

Over the last decades, the field of neurodegeneration research collected substantial evidence allowing for the classification of well-known NDDs - such as Alzheimer's disease (AD), Parkinson's disease (PD), Dementia with Lewy Bodies (DLB), Fronto-Temporal Dementia (FTD), Amyotrophic Lateral Sclerosis (ALS), Huntington's disease (HD) and Polyglutamine (polyQ) disorders - as cerebral proteinopathies, referring to the aberrant accumulation and aggregation of specific proteins within the neurons, leading to cellular dysfunction and eventually neuronal death.[18,21–32]

Indeed, despite the distinct clinical symptoms exhibited by each NDD, postmortem neuropathological evaluation has unveiled first-hand observation of protein pathology, correlating the presence in patients' brains of histopathologically appreciable proteinaceous assemblies with neuronal collapse for all the aforementioned disorders.[33–44] Moreover, experiments involving the intracerebral injection of human brain homogenates from patients in healthy animals have demonstrated the formation of appreciable amyloid lesions also in the new host, suggesting that the introduction of misfolded proteins is able to seed the aggregation process. Comparable results have been achieved through the overexpression of suspected causative proteins in various animal models or by injecting recombinant protein aggregates into them.[45–51] Alzheimer's disease and other cerebral proteinopathies seem then to arise from the misfolding and corruption of specific endogenous proteins.

Among the extensively studied proteins implicated in the pathological processes of NDDs, there are amyloid-beta (aβ), tau, α-synuclein (α-syn), TAR-DNA binding protein (TDP-43), polyQ-expanded proteins (huntingtin and ataxins), and prion protein (Table 1.1).

These proteins constitute the primary components of various pathological structures such as amyloid plaques and neurofibrillary tangles (AD), Lewy bodies (i.e. PD, FTD, and DLB), neuronal cytoplasmic and intranuclear inclusion (ALS), nuclear inclusions (polyQ diseases), and spongiform plaques (prionopathies as transmissible spongiform encephalopathy, kuru, etc) (Fig. 1.1).[2,52,53] These understandings have opened important avenues for targeted therapeutic interventions aimed at mitigating the aggregation of the specific proteins associated with these devastating neurological conditions. Nevertheless, despite notable progress in the field, several aspects remain elusive, posing an ongoing challenge to the scientific community in the quest for an effective treatment. This challenge is underscored by the intricate nature of the molecular processes leading to protein aggregation, as thoroughly explored in the subsequent section 1.1.3.

Neurodegenerative diseases	Aggregated proteins	Lesions and accumulation sites
Alzheimer's disease (AD)[21,22]	Amyloid-β peptide	Extracellular amyloid plaques
	Tau	Cytoplasmic NFT
	α-synuclein	Cytoplasmic Lewy bodies
Parkinson's disease (PD)[23]	α-synuclein	Cytoplasmic Lewy bodies
Huntington's disease (HD)[25]	Huntingtin (polyglutamine expansion)	Nuclear inclusions
Amyotrophic lateral sclerosis (ALS)[24]	SOD1, Fus, TDP43	Cytoplasmic and nuclear inclusions
Dementia with Lewy bodies (DLB)[54]	α-synuclein	Cytoplasmic Lewy bodies
Spino-cerebellar ataxias (SCAs)[27]	Ataxins (polyglutamine expansion)	Nuclear and perinuclear inclusions
Prion diseases[26]	Protease-resistant prion proteins	Extracellular prion plaques

Table 1.1 NDDs and related causative proteins. Adapted from Nigam *et al.* 2015.[55]

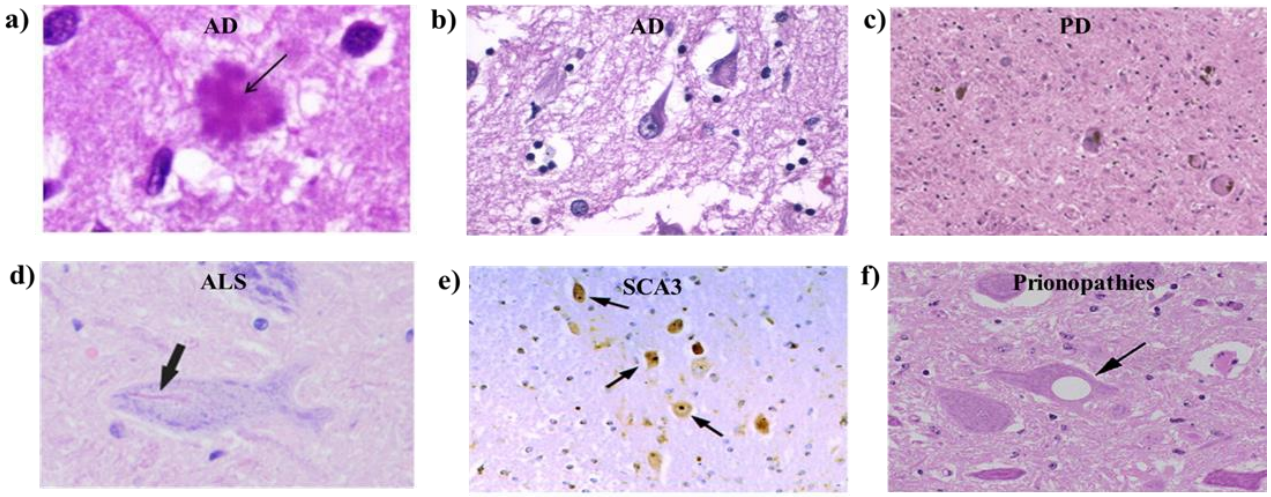


Figure 1.1 Type of protein deposits in NDDs. a) Senile plaques of Aβ with amyloid core in AD brain;[56] b) Phosphorylated tau in neurofibrillary-tangles in AD hippocampal neurons; c) α-syn accumulates in Lewy body in substantia nigra of PD patients;[57] d) TDP-43 nuclear inclusions in ALS;[58] e) nuclear inclusions in ventral pons in SCA3 brain;[59] f) spongiform degeneration and vacuolation in Creutzfeldt-Jakob disease brain.[60]

1.1.3 Failures and mechanisms leading to protein aggregation

The reaching and the maintenance of the proper folding by a protein is essential for the correct exploitation of its physiological functions, that are strictly connected to its unique three-dimensional conformation. To ensure adequate protein folding, sequestration, and degradation, the cell has developed an elaborate protein quality control system, with a highly regulated cascade of events. In healthy conditions, the biologically functional folding is thermodynamically favoured, because of the long-range molecular interactions established between the amino acids (aa). Indeed, these interactions bring to the formation of the conformation with the lowest energy content (i.e. the most stable one), the so-called native form.[61,62] Thus, many proteins ball up into well-packed folded states, where hydrophobic aa are predominantly located in the protein's core, while polar aa are more commonly exposed on the surface, reducing the entropy of the system.

The complexity of the hierarchical folding process is pictured as a funnelled energy landscape,[63] a conceptual model that provides a representation of how a protein navigates through different energy states during the folding process, ultimately reaching its native and functional structure (Fig. 1.2). The folding process is not a rigid step-by-step progression, but rather a dynamic exploration of different structures and energy states: the higher regions of the funnel represent unfolded or partially folded states of the protein. As the amino acid chain folds, it explores various conformations with progressively reduced energy before reaching the native state, which can be reached by multiple pathways and conformations available.[62,63] Errors in this intricate process can lead to the formation of incompletely or incorrectly folded proteins (i.e. misfolded), that expose hydrophobic regions to the solvent, increasing the probability to form cytotoxic self-aggregates.[64] Once assembled, these higher-order aggregates are resistant to disintegration and degradation being the most stable and energetically favourable structures.[65,66]

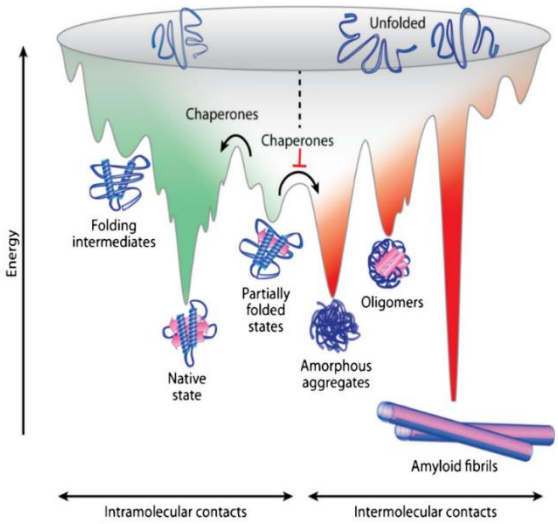


Figure 1.2 Energy landscape of protein folding and misfolding. Protein folding and protein aggregation represent competing processes as a protein converts from an unfolded state to its native state. The folding dynamics are influenced by the energetic landscape, where thermodynamic constraints drive the protein to adopt a conformation that minimizes the free energy of the system. Chaperones play a crucial role in assisting the protein in attaining its functional structure. Intramolecular interactions that are energetically favourable (in green) contribute to increased conformational stability as the protein folds toward its native state. Throughout this folding process, proteins may temporarily assume energetically favourable yet non-native conformations, resulting in the formation of kinetically trapped states. These are susceptible to intermolecular interactions (in red) that can lead to protein aggregation, giving rise to various forms such as amorphous aggregates, β-sheet-rich oligomers, and highly stable amyloid fibrils. Image from Muntau *et al.* 2014. [67]

Although age is surely a critical factor, NDDs-related proteins do not misfold spontaneously, but the increase in their aggregation propensity requires triggers, most of which are likely still unidentified. Both environmental and genetic factors may affect the folding process, thereby rendering proteins more aggregation-prone. If protein folding fails - due to misprocessing phenomena, aberrant interactions, changes in environmental conditions (i.e. pH or temperature), and chemical modification (oxidation, proteolysis)[34] - chaperones may sequester misfolded proteins and target them to the cell degradation system.[68] When this system fails, as in the case of NDDs, misfolded proteins, and their toxic assemblies accumulate into proteinaceous inclusions. This may result in a loss of function, due to the disappearance of functional protein, and/or in a toxic gain of function, associated with the appearance of protease-resistant aggregates, addressing the cell to death.[69,70]

Based on *in vitro* kinetic studies, multiple mechanisms have been put forward to explain protein aggregation. Among them, the most important are nucleated polymerization, nucleated conformational conversion, and native-like aggregation.[71,72] The nucleated polymerization model describes the growth of fibrillar structures through homogeneous (primary) nucleation, and depends on completely or partially disordered monomers. These cluster into nuclei that, once reached a critical size, fast assembly in large supramolecular structures.[73,74] On the other hand, the conformational conversion model assumes that the protein is stable both as folded and misfolded. Conformational changes are then required for the formation of the initial oligomeric species.[53] Lastly, in the native-like aggregation, globular protein, characterized by compact structures with aggregation-prone region normally buried in the core, converts into partially unstructured ensembles that are competent for aggregation. Local unfolding, thermal fluctuation, or ligand release can push globular protein to populate native-like conformation more aggregation-prone, without crossing major energy barrier.[75] Alternative pathways are likely to exist for the stage at which conformational conversion from a non-amyloid to an amyloid conformation occurs.[76]

Multiple aggregation pathways are accessible and selected depending on experimental conditions, protein sequence, and conformational state adopted by the amyloidogenic monomer.[71] The resulting supramolecular aggregates present different structures and morphologies[77] ranging from amorphous structures, mostly made up of disordered polypeptide chains,[78] to highly organized amyloids.[72] Both species occur after protein homeostasis failure, and evidence shows that the pathways are interconnected influencing each other, even if their relationship needs to be clarified.[79]

Despite NDDs-driving proteins present obvious differences in size, aa composition, sequence, or structure, they all convert into final aggregates that share the same internal structure of amyloid aggregates, characterized by a highly organized architecture rich in cross β-sheet structure (Fig. 1.3).

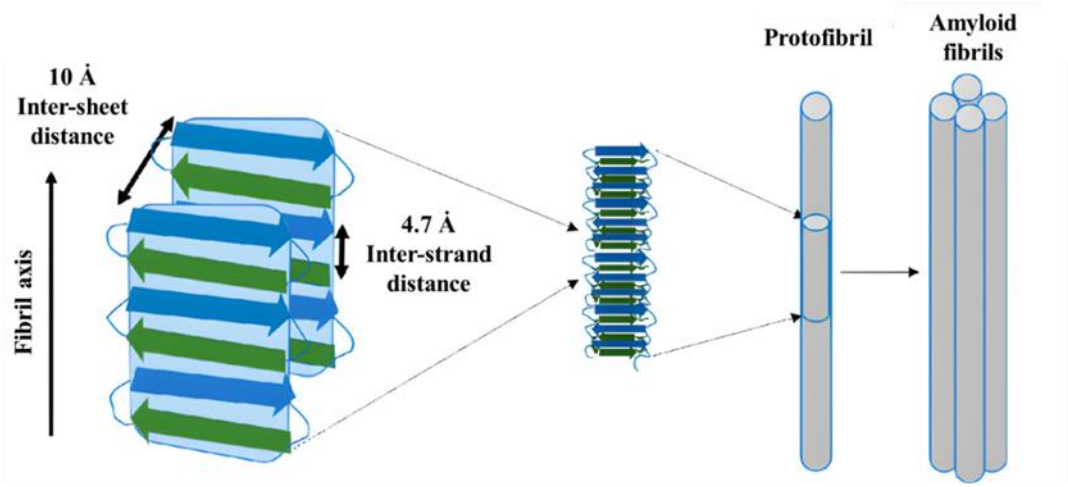


Figure 1.3 Cross β-sheet structure. The core of cross β-sheets consists of highly organized hydrogen-bonds arrangement, in which β-strands run perpendicular to the protofilament axis, resulting in a series of β-sheets that propagate along the direction of the fibril, allowing aggregates maturation in larger structures with unique kinetic stability, defined as amyloids. Image from Vahdat *et al.*, 2019.[80]

The kinetics of *in vitro* nucleated self-assembly processes, including amyloid formation, often follows a sigmoidal trend that can be attributed to the nucleation and growth stages of the assembly process (Fig. 1.4).[12,81] The observable quantities associated with aggregation are negligible during the initial phase of the aggregation process, known as the "lag phase". The signal then undergoes an abrupt growth phase that reaches a plateau called "saturation phase".[82–84]

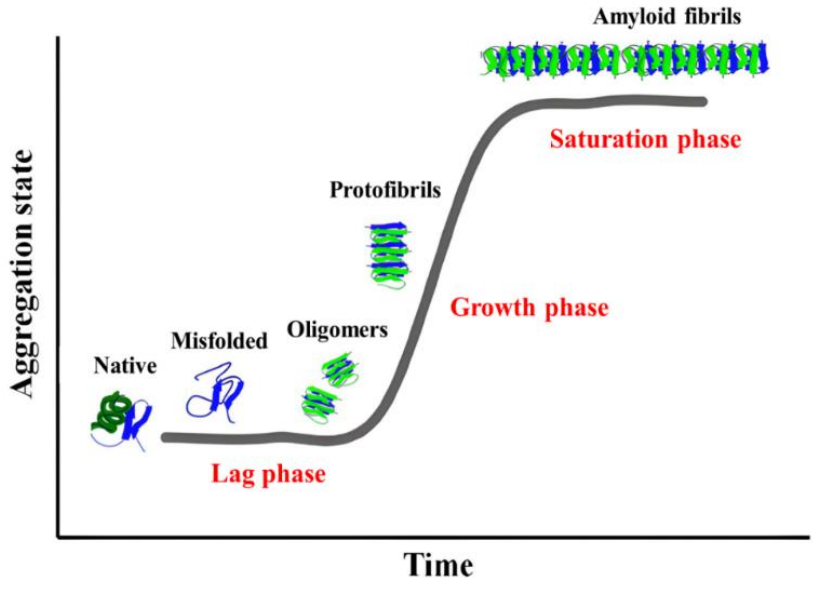


Figure 1.4 Kinetics profile of nucleation-dependent fibril formation process. The sequence of events along the amyloid fibril formation pathway includes lag phase, growth phase, and saturation phase. The lag phase represents the rate-limiting step of the process. During this phase, pathological monomers assemble to form small oligomeric aggregates, soluble and diffusible species that represent the nuclei of aggregation. The length of the lag phase depends on several factors, including chemical-physical properties of the system environment, and protein features such as amino acid sequence, hydrophobicity, and the presence of unfolded regions.[12,85] The lag phase is followed by the growth phase, during which β-structured oligomers associate into protofibrils and finally into stable amyloid fibrils. All the phases consist of equilibrium of multiple species. Image from Iannuzzi *et al.* 2015.[86]

Although evidence clearly shows that NDD-driving proteins convert in amyloid aggregates, solid literature reports as more toxic highly diffusible species, including soluble monomeric, fibrillar, and oligomeric protein species.[87,88] These intermediate species are reported to be the primary toxic agents underlying most proteinopathies, whereas amyloid fibrils may be considered inert or even cytoprotective, inhibiting certain types of toxicity, including generation of reactive oxygen species (ROS) and alteration of cell membrane permeability.[89–96] The prevalence of amyloid aggregates as toxic species in early literature could then be explained by their abundance in human postmortem histopathology.[97]

At any rate, the debate on the identification of the right core elements of pathological progression in NDDs is still open. Any scientific progression on the topic is thus crucial for future therapeutics development.

1.1.4 Genetics behind NDDs

While age stands out as the primary risk factor contributing to the onset of NDDs, years of studies indicate that the interplay between an individual's genetic composition and environmental elements, up to an emerging role for epigenetics components, equally play a significant role in elevating the risk of NDDs.[98]

Aetiology of NDDs includes mutations affecting both the genes encoding for amyloidogenic proteins and others involved in upstream and downstream-related pathways. For instance, approximately 10% of early-onset AD patients (diagnosed before the age of 65 years) present a genetic aetiology of up to 100%,[99–101] while 15% of PD patients are affected by monogenic disease form.[102,103] Moreover, advancement in genome sequencing and genome-wide association studies have enabled the identification of 90 independent risk signals associated with sporadic PD,[104] and 11 new susceptibility loci associated with AD,[105] confirming a substantial role for genetic risk in the arousal of these pathologies.[103,106]

The discovery in 1997 of mutations in the SNCA gene, responsible for encoding α-syn, marked the onset of understanding autosomal-dominant PD.[107] Subsequent identification of various SNCA mutations revealed their association with aggressive early-onset cases.[108]

In the case of AD, all three known familial forms exhibit autosomal-dominant inheritance, with mutations affecting proteins involved in Aβ generation, the primary component of amyloid plaques. Pathogenic mutations in the Amyloid Precursor Protein (APP, Aβ precursor protein), were first identified in 1991.[109–111] The spectrum of these mutations underscores various mechanisms leading to increased levels of Aβ, causing AD or related phenotypes. With the discovery of APP mutations, it was predicted that other genes involved in Aβ processing would be responsible for subsets of

familial cases. Indeed, this proved to be the case with the identification of PSEN1[112] and PSEN2,[113,114] both encoding components of the γ-secretase complex, normally involved in APP cleavage.

Apart from monogenic mutations, which are relatively easy to identify owing to a clear heritage pattern, polymorphisms among multiple genes are also recognised as significant contributors to NDDs aetiology. For instance, ApoE e4, an allele of Apolipoprotein E, normally associated with Aβ and tau clearance, present in 5–35% of the population, elevates the risk of AD threefold if heterozygous and twelvefold if homozygous. Moreover, the ApoE e4 allele exacerbates PD dementia and Lewy pathology.[2,54]

Also, ALS and FTD present genetic risk, with respectively 10% and 40% of cases with familial history.[115,116] It is now recognized that ALS and FTD are two diseases that form a broad neurodegenerative continuum, as suggested by the clinical observation that both disorders can be present within the same family or coexist in the same patient.[117] Over the past decade, cross-sectional studies have indicated that cognitive impairment linked to FTD may develop in as many as 50% of ALS patients. Likewise, motor dysfunction can manifest in up to 30% of FTD patients.[118] Thus, it is not surprising that the two conditions share also some genetic critical players, such as TADBP (i.e. gene encoding for protein TDP-43), FUS (i.e. encoding for Fused-In-Sarcoma protein), SOD1 (i.e. superoxide dismutase 1), and C9orf72. All these genes are involved in RNA metabolism, autophagy, and vesicles and inclusions formation, all typical hallmarks of these disorders.[117] In particular, mutations in SOD1 and FUS are associated with predominantly lower motor neuron symptoms and do not exhibit TDP-43 inclusions.[119] In contrast, carriers of the pathogenic mutation in C9orf72, a hexanucleotide repeat expansion in a gene located on chromosome 9p21, show increased cognitive and behavioural impairment, and characteristic distribution of TDP-43 pathology.[120] This repeat expansion is the most common genetic cause of ALS, FTD, and ALS/FTD responsible for 11% of all ALS and 13% of all FTD cases.[117]

Among NDDs, Huntington's disease (HD) - and all the other hereditary spinocerebellar ataxias belonging to polyQ diseases - exemplify a peculiar case, as they occur exclusively in the presence of a genetic factor (i.e., expansion of CAG repeats, encoding for the glutamine tract), that triggers the molecular mechanism of pathological protein aggregation. Indeed, pathogenic expansion is detected in 99% of cases with typical HD symptoms in large cohorts.[121] This feature makes polyQ diseases a good model for studying pathological protein aggregation, given the opportunity for controlled and easier *in vitro* manipulation of the primary triggering factor: the protein with an expanded polyQ tail. Based on this, Ataxin-3, the protein responsible for Spinocerebellar Ataxia type 3, was selected as a model in this PhD project. Indeed, Ataxin-3 provides an optimal system which recapitulates most of

the features observed in misfolding diseases. Moreover, it has interesting aggregation properties since it aggregates spontaneously when in isolation without the need of destabilizing agents. Ataxin-3 is also the smallest protein among those related to polyglutamine diseases, making it more amenable to straightforward manipulation through recombinant techniques and *in vitro* aggregation assays.

1.2 Spinocerebellar Ataxia type 3

1.2.1 Polyglutamine diseases

Spinocerebellar Ataxia type 3 belongs to the polyglutamine (polyQ) diseases family, a subset of nine inherited NDDs among proteinopathies (Table 1.2). PolyQ diseases are all rare genetic disorders, with average frequency of 1-10 cases per 100.000 people,[122] and comprises Huntington's disease (HD), dentato-rubral-pallido-luysian atrophy (DRPLA), spinal and bulbar muscular atrophy (SBMA) and six distinct spinocerebellar ataxias, including SCA1/2/3/6/7 and 17.[123,124] They have all dominant autosomal transmission (except SBMA, which is X-linked and sex-limited), affect mainly the central nervous system, and unfortunately are all fatal diseases with still no treatment approved to alter the disease course.[125]

Disease	Protein	Pathogenic Q length	Protein architecture
HD[126,127]	Huntingtin	36-121	
SCA1[128,129]	Ataxin-1	40-83	
SCA3[130,131]	Ataxin-3	45-84	
SCA2[132]	Ataxin-2	35-59	
SCA6[133]	A2	21-29	
SCA7[134,135]	Ataxin-7	37-306	
SCA17[136,137]	TBP	47-63	
DRPLA[138]	Atrophin 1	49-88	
SBMA[139]	Androgen receptor	40-62	

Table 1.2 PolyQ diseases and related proteins. The polyQ tract is indicated in black. Adapted from Saunders *et al.* 2009.[140]

As all NDDs, polyQ diseases are accompanied by progressive death of neuronal cells in distinct regions of the brain, with consequent neurological symptoms such as cognitive impairment and physical/motility disorder.[141] The pathological features progressively worsen until the death of the patient which usually occurs in 15 to 20 years from the disease's onset.[142]

In polyQ disorders, protein aggregates (or inclusion bodies) accumulate within the nuclei or cytoplasm of affected neuronal cells. The dysfunctional accumulation is related to the abnormal expansion of a polyQ stretch within the causative protein sequence, due to tandem CAG nucleotide repeats in the exonic part of the coding gene.[3,143,144] The length of the expansion correlates with the age of onset of the disease and with the severity of the symptoms.[142,145–147]

Apart from the expansion of the polyQ tract, polyQ-containing proteins share no similarities in size, biological functions, and cellular localization, supporting that the expanded polyQ tract itself plays a crucial role in pathological aggregation.[3,123] Although the details of the molecular mechanism are not yet well understood, there are several hypotheses to explain this correlation. The underlying mechanism is thought to reflect the toxic dominant properties of the causative proteins that, when harbouring the expanded polyQ, interact differently with their natural binders, becoming more prone to aggregate.[4] This may overload the protein quality control system, creating proteotoxic stress and mitochondrial dysfunction.[148] Other toxic effects may be directly caused by a loss of function mediated by the sequestration of functional protein in supramolecular assemblies or by the production of toxic polyQ-containing fragments.[123]

Other studies suggest an even more complex scenario where also polyQ-flanking domains play a crucial role, mediating the formation of intermediate soluble aggregates. According to this, the hypothesized aggregation mechanism consists of two steps: the first involves polyQ-flanking domains, that produce SDS-soluble and small intermediates, and a further step, directly mediated by the expanded polyQ tract, that leads to the formation of SDS-insoluble and stable inclusions.[7,140,149–153] Additional complexity is given by *in vitro* and *in vivo* studies, including research on *Drosophila melanogaster* and transgenic mice, which suggest a role in polyQ diseases also for caspases and calpains proteases. Proteolytic cleavage of polyQ proteins by these proteases might indeed contribute to the production of toxic polyQ-containing fragments, which are highly reactive and readily undergo to misfolding and aggregation.[154–159]

In spite of this complex scenario, it is crucial to find therapeutic strategies to interfere with the pathogenic cascade of polyQ diseases, which although rare are devasting and fatal for affected individuals.

1.2.2 The case of Spinocerebellar Ataxia type 3: clinical symptoms and pathophysiology

Spinocerebellar Ataxia type 3 (SCA3) is the most common among genetic ataxias, with a global estimated prevalence of 1-5 cases per 100,000 people, and significant geographical and ethnic variations. A higher number of cases has been reported in Brazil, Portugal, and Japan.[160–163] The first description of SCA3 occurred in 1972 among USA immigrants of Azorean origin,[164] and it was

initially named Machado-Joseph disease from the name of the two families where it was first identified.

SCA3 is caused by the expansion of CAG triplets in the exon 10 of the ATXN3 gene, that codes for the protein Ataxin-3 (Atx3).[130,165] In healthy individuals, CAG repeats range from 12 to 44, whereas in affected patients the CAG repeats range from around 56 to 87.[147,166,167] Individuals harbouring CAG repeat lengths ranging from 45 to 55 present instead of incomplete penetrance of SCA3 symptoms.[168]

The most affected brain regions in SCA3 patients are the cerebellar dentate neurons, basal ganglia, brainstem, and spinal cord (Fig. 1.5a).[169–172] Symptoms onset typically begin in adulthood between the third to the fifth decade of life and progresses slowly with age.[173] Patients suffer from impaired balance and coordination (i.e. ataxia), gait abnormalities, and dysarthria (Fig. 1.5b). They may also experience muscle stiffness, weakness, and tremors, further complicating motor functions. Moreover, SCA3 is often associated with ophthalmological features, such as oculomotor disturbances and double vision. In advanced stages, cognitive impairment, depression, and psychiatric symptoms also occur, affecting the patient's overall quality of life.[174]

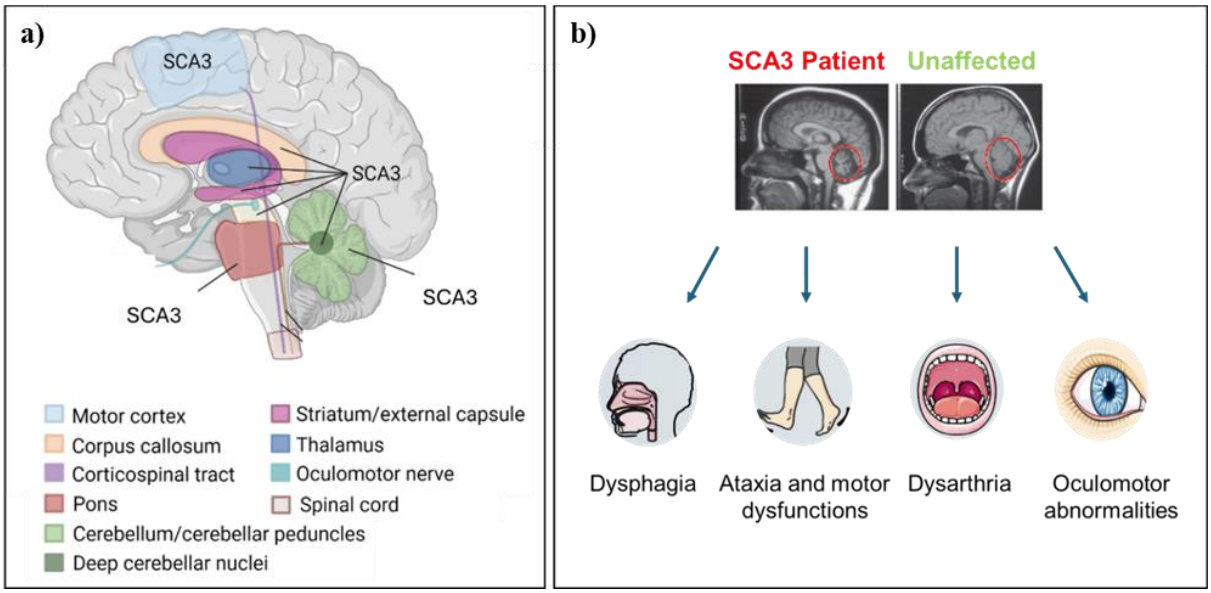


Figure 1.5 Brain areas interested by SCA3 and clinical symptoms. a) Central nervous system areas interested in SCA3 alterations. b) Among SCA3 wide clinical spectrum, progressive cerebellar ataxia is normally present. Other symptoms include pyramidal syndrome, peripheral neuropathy, oculomotor abnormalities, extrapyramidal signs, motor-speech impairments (dysarthria), swallowing difficulties (dysphagia), and sleep disorders. Figure edited with Biorender. Adapted from a) Putka *et al.* 2023[175,176] and b) Wan *et al.* 2020.[175]

As with the other polyQ diseases, in SCA3 the CAG repeat length inversely correlates with age of disease onset and directly correlates with the severity of the symptoms.[147,163,166]

Genetic tests are the most definitive method for diagnosing SCA3, accompanied by neurological examination and imaging studies. The histopathological hallmark of SCA3 is represented by the accumulation of ubiquitinated protein aggregates containing mutant polyQ-expanded Atx3, that are

found throughout vulnerable brain regions in SCA3.[59,177,178] Intracellular inclusions contain both normal and mutant Atx3, as well as other proteins, including ubiquitin, heat-shock proteins, and transcription factors.[125,179–181] Nuclear inclusions make up the majority of aggregates, though smaller cytoplasmic inclusions and distal axonal aggregates also occur.[181–183]

To date, there is no effective treatment for SCA3 and the average life expectancy of SCA3 patients is set around 20 years after onset.

1.2.3 Ataxin-3 architecture

SCA3 causative protein Atx3 is encoded by the ATXN3 gene, located on the long arm of human chromosome 14 (14q32.1). The ~48 Kb long ATXN3 gene possesses 11 exons, with the CAG repeat located in the 10th exon,[184,185] and is subjected to alternative splicing.[186] The protein Atx3 produced by the transcripts is about 42 kDa, depending on the polyQ expansion, and is ubiquitously expressed throughout the body.[187]

Atx3 is a modular protein consisting of an N-terminal globular domain, called Josephin (Jos), stretching from aa 1 to 182, followed by two ubiquitin-interacting motifs (UIMs) and the expandable polyQ tract. A third UIM is present downstream of the polyQ tract in the brain-predominant isoform (Fig. 1.6).[177]

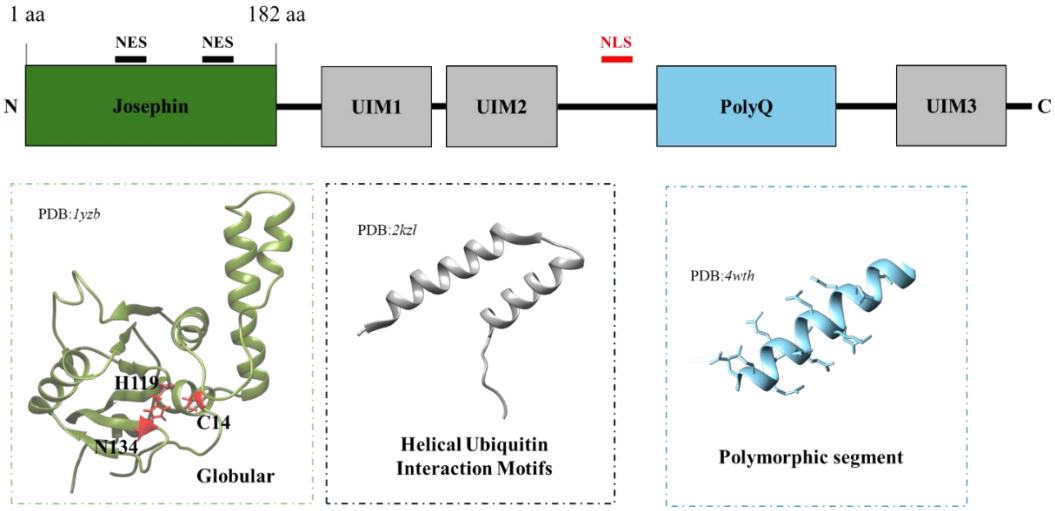


Figure 1.6 Atx3 domain organization. Schematic representation of the modular structure of Atx3 composed of globular Josephin domain (Jos, green) and a C-terminal flexible tail containing two or three ubiquitin interaction motifs (UIMs, grey), depending on the isoform, and the polymorphic polyQ stretch (blue). Jos is a cysteine protease with a catalytic triad consisting of C14, H119 and N134. Two NES sequences (77-99 aa, 141-158 aa, in black) and a NLS (282-285 aa, in red) regulate Atx3 nuclear-cytoplasmic shuttling. Adapted from Carvalho *et al* 2018.[188]

The polyQ region has been observed to exhibit conformational flexibility, adopting a variety of structural states, ranging from random coil to α-helical structures.[189] This polymorphic characteristic of the polyQ region in Atx3 is not unique and is also shared by polyQ segments found in other polyQ-related proteins.[190]

The Josephin catalytic domain belongs to the papain-like cysteine proteases family, with the active site consisting of strictly conserved aa, i.e. Cys14, His119, and Asn134, representing the catalytic triad involved in the substrate cleavage (Fig. 1.6).[6,191,192] The first structure of Josephin was solved in 2005 by NMR and consists of two subdomains, one rich in α-helices and the other rich in β-strands forming an antiparallel β-sheet, separated by a cleft (PBD: 1yzb).[7] Despite being a predominantly rigid structure, Jos also possesses a flexible helical hairpin formed by α-2 and α-3 helices which samples multiple conformations and can be found in a closed[191] (PDB 2aga), half-closed[193] (PDB 2dos), and open[7] (PDB 1yzb) conformation.

NMR structure in solution was also solved for UIM1 and UIM2, which are separated by a short 2 aa spacer, and revealed that they can fold into two α-helices separated by a flexible linker (PDB: 2klz).[194] Upon ubiquitin binding, this structure adopts a typical helix-loop-helix folding pattern, where hydrophobic interactions dominate the complex formation. The presence of the third UIM in the most prevalent brain isoform seems not to affect Atx3 binding affinity for Ub substrates, since no differences in the catalytic activity were observed when 2 UIMs- and 3 UIMs-containing isoforms were compared.[194]

A nuclear localization sequence (NLS) is situated before the polyQ region and most likely allows translocation to the nucleus also of C-terminal fragments generated by cleavage of expanded Atx3, probably contributing to the formation of intranuclear aggregates.[195] Nucleus-cytoplasm shuttling is regulated by the presence of two nuclear exporting signal motifs (NES) in the N-terminal region (aa 77-99 and aa 141-158), that address the protein to the cytoplasm, counterbalancing nuclear translocation mediated by the NLS. This mechanism is finely regulated by interaction with specific regulators.[59,196]

1.2.4 Physiological roles of Ataxin-3

Despite Atx3 involvement in SCA3 being well assessed, its physiological roles are still debated. Atx3 is ubiquitously expressed throughout different organs, tissues, and cell types, and takes part in several cellular pathways, establishing interaction both through its N-terminal domain and C-terminal regions (Fig. 1.7).[197]

One of the many suggested functions sees Atx3 involved as de-ubiquitinase in the proteosome-mediated protein degradation pathway. As cysteine protease, it is able to cleave Ubiquitin-Ubiquitin isopeptide bonds in proteins tagged with Lys48-linked polyubiquitin chains for proteosome degradation, recycling polyubiquitin (polyUb) chains and restoring the cytosolic pool of monomeric Ubiquitin (Ub).[198] Atx3 also recognizes Lys63-linked polyUb, which are involved in protein function regulation, subcellular localization, protein–protein interactions, and DNA repair.[199]

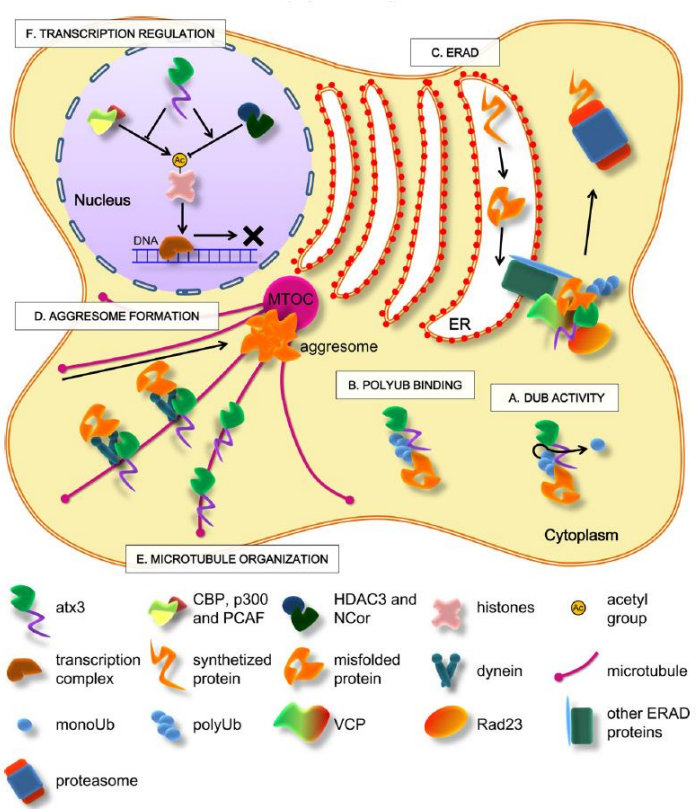


Figure 1.7 Cellular pathways involving Atx3. a) The deubiquitinase (DUB) activity of Atx3, coupled with b) its capability to bind polyUb chains, and c) the documented interactions with components of the ERAD and UPP protein homeostasis systems, suggest a biological role in the proteome quality control mechanism. d) The association with proteins linked to aggregasomes, along with its co-localization with aggregasome and pre-aggregasome particles, as well as e) its interaction with cytoskeletal proteins, provides evidence for Atx3's involvement in aggregasome formation and microtubule organization, respectively. f) Lastly, its interaction with transcription activators, repressors, and histones signifies a role in the regulation of transcription. Image from Matos *et al.,* 2011.[3]

The mechanism through which Atx3 is able to bind polyUb chains is still debated. Jos catalytic triad is flanked by two Ub binding sites, Site 1 (residues I77, L91, W120, F163) and Site 2 (residues Y27, F28, W87), contiguous but placed on the opposite surfaces (Fig. 1.8).[7] Site 1, located in proximity to the catalytic region, plays a pivotal role in polyUb cleavage activity by binding to a Ub moiety and facilitating its precise positioning within the catalytic site.[7] Meanwhile, Site 2 governs the preference for linking ubiquitin chains and engages with the N-terminal domain of the ubiquitin-like domain of HHR23A and HHR23B, the human homologues of yeast DNA repair protein Rad23, other actors of the ubiquitin-proteasome pathway.[200–202] Binding affinities of Jos and monomeric Ub are very weak, supporting Atx3 mild deubiquitinase activity *in vitro*.[203] The two conserved UIMs located upstream of the polyQ region seem also to be involved in Ub binding operating cooperatively, as shown by *in-vitro* NMR and ITC studies.[194] Moreover, the identification in the UIMs of several serine residues as phosphorylation sites, supports the hypothesis of finer regulation by post-translational modifications that could influence Atx3 interactions.[204,205]

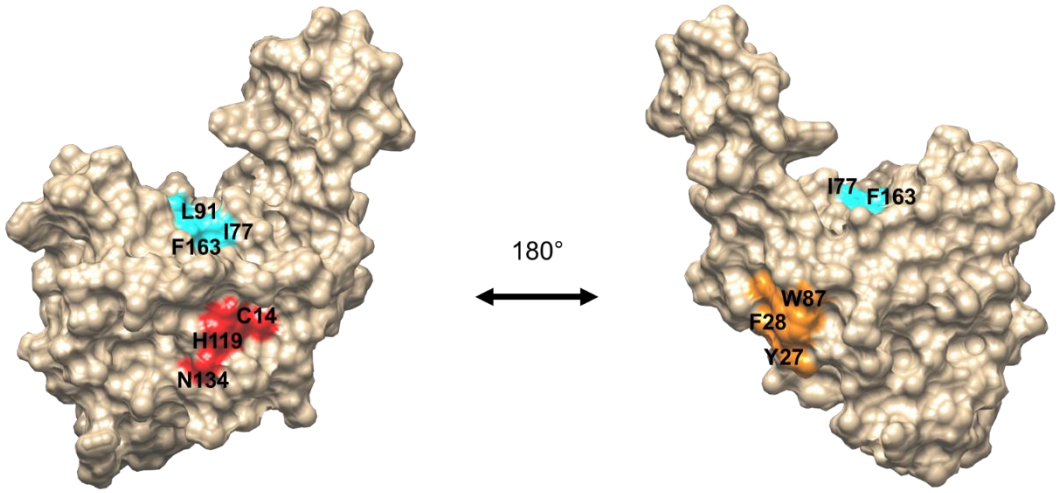


Figure 1.8 Relevant Josephin sites involved in Ub chains binding. Surface representation of Josephin domain (PDB: 1yzb).[7] The active site residues (i.e., C14, H119, and N134) are shown in red, Ub binding site 1 (i.e. I77, L91, W120, F163) is shown in cyan and Ub binding site 2 (i.e. Y27, F28, W87) is shown in orange. Site 1 is also defined as proximal Ub binding site, due to its contiguity to the active site; conversely, site 2 is defined as distal site.

Atx3 also contains a sequence rich in arginine ('RKRR') in its C-terminal region which is known to interact with Valosin-containing protein (VCP, also known as p97), an AAA ATPase involved in multiple cellular pathways, as the extraction of ubiquitinated proteins from endoplasmic reticulum to address them to degradation.[191,206,207] Intriguingly, targeting Atx3-VCP binding by protein decoy in *Drosophila* model of SCA3 ameliorates the symptomatology, reducing the amount of Atx3 aggregates.[208,209]

Atx3 interactome hints at additional biological functions (reviewed in Matos *et al.* 2011).[3] This includes interactions with histones,[7,210] and transcription repressors[211] and activators,[210] suggesting a potential role of Atx3 in transcription regulation. Moreover, interactions with dynein,[212] tubulin, and MAP2,[213,214] and the observed morphological changes and adhesion loss in Atx3-silenced cell lines, indicate involvement of Atx3 in cytoskeletal organization.

In SCA3, the intricate intracellular pathways involving Atx3 are strongly affected being many players sequestrated in the nuclear and perinuclear inclusions formed by Atx3 aggregates. This results in a loss of function, due to the disappearance of functional proteins, and in a toxic gain of function associated with the appearance of protease-resistant aggregates, addressing the cell to death.[69,70]

1.2.5 Ataxin-3 aggregation

Several studies *in vitro* demonstrate that expanded Atx3, responsible for SCA3, forms fibrils enriched in β-structures, able to bind amyloid-specific dyes.[215] However, both non-expanded Atx3 and isolated Jos domain are also able to fibrillate under physiological conditions.[149,216–218] Moreover, *in cell* studies demonstrate that when the Jos domain of Atx3 is substituted with a similarly sized domain,

the formation of large aggregates is inhibited. This body of evidence strongly supports the hypothesis that even though the polyQ tract is *per se* toxic and linked to the pathology, also polyQ flanking regions play a crucial role in the supramolecular assembly process of Atx3,[219] as also observed in other polyQ proteins such as Atx1 and huntingtin.[216,217,220–224] Accordingly, the Atx3 fibrillogenic pathway has been characterized kinetically as nucleation-dependent polymerization consisting of a two-stages aggregation mechanism which involves also the polyQ-flanking Jos domain (Fig. 1.9).[152,153,225,226]

The first aggregation stage is polyQ independent and leads to the formation of short, ThT-positive, and SDS-soluble protofibrils. The process is mediated by aggregation-prone regions located in the Josephin domain, and accordingly, it occurs in both expanded and non-expanded Atx3.[149,216,226,227] In the initial phases, Jos domain retains a native-like secondary structure but is deployed of its catalytic activity pointing to a subtle conformational change before fibril assembly.[149] The self-association of Jos is initiated through dimerization. The more compact dimer further sustains linear oligomer assembly by sequential monomer addition.[227] In the second aggregation stage, these oligomers convert to amyloid-like fibrils due to the expanded polyQ tract which leads to the formation of longer, mature, and detergent-resistant fibrils, stabilized by side-chain mediated hydrogen bonds.[152,228]

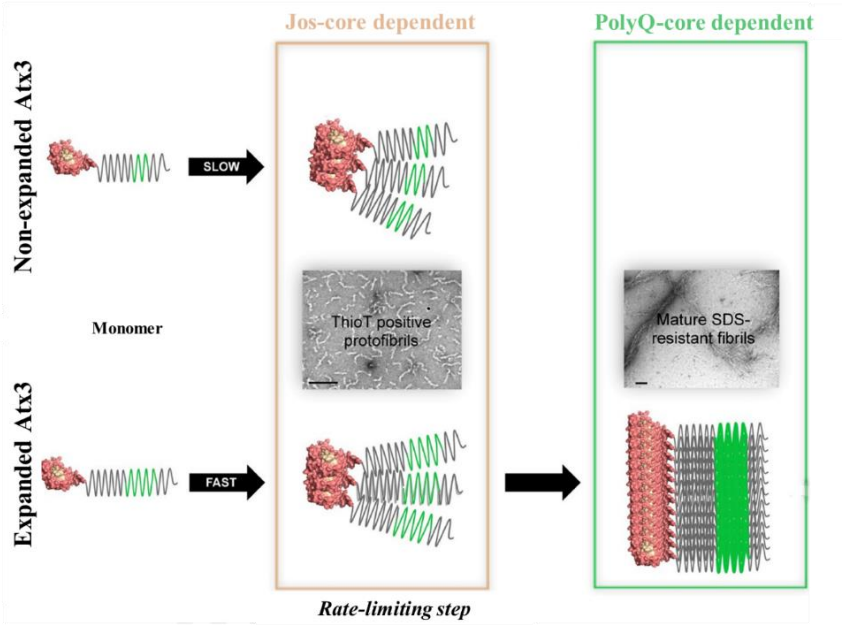


Figure 1.9 Proposed multi-steps aggregation mechanism for Atx3. The first aggregation step is controlled by the Jos domain, forming ThT-positive and detergent-soluble protofibrils. A second aggregation step is controlled by expanded polyQ tract that destabilizes the Jos domain, leading to the formation of ThT-positive and SDS-resistant mature fibrils. Image from Silva *et al.* 2017.[229]

According to *in silico* and experimental evidence, polyQ expansion favours fluctuation in dynamic regions of Jos, influencing, in particular, three identified highly amyloidogenic peptides (76-81 aa, 91-96 aa, 153-160 aa) positioned on the opposite sides of the catalytic cleft. This results in a transient

disruption of the hydrophobic packing in Jos between the helices α1 and α4 and the domain core.[230] Moreover, mutations of single key residues on these peptides (i.e. I77A and S81A, clustered within helix α4, and L93A located on strand β1) were shown to strongly affect the capability of expanded Atx3 to aggregate.[230] Interestingly, these aggregation-prone regions are highly conserved through evolution and contain two distinct surfaces involved in Ub binding (i.e. proximal Ub site 1, I77, L91, W120, F163; and distal Ub site 2, Y27, F28, W87).[203] Accordingly, the addition of an excess of ubiquitin to Jos significantly slows down the aggregation kinetics.[231] Thus, the same regions are important for physiological functions but, if left exposed, promote aggregation, opening new possibilities to the design of potential therapeutics.[8–10]

Atx3 aggregation in SCA3 is also suggested to be caused by toxic polyQ-containing fragments, believed to represent the starting point for aggregation, so mediating the pathogenicity. These truncated species were shown to display an increased toxicity and aggregation propensity when compared to the respective full-length protein.[154,155,157,232–234] Analogously to studies on HD, early results on SCA3 showed that the expression of polyQ-expanded truncated Atx3 in cell models and mice showed a markedly exacerbated phenotype compared to the expression of full-length Atx3.[235] Moreover, Atx3 aggregates in SCA3 post-mortem tissue were clearly detectable with antibodies directed against the polyQ-containing C-terminus but not against the N-terminus of Atx3.[177]

The formation of these toxic fragments could be mediated by proteolysis acted by caspases and calpains, or by partial degradation of full-length protein by autophagic-proteasomal system.[154,236–238] Many aspects of Atx3 aggregation in SCA3 remain then controversial, although significant efforts devoted to comprehending the molecular mechanism of the process. More and incisive advancements are needed for the development of an effective treatment.

1.2.6 Current therapeutic strategies against SCA3

Despite the significant efforts devoted to comprehending and mitigating the processes underlying neuronal dysfunction and degeneration in SCA3 patients, currently, there is still no cure for SCA3, and treatments administered to patients are aimed at limiting the severity of symptoms. For example, parkinsonism, restless legs syndrome, spasticity, sleep disorders, and depression can be treated pharmacologically.[239] Dystonia and spasticity can be managed with local botulinum toxin injections[240] and speech therapy may also be of benefit for managing dysarthria.[241]

Data show that, though arising from a simple, monogenic factor, SCA3 pathophysiological picture is more complicated than expected, with polyQ expanded Atx3 disturbing several cellular systems, such as ubiquitin-proteosome quality control system and autophagy pathways, and also impairing mitochondrial and calcium homeostasis (extensively reviewed in Matos *et al.* 2019).[174]

Current therapeutic strategies against SCA3 are addressed to eliminate the pathological expanded Atx3 from the cells by suppressing its production,[242] increasing its degradation,[243] or by interfering with the putative toxic roles of its proteolytic cleavage or aggregation.[3,151,244,245]

Approaches aimed at disease gene suppression seem promising, being mice lacking ATXN3 phenotypically normal.[198] Delivery of antisense oligonucleotide mRNA or duplex RNA has been tested to inhibit aberrant Atx3 expression and slow down the progression of the disease in mouse models.[246]

Other approaches focus on downstream stages, inducing expanded Atx3 autophagy, attempting to reduce the deleterious effect caused by its accumulation,[247] or inhibiting further post-translational modifications as proteolytic cleavage by caspases and calpains, that produce soluble toxic polyQ-containing fragments.[154,159,248]

Debated results come from several studies in *Drosophila*, where molecular chaperones were used to control Atx3 misfolding. Several chaperones were found to significantly reduce the accumulation of mutant Atx3, resulting in lower toxicity.[249–251] For instance, aB crystallin, a small heat shock protein, is a suppressor of SCA3 toxicity in *Drosophila*,[250] and is able to transiently interact with Jos *in vitro*, altering the first polyQ independent aggregation step.[141] However, later studies show that overexpression of molecular chaperones mitigates toxicity mediated by C-terminal Atx3 fragments rich in polyQ, otherwise increasing aggregation of full-length expanded Atx3.[252] These results reinforce the idea that expanded Atx3 mainly aggregates inside cells, but there are also other toxic species formed somewhere along the self-assembly pathway.

A deeper comprehension of the physiological role of Atx3 is indispensable to elucidate further the complex toxic cascade and the entailed molecular networks that could be affected by aberrant aggregation. New insights on Atx3 interactions at the molecular level, will surely help in elaborating more precise strategies to tackle pathological mechanisms.

1.3 Project aim and strategies

The focus of this PhD project was the investigation of the molecular mechanisms of Ataxin-3 aggregation, to gain new insights into its complex and still controversial supramolecular assembly process.

Ataxin-3 is the protein responsible for Spinocerebellar Ataxia type 3 and also provides an optimal system which recapitulates most of the features observed in misfolding diseases. Moreover, it has interesting aggregation properties since it aggregates spontaneously when in isolation without the need of destabilizing agents. Ataxin-3 is also the smallest protein among those related to

polyglutamine diseases, making it more amenable to straightforward manipulation through recombinant techniques and *in vitro* aggregation studies.

More specifically, within this PhD project it was investigated the effect of temperature and buffer composition on different steps of Atx3 self-assembly, being two variables strongly influencing protein stability, folding, and aggregation propensity. This had the double aim of revealing new insights on Atx3 aggregation, and setup a robust and reproducible aggregation assay for screening anti-aggregating compounds.

Moreover, this PhD project also aimed at characterizing Ataxin-3 in complex with polyubiquitin chains, a recognized Atx3 partner,[5–7] and investigating the effect of these chains on Atx3 aggregation. Indeed, evidence suggests an existing competition among proteins' normal function and their tendency to aggregate, thus suggesting that exploring interactions between neurodegeneration-related proteins and their natural binders might offer promising therapeutic perspectives.[8–10]

2 Materials and Methods

2.1 Ataxin-3 and Josephin constructs

In the present study, three full-length Atx3 variants with distinct polyQ tract length, and the solely Jos domain, were investigated. In particular, the following Atx3 variants were used: Atx3-Q14, Atx3-Q55, and Atx3-Q70, where the number associated with "Q" reflects the number of glutamine residues within the polyQ tract. The three variants were chosen to represent the non-pathological, mildly toxic, and fully toxic Atx-3, respectively.

Expression plasmids for Jos, Atx3-Q14, and Atx3-Q55 were already available in Dr Alfano's lab (Annex III-V), while the expression vector for Atx3-Q70 was produced within this project (Annex VI). In detail, Jos construct consists of fragment 1–182 of human Atx3 cDNA, cloned into a modified pET9d plasmid vector to produce Jos with an N-terminal tag comprising six histidines (6xHis-tag), glutathione S-transferase (GST) fusion protein, and the cleavage site for the recombinant Tobacco Etch Virus (TEV) protease (ENLYLQG). This allows the enzymatic removal of the tag from purified recombinant products since it might interfere with biophysical and structural studies. Atx3-Q14 and -Q55 cDNA were cloned in pMAL expression vectors, to produce recombinant proteins fused with the highly soluble Maltose Binding Protein (MBP). In these two constructs, the 6xHis-tag is positioned downstream of the MBP, and it is followed by the TEV protease cleavage site. Domains organization of constructs is reported in Fig. 2.1.

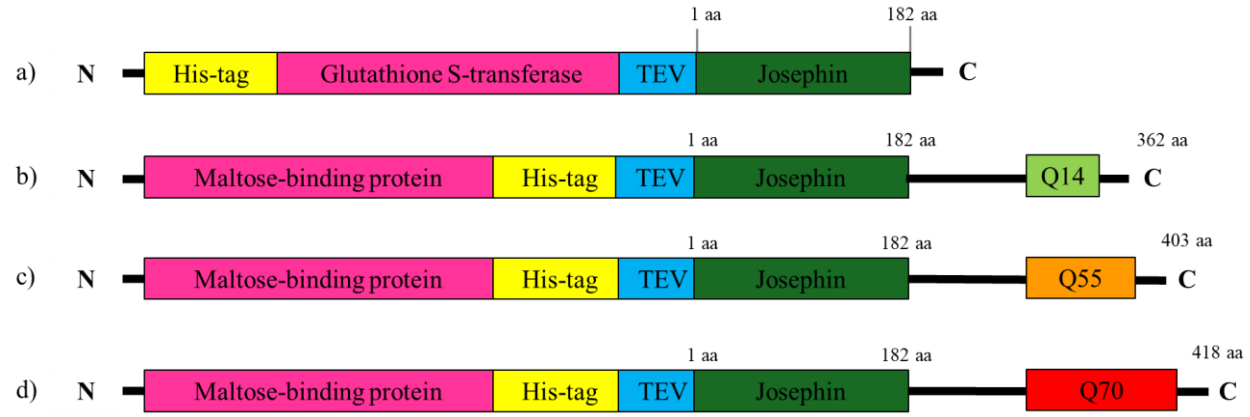


Figure 2.1 Organization of protein constructs used in the present work. N-terminal His-tag, consisting of six histidine residues (yellow), the site recognized by TEV protease (cyan), Jos catalytic domain (dark green), glutathione S-transferase (GST) and maltose-binding protein (MBP), used to increase protein solubility (pink), polyQ expansion (Q14, light green; Q55, orange; Q70, red). a) Josephin domain; b) Non expanded Atx3-Q14; c) expanded Atx3-Q55; d) expanded Atx3-Q70.

2.2 Cloning for Atx3-Q70 variant

To obtain an expression vector for Atx3-Q70 variant, expansion of CAG repeats in Atx3-Q55 construct was performed using Synthesis of Long Iterative Polynucleotides (SLIP) method, as described in Figura *et al.,* 2015,[253] with minor modifications (Fig. 2.2).

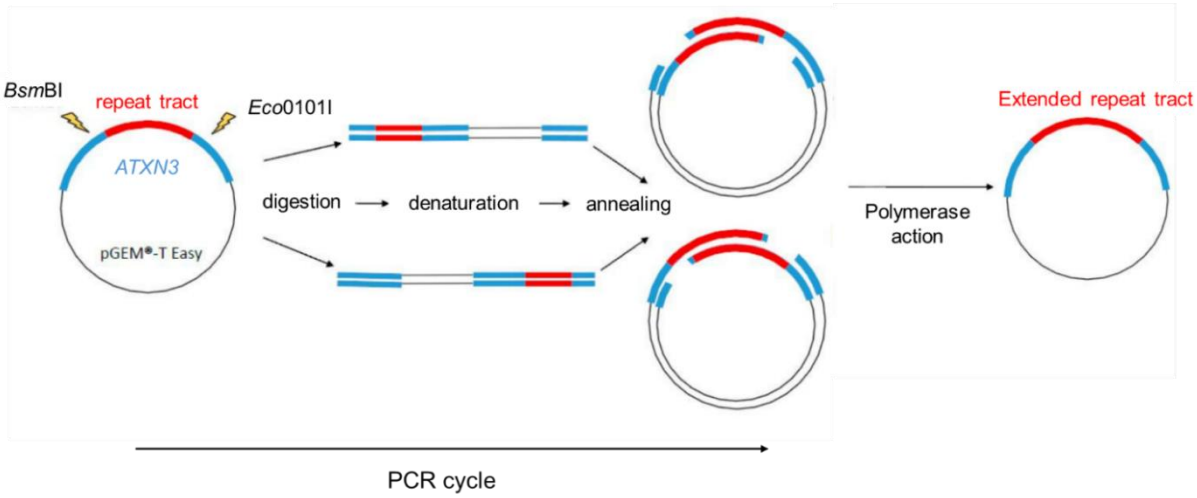


Figure 2.2 Schematic representation of the cloning method used to expand CAG repeats. Firstly, the starting construct was obtained by cloning Atx3-Q55 ORF in pGEM-T Easy plasmid. Then, *Bsm*BI and *Eco*0109I restriction sites, located upstream and downstream to the polyQ tract of Atx3, were used to independently and parallelly digest two separate aliquots of the starting construct. The digested constructs were mixed and used for a single PCR amplification cycle. Polymerase action filled the gap caused by the non-perfect alignment of the constructs. This technique allows both CAG expansion and contraction events. Figure from Figura *et al.,* 2015.[253]

Briefly, the starting construct containing Atx3-Q55 was prepared transferring Atx3-Q55 cDNA in pGEM-T-Easy vector (Promega, Madison, WI, USA) between *Eco*O109I and *Bsm*BI restriction sites according to manufacturers' instructions. The two restriction sites, located upstream and downstream of the polyQ sequence, were required for employing the SLIP method. The amplification of the pGEM-T-Easy vector containing Atx3-Q55 cDNA (pGEM-T-Easy(Atx3-Q55)) was performed in SCS110 *E.coli* dam-/dcm- strain (Agilent) because *Eco*O109I restriction enzyme activity is blocked by dcm methylation. The amplified plasmid was recovered by PureYield™ Plasmid Miniprep System Kit (Promega). Then, two aliquots of pGEM-T-Easy(Atx3-Q55) were independently and parallelly digested upstream and downstream of the repeats tract. In detail, *Eco*O109I and *Bsm*BI digestions were performed separately using 100 ng of DNA and 6 and 3 units of enzyme (NEB) respectively, incubating 12 h at 37°C. Thus, the successful digestion was checked by running the digestion mixture on 1% agarose gel in TAE buffer. The digested products were then recovered from the gel bands by using QIAquick® Gel extraction Kit at room temperature, according to the manufacturer's instructions. Digested and purified pools were mixed together and used in 25 µL PCR master mix, composed as reported in Table 2.1.

Component	Volume	Final Concentration
Q5 reaction buffer	5 µL	1 x
dNTPs	0.63 µL	0.25 mM
Q5 hot Start High fidelity DNA polymerase	1.25 µL	2.5 U
Digested mix	18 µL	100 ng

Table 2.1 Master mix composition for single PCR step to extend CAG repeats.

A single PCR cycle was performed alternating 3 min at 98°C (denaturation step), 10 min at 50 °C (annealing step), and 3 min at 72°C (extension step). Q5 hot Start High fidelity DNA polymerase (NEB) was selected to gap-fill repeats extension. The polymerase first removes mismatched sequences immediately adjacent to the repeat tract and then extends the 3′ end, resulting in repeat tract elongation.[254] PCR product was used to transform DH5-α Competent Cells (NEB) by heat-shock method. Colonies were screened by PCR using the following primers (Table 2.2):

Forward Atx3 colony primer	Reverse Atx3 colony primer
5'- GGAAGAGACGAGAAGCCTAC- 3'	5'-TCACCTAGATCACTCCCAAGT-3'

Table 2.2 Primer sequences for colony screening.

Each screened colony was resuspended in 5 µL PCR MasterMix (1x DreamTaq Buffer, 0.5 µM primers, 0.2 mM dNTPs, 0.5 U DreamTaq Polymerase from ThermoFisher Scientific) and subjected to 30 cycles PCR consisting of alternating 30 sec at 95°C (denaturation), 30 sec at 60°C (annealing) and 1 min at 72°C (extension). The extent of the repeat length change is difficult to control or predict and the method does not exclude CAG contractions. The quality of DNA was assessed by 3% agarose gel in TBE buffer, and the exact number of CAG triplets was confirmed by sequencing (Annex I), using the following primers (Table 2.3):

Reverse Atx3 pGEM primer	Reverse M13 primer
5'-CTACTTCTTCCCTTCTGTTTTCAAATCATTTC-3'	5'-CAGGAAACAGCTATGACC -3'

Table 2.3 Primer sequences for construct sequencing, targeting CAG expanded region.

2.3 Preparation of recombinant Ataxin-3 and Josephin samples

Expression in *Escherichia coli* and protein purification were achieved as previously published with minor adjustments.[255] *E.coli* Rosetta2 (DE3) pLysS (NEB) and BL21 (DE3) (NEB) were used for Atx3 and Jos protein expression, respectively. In particular, *E.Coli* Rosetta 2 (DE3) pLysS strain (NEB) was chosen for Atx3 to enhance the expression of eukaryotic proteins that contain rare codons, and for reducing background expression of the proteins before induction, thus avoiding potential toxicity sources for the cells. Competent cells were transformed by the heat-shock method, according to the manufacturer's instructions. A single colony was resuspended in 5 ml LB supplemented with antibiotics and used to test different expression conditions. Bacteria growth was achieved by incubating the cell culture at 37 °C under shaking. Protein expression was induced when OD_{600}: 0.6-0.8, adding 0.5 mM and 1 mM IPTG for Atx3 and Jos, respectively. Cells were harvested after 4 hours at 37°C, except for Atx3-Q70 construct for which expression was carried out at 18°C, overnight.

Bacterial pellets were resuspended in 20 mM TRIS, 300 mM NaCl, 5% glycerol, 1 mM DTT, 5 mM MgCl2 at pH 7.2, supplemented with 0.5 mg/ml Lysozyme, 10 µg/ml DNase I, and Complete EDTA-free protease inhibitor cocktail. After short sonication by ultrasonic homogenizer, and centrifugation, the soluble fraction was purified by Immobilised Metal ion Affinity Chromatography (IMAC) using HisTrap FF SP columns pre-packed with Nickel-Sepharose resin (Cytiva). Proteins binding to the column was performed in 20 mM Tris-HCl, 300 mM NaCl, 5% Glycerol, 1 mM DTT, pH 7.2 (Buffer A). To wash out unspecific bounded components, before elution, the column was washed with 20 column volume of Buffer A supplemented with 5 mM Imidazole. or 1 M NaCl for Jos and Atx3 variants, respectively. His-tagged proteins were collected after isocratic elution with 20 mM TRIS, 300 mM NaCl, 5% glycerol, 1 mM DTT, and 250 mM imidazole, pH 7.2 (500 mM imidazole for Atx3). The N-terminal tag was removed by overnight digestion with Tobacco Etch Virus (TEV) protease at 4°C in a ratio of 1:10 TEV: Atx3 and 1:20 TEV: Jos. TEV digestion was performed during sample dialysis in 20 mM Sodium Phosphate, 300 mM NaCl, 1 mM DTT, pH 7.2. Digested proteins were then separated from undigested protein, tags, and TEV by reverse IMAC and recovered from the flow through. This was then concentrated, aliquoted, and stored at – 80°C until use (Suppl. Fig. S1-S3).

Atx3-Q70 was purified according to the previous protocol, with some modifications, including the introduction of 10 µg/ml RNase during cells lysis, elimination of DTT from purification buffers, and an additional purification step by anionic exchange chromatography. Moreover, to reduce protein high-propensity to aggregate, all the purification steps were performed at 4°C and keeping 5% glycerol also during o.n. TEV digestion. This was performed under dialysis using a 1:5 protein: TEV ratio. The untagged sample was recovered in the flowthrough of reverse IMAC, and diluted with 20 mM Tris-HCl pH 7.2, and 5% glycerol, such to reduce NaCl concentration down to 50 mM for further purification by Capto Q ionic exchange column (Cytiva). Elution from Capto Q was performed by 60 CV linear gradient up to 100% of buffer B (20 mM Tris-HCl pH 7.2, 1 M NaCl, 5% Glycerol). Fractions around 20/30% buffer B concentration were selected, aliquoted, and stored at –80 °C until use (Suppl. Fig. S4).

Shortly before the biophysical investigation, all protein samples were subjected to a final purification step by size exclusion chromatography (SEC) to isolate monomeric protein from aggregates. SEC was performed by Hi Load 16/600 75s or Increase 10/300 75s pre-equilibrated in aggregation buffer (i.e. 20 mM Sodium Phosphate, pH 6.5 or 20 mM Hepes pH 7.5, 150 mM NaCl). Protein quality was assessed by SDS-PAGE during each purification step. Protein concentration was determined by UV absorption at 280nm using the predicted molar extinction coefficients (Josephin ε: 25440 M^{-1} cm^{-1}, Ataxin-3 ε: 26930 M^{-1} cm^{-1}).

2.4 Screened compounds

In this work it was extensively evaluated the effect on Atx3 aggregation of Dithiothreitol (DTT), Tris(2-carboxyethyl) phosphine hydrochloride (TCEP), and 3,4-Dihydroxy-1,6-dithiane (hereafter oxidized DTT, DTTox). DTT was purchased as powder from Fisher Scientific (product 10386833), while TCEP was purchased from Sigma-Aldrich (product C4706). Stock solutions of DTT and TCEP were prepared dissolving the powder in 20 mM Sodium Phosphate, pH 6.5 or in 20 mM Hepes, pH 7.5, 150 mM NaCl, and stored at -20°C. DTTox was obtained by incubating 4.76 mM DTT solution at 60°C for 120 hours. The efficacy of the oxidation protocol was assessed by the acquisition of mono-dimensional ^{1}H NMR spectrum, as well as the stability at 37 °C of freshly dissolved DTT and TCEP. NMR experiments were acquired on 1 mM samples by 800 MHz Bruker spectrometer, in testing buffer supplemented with 10% D_2O. For evaluation of the compound's stability at 37 °C overtime, 550 µl samples were incubated in 5 mm diameter NMR tubes (Norell select series, Sigma Aldrich) inside the magnet maintained at 310 K. Sequential mono-dimensional ^{1}H NMR spectra were acquired every 6 h up to 160 h. Structure and representative ^{1}H NMR spectrum at 37 °C for each of the three afore-mentioned compounds are shown in Fig. 2.3.

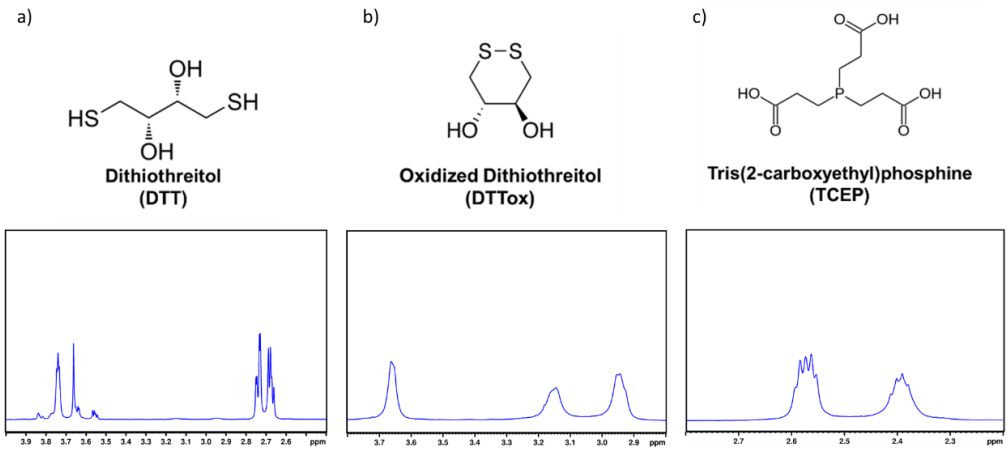


Figure 2.3 Structures and representative 1D ^{1}H NMR spectra of screened compounds. a) DTT; b) DTTox; C) TCEP. NMR spectra were acquired at 800 MHz proton frequency in 20 mM Sodium Phosphate buffer at 37°C.

2.5 Thioflavin T aggregation assays

Protein aggregation was monitored by Thioflavin T (ThT, Sigma-Aldrich T3516) fluorescence emission, still considered the golden standard approach to assess amyloids formation kinetics.[256–258] ThT works as a molecular rotor: when it is bound to cross-β sheets structures, the torsional motion of the ThT benzothiazole and aminobenzene rings relative to each other is decreased, favouring the reaching of excited states, giving a strong fluorescence signal whose intensity in first approximation is proportional to the amyloid structures content of the sample (Fig. 2.4).[259]

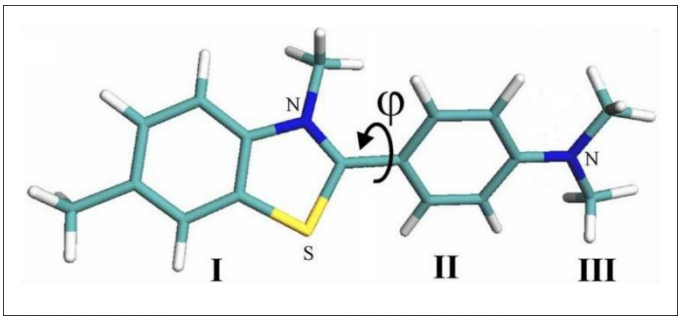


Figure 2.4 Structure of Thioflavin T molecule. ThT molecule is characterized by a benzothiazole ring (I), a benzene ring (II), and a dimethylamino group (III). Atoms forming torsion angles ϕ (N5-C6-C12-C13) are used to characterize the potential ThT configurations in the ground and the excited states. Image from Sulatskaya *et al.,* 2010.[260]

ThT fluorescence intensity measurements were performed either at 50°C or at 37°C on freshly purified protein samples concentrated 14 µM, in aggregation buffer containing 26 µM ThT.

Experiments at 50 °C were performed in large volume sample (1.2 ml) in 20 mM Sodium Phosphate pH 6.5, supplemented with 1 mM TCEP. Protein solution was incubated at 50 °C, without stirring, in 1 cm path quartz in Jasco FP-8500 spectrofluorometer equipped with a Jasco ETC-815 Peltier as temperature controller. Spectra were recorded every 5 minutes under excitation at 440 nm in the range 435‑700 nm, using 0.5 nm wavelength intervals, excitation bandwidth of 2.5 nm, emission bandwidth of 5 nm, scan speed of 100 nm/min, and integration time of 1 s. For each sample, absorption and fluorescence spectra were acquired at room temperature both before and after the kinetics experiments to exclude a significant effect of ThT hydroxylation,[261] which may hinder a reliable evaluation of the fibrils amount.

Measurements at 37°C were performed in static mode using a CLARIOstar Plus MicroPlate reader (BMG LabTech), incubating the samples in sterile clear-bottomed 96 well black plates (Perkin/Elmer), sealed with a transparent tape to prevent samples evaporation (Perkin Elmer, 6050185). Measurements were performed at least in triplicate, using 200 µl sample for each replicate. ThT emission was measured every 30 minutes at 480 nm under excitation at 440 nm, from the plate bottom. Aggregation kinetics experiments were performed either in 20 mM Sodium Phosphate, pH 6.5, or in 20 mM Hepes, pH 7.5, 150 mM NaCl.

Compounds for screening were added to the sample at a final concentration ranging from 0.25 mM to 1 mM.

2.6 Atx3-Q14 stability tests

Atx3-Q14 stability at 37 °C and 50 °C was monitored by timepoints evaluation by SDS-PAGE. 14 µM freshly purified protein was diluted in 20 mM Sodium Phosphate, pH 6.5 supplemented with 1 mM TCEP and aliquoted in PCR tubes, fulfilling them up to 330 µl and sealed such to prevent

evaporation. Protein samples were incubated in parallel in two static thermo-blocks, one set at 37 °C and one at 50 °C. Aliquots of 5 µl were then taken from the bottom of each PCR tube, corresponding to a different timepoints (i.e. 2 h, 6 h, 12 h, 24 h, 48 h, and 72 h), and run on SDS-PAGE.

The effect of different reducing agents on Atx3-Q14 stability during incubation at 37°C was monitored by timepoints evaluation by SDS-PAGE. 14 µM freshly purified protein was diluted in 20 mM Sodium Phosphate, pH 6.5 supplemented with 1 mM TCEP or 1 mM DTT. Stock solution was aliquoted in PCR tubes, fulfilling them up to 330 µl, and sealed such to prevent evaporation. Aliquots of 5 µl were then taken from the bottom of each PCR tube, corresponding to a different timepoints (i.e. 61 h, 67 h, 79 h, 85 h, 97 h, and 160 h) and run on SDS-PAGE.

2.7 Atomic Force Microscopy measurements

Aggregates morphology of Jos wild type incubated in different reducing conditions, as well Jos in complex with di- and tri-Ub chains, was investigated by Atomic Force Microscopy (AFM). Protein samples were deposited on mica substrate previously functionalized with 1 M $MgCl_2$, to favour protein adhesion to the substrate. The endpoints of the ThT kinetics were diluted in 20 mM Hepes, pH 7.5, and. deposited on mica substrate. The excess of solution was removed after 10 min of incubation, then sample-loaded mica was rinsed with milliQ water and gently dried by flushing gaseous N_2. All the phases of sample preparation were performed at RT. AFM measurements were carried out in the air using a Bruker FAST-SCAN microscope. Scans were obtained in the soft tapping mode using a FAST-SCAN-A probe with an apical radius of about 5 nm. Each image was obtained with a resolution comparable to the tip radius, spring constant of 18 N/m, and resonance frequency of 1,400 KHz. Each sample was typically characterized by acquiring images from three distinct regions of the sample. Dilutions were typical for each sample. In detail: Jos incubated with different reducing agents (1:100), monomeric Jos (1:500), triUb (1:1000), complex (1:1000).

2.8 Saturation Transfer Difference NMR

Saturation Transfer Difference NMR (STD-NMR) experiments were employed to assess the binding affinity of monomeric Jos for DTT, TCEP, or DTTox, being STD-NMR one of the most suitable NMR techniques to study weak ligand-protein interactions in solution (equilibrium dissociation constant (K_D) ranging from 10^{-8} mol L^{-1} to 10^{-3} mol L^{-1}). The main advantage of this technique consists of focusing just on the NMR signals of the ligands, requiring only a small amount of non-labelled protein. Two experiments are acquired for STD-NMR analysis: i) a spectrum in which the protein is selectively saturated (i.e. on resonance spectrum), by irradiating at a region of the spectrum that contains only resonances typical of the protein; ii) a spectrum recorded without protein saturation

(i.e. off resonance or reference spectrum). It is then performed a subtraction of the on resonance spectrum to the reference one. Based on the existence of an equilibrium among free and bound ligand, the difference spectrum reports only the signals of the ligand that received saturation during the irradiation. This irradiation transfer happens only if protein-ligand binding occurs.[262,263]

Here. samples for STS-NMR experiments were prepared adding 10 µM freshly purified protein to 500 µM compound to ensure a protein: ligand stoichiometric ratio of 1:50. Experiments were recorded at 310 K on an 800 MHz Bruker NMR spectrometer locking on D_2O added to the samples. Saturation was achieved with irradiation frequency of – 40.0 ppm (off resonance experiment) and 9.00 ppm (on resonance experiments) for DTT and DTTox, or 10.00 ppm for TCEP. All the experiments were acquired using 3 s saturation time and 256 scans. Free compound samples were used as controls. All experiments were performed in 20 mM Sodium Phosphate, pH 6.5.

DTTox binding affinity for Jos aggregates was also assessed. In this case, 14 µM of freshly purified monomeric Jos was added to 700 µM DTTox, still maintaining a 1:50 protein: ligand ratio. The mixture was incubated at 310 K into the NMR spectrometer and monitored for 160 h, acquiring one STD-NMR experiment each hour.

2.9 NMR Chemical shift perturbation analysis

NMR Chemical Shift Perturbations (NMR-CSP) analysis was employed to test the binding of DTT, TCEP, and DTTox to the Jos domain, being NMR-CSP a pivotal approach in elucidating molecular interactions, particularly between small molecules and proteins. Indeed, when a ligand binds to a protein, it induces changes in the local electronic environment of specific atomic nuclei within the protein. These alterations manifest as shifts in the NMR resonance frequencies, i.e. Chemical Shift Perturbations.

The most employed NMR experiment in CSP analysis is the ^{15}N Heteronuclear Single Quantum Correlation (^{15}N-HSQC), which provides the correlation between nitrogen and amide protons. The resulting spectrum is two-dimensional (2D) with one axis for the proton (^{1}H) and the other for the ^{15}N, and contains a peak for each unique proton attached to the hetero-nucleus. Each amino acid in the protein sequence can therefore produce an observable peak in the spectrum related to its own backbone amide, except proline which misses the amide proton. In addition, sidechains with nitrogen-bound protons will also produce peaks (i.e. glutamine, asparagine, lysine, and arginine).[264]

In CSP analysis, it is performed a comparison of the HSQC of the free protein with one of the proteins bound to the ligand, which can lead to the observation of changes in the chemical shifts of some peaks. The chemical shift change, which can be measured accurately, is very sensitive to structural change and provides the information about almost genuine binding when interactions between the

protein and the ligand occur. By comparing the NMR spectra of the free protein with those in the presence of the ligand, it is then possible to identify shifted peaks, revealing those protein residues influenced by the binding event. The magnitude of the chemical shift perturbation is correlated with the strength of the interaction, aiding in the determination of binding affinities.[265] This method is very useful in structural biology and drug discovery, enabling the mapping of binding sites and the characterization of conformational changes associated with molecular recognition events. In essence, CSP analysis serves as a valuable tool for unravelling the intricacies of molecular interactions at the atomic level.

To perform CSP analysis, ^{15}N-labelled Jos was produced by growing transformed *E. coli* competent cells in minimal medium containing ^{15}N ammonium sulphate as the sole nitrogen source. Protein expression and purification were carried out as described in 2.3. Bi-dimensional ^{1}H-^{15}N HSQC spectra were recorded at 298 K on an 800 MHz Bruker NMR spectrometer. 50 μM ^{15}N Jos sample in 20 mM Sodium Phosphate pH 6.5 was progressively titrated with the compound, following stoichiometric ratio protein: compound of 1:0, 1:0.25, 1:0.5, 1: 0.75, 1:1, 1:1.5. ^{1}H-^{15}N HSQC spectra were acquired for each titration point, and processed with CcpNmr Analysis 2.4.2 software.[266] Resonances assignment was performed by semi-automatically transferring the published chemical shifts list (bmrb accession number: 6241)[7]. CSP were analysed according to the following equation:

$$\mathrm{CSP}_i = \sqrt{(\Delta\delta_{Hi})^2 + \alpha(\Delta\delta_{Ni})^2}$$

where $\Delta\delta_{Hi}$ and $\Delta\delta_{Ni}$ indicate the observed chemical shift changes for the proton and the nitrogen of the residue *i*, respectively, weighted by a factor α. According to standard practice, a cut-off greater than or equal to two folds of the standard deviation (2σ) was set to establish which chemical shift perturbations are large enough to be considered a significative indicator of binding.[264]

2.10 Josephin point mutants

Four points mutants of Jos were prepared within this project using Site Directed Mutagenesis by PCR. In each mutant, one of the four cysteine residues in Jos sequence was substituted by an alanine residue: C14A, C18A, C114A, and C172A. Dedicated primers were designed as reported in Table 2.4 and purchased from Integrated DNA Technologies, IDT.

Site-directed mutagenesis was performed using Q5 hot start high fidelity DNA polymerase (NEB). The PCR mixture was prepared according to manufacturers' instructions: 500 nM of corresponding forward and reverse primers, 0.42 ng Jos WT plasmid as template, 200 μM dNTPs, Q% buffer, Q5 GC enhancer, and 0.02 units of polymerase. After initial denaturation for 30 sec at 98°C, 30 PCR cycles were performed, annealing at 65°C for 20 sec, and 2 min 20 sec of extension at 72°C. The PCR

product was then treated with a mix of Kinase, Ligase, and DpnI enzymes from New England Biolabs (KLD Kit, NEB Cat N.° M0554S). After transformation in *E.coli* DH5α competent cells (NEB Cat N.° C2987H), plasmids were isolated from the resulting colonies by plasmid extraction and purification kit (QIAquick® Gel extraction kit, Cat N.° 28704). The desired modification of cysteine with alanine residues was confirmed by Sanger sequencing (Annex II). All the steps were monitored by agarose gel electrophoresis.

Mutation	Forward primer sequence	Reverse primer sequence
C14A	5' -GGCTCACTTGCTGCTCAACATTGCCTGAATAACTTATTG- 3'	5'-ATGTTGAGCAGCAAGTGAGCCTTCTTGTTTCTCGTGGAA-3'
C18A	5'-GCT CAA CAT GCG CTG AAT AAC TTA TTG-3'	5'-ACA AAG TGA GCC TTC TTG TTT CTC GTG-3'
C114A	5'-CAT TTA TAG CGA ATT ATA AGG AAC ACT G-3'	5'-ATC TTT CAT TTA TAG GAT CGA TCC TG-3'
C172A	5'-CCA GAT GCA GAA GCT GAC CAA C-3'	5'- CAGATCACCCTTAACAACAAATATAGAATAACCTTC-3'

Table 2.4 Primers designed for Jos point mutants to substitute cysteine residues with alanine.

2.11 Circular Dichroism

Circular dichroism (CD) measurements were performed to evaluate if Jos point mutants maintain the same secondary structure of the wild type protein. Each mutant was tested at a concentration of 14 µM in 20 mM Sodium Phosphate pH 6.5. CD measurements were performed in 1 mm quartz cuvette (Hellma Analytics), at RT, using a Jasco J-1500 spectropolarimeter equipped with a Jasco CTU-100 temperature controller. Each spectrum was recorded from 260 to 195 nm, using scan speed 50 nm/min, 2 nm bandwidth, and data pitch 0.1 nm. Each spectrum is the result of an average of 10 accumulations.

2.12 Evaluation of crosslinking disulfide bridges

Aggregation endpoint of wild type and mutants Jos samples after ThT-assays at 37°C were run on SDS-PAGE to evaluate the presence of crosslinking disulfide bridges in the final aggregates. Briefly, aliquots from final Jos samples after incubation at 37°C in different reducing conditions (i.e. in absence of reducing agents, and in presence of 1 mM DTT, 1 mM DTTox, or 1 mM TCEP) were taken and treated with non-reducing loading buffer. After boiling for 5 minutes, 5 µL samples were loaded on NuPAGE Novex, Bis-Tris 4-12% gel (Thermofisher). All gels were stained using Instant Blue staining solution (Abcam).

2.13 Lysine48-linked di- and tri-ubiquitin chains production and purification

Lysine48-linked di- and tri-ubiquitin chains (K48-diUb and K48-triUb, respectively) were produced to investigate their binding mode to Jos and their impact on Jos pathological aggregation. The chains were produced by enzymatic reaction, according to the protocol first reported by Pickart group with minor adjustments (Fig. 2.5).[267–270] Briefly, 25 mg of lyophilized monoUb, purchased by Sigma (Cat N.° U6253) and derived from bovine erythrocytes, was resuspended in 50 mM Tris-HCl, pH 7.5 obtaining a 2.5 mM monoUb solution. This was used as a substrate for the polymerization reaction acted by E1 and CdC34 enzymes. Reagents were mixed together as concentrations reported in Table 2.5, and incubated at room temperature for 1 hour to allow the reactions summarized in Fig. 2.5. Reaction was then stopped by adding DTT at a final concentration of 50 mM, and the mix was incubated on ice for 30 min to release all E1- and E2-linked polyUb chains. Precipitated enzymes were then separated by centrifugation at 5000 rpm for 5 min at 4 °C, and discarded, while K48-diUb and K48-triUb were recovered from the supernatant.

monoUb	480 µM
E1	2 µM
CdC34	40 µM
$MgCl_2$/ATP	10 mM
DTT	0.5 mM
Reaction buffer 10 x	1 x

Table 2.5 Reaction mix for Ub polymerization. Reaction buffer 1x consists of 50 mM Tris-HCl, pH 7.5

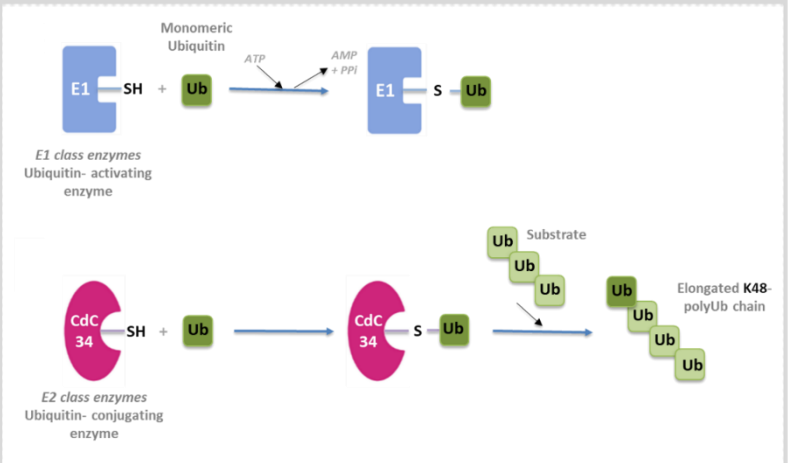


Figure 2.5 K48-polyUb chains production by enzymatic reaction. E1 is the ubiquitin-activating enzyme and CdC34 is the specific ubiquitin-conjugating enzyme for Lysine48-linkage. The formation of the polyUb chain occurs through the linkage between the lysine 48 residue of one Ub monomer and the C-terminal glycine residue of another Ub monomer. In detail, in the first step, the C-terminal Gly residue of one Ub moiety is activated in an ATP-dependent process by E1. Ub binds a Cys residue of E1 forming a thiol-ester linkage, releasing AMP. In the second step, activated Ub is next transferred to an active Cys residue of E2 (Cdc34), that directly transfers Ub to the lysine 48 of a successive Ub in order to obtain Lys48-linked polyUb chains.[271]

The separation of ubiquitin chains of different lengths was performed by ion exchange chromatography. PolyUb chains pool was diluted 1:15 with 50 mM Sodium Acetate, pH 4.5, and loaded onto a 5 ml prepacked cationic exchange chromatography Hi trap SP column (Cytiva, Cat. N. 17115201). A linear gradient of 48 column volumes (CV) of 50 mM Sodium Acetate, 0.5 M NaCl, pH 4.5 was performed to separate the polyUb of different lengths. The eluted sample was then concentrated and subjected to size exclusion chromatography (SEC) for final polishing. SEC was performed in 20 mM Hepes, pH 7.5, 150 mM NaCl, and 1 mM DTT using Increase 10/300 75s (Suppl. Fig. S5). Fractions were pooled based on SDS-PAGE analysis. Sample quantification was assessed by BCA assay (Thermo Scientific), because of the absence of tryptophan in the protein sequence. K48-triUb pool was lyophilized by a vacuum concentrator system (Speedvac, Eppendorf) and stored at -80°C until use.

2.14 Preparation of Jos-diUb and Jos-triUb complexes

Jos-diUb and Jos-triUb complexes were produced for SAXS measurements, ThT assays, and crystal trials. Jos point mutant C14A was used for sample preparation, inhibiting the catalytic activity of Jos for cleavage of polyUb substrate. Freshly purified Jos C14A was incubated for one hour at RT either with K48-diUb or K48-triUb in a molar ratio of 1:1.2 (in the case of ThT assays and SAXS) or 2:1 (in the case of crystallographic trial). Complexes were formed in different buffers, according to the technique used for the following characterization: 20 mM Tris, pH 6.5, 1% sucrose, 2 mM TCEP (SAXS), 20 mM Hepes, pH 7.5, 150 mM NaCl (ThT assays, Suppl. Fig. S6) and 20 mM Sodium Phosphate, pH 6.5, 1 mM TCEP (crystal trials). After complex formation, the component in excess was removed by SEC. Cytiva Increase 10/300 75s column was used in the case of ThT assays and crystal trials samples. Agilent AdvanceBio SEC 130Å column was used in the case of SAXS. Samples quality was assessed by SDS-PAGE analysis. Concentration was measured by UV absorption at 280 nm using the predicted molar extinction coefficients (complex Jos-diUb MW: 38.1 kDa, ε: 28420 M^{-1} cm^{-1}; complex Jos-triUb MW: 46.6 kDa, ε: 29910 M^{-1} cm^{-1}).

2.15 SAXS data collection and analysis

Small-angle X-ray scattering (SAXS) was employed to characterize Jos-diUb and Jos-triUb complexes structures. SAXS data were collected at the European Synchrotron Radiation Facility (ESRF), BioSAXS beamline 29 (BM29) at 20°C and at a wavelength of 0.99 $Å^{-1}$, with an incident energy of 12.5 keV. To exclude aggregated species, the samples were subjected to SEC in line with SAXS. In detail, samples were applied to an AdvanceBio SEC 130Å (Agilent) at a concentration of 5 mg/ml and run at a flow rate of 0.2 ml/min in 20 mM Tris pH 6.5, 1% sucrose, and 2 mM TCEP.

During the elution, 750 frames were taken with 2 s exposure. Individual frames were systematically controlled for systematic changes and averaged for better statistics. Analogous patterns of the buffer were collected before and after every collection of the samples and used to subtract any contribution from the solvent and the sample environment. Absolute intensity units were determined by measuring the scattering signal of water (0.01632 cm^{-1} at 20 °C). The ESRF in-house software BsxCuBE (Biosaxs Customized Beamline Environment), connected to EDNA data processing pipeline,[272] was used to control the real-time data display (two dimensional and one dimensional), and to provide the first automatic data processing. Normalization of beamline to water and calibration were performed. The averaged buffers were then subtracted from the averaged samples to give a final subtracted data set for each sample. Java-based software ScÅtter IV (https://bl1231.als.lbl.gov/scatter/) was used to upload SAXS buffered subtracted data files (*.dat), and perform primary data analysis including Guinier analysis, normalized Kratky plots, Porod-Debye volume and Dmax calculations (Fig. 2.6).[273] Further data analyses were performed by the online-available ATSAS package version 3.2.1 (https://www.embl-hamburg.de/biosaxs/software.html). The maximum distance (Dmax) was selected by letting the P(r) curve decay smoothly to zero. R_g was estimated from the SAXS profile according to Guinier's estimation.

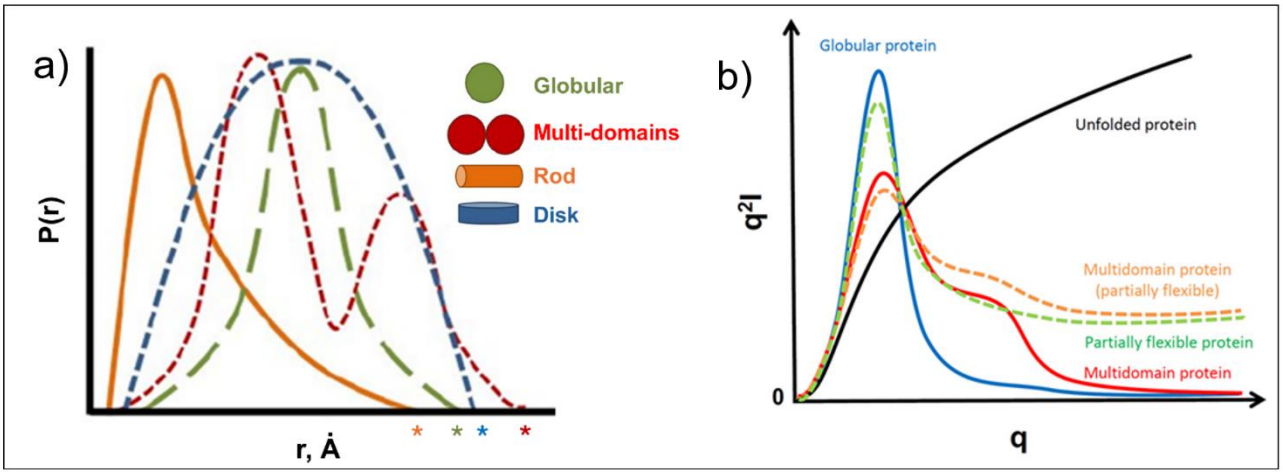


Figure 2.6 Schematic representation of P(r) and Kratky plot. The scattering patterns generated from a dilute solution of macromolecules are typically presented as radially averaged one-dimensional curves, that are used to extract information about a three-dimensional particle. When interpreting SAXS data, more curve-fitting routines are applied to lend robustness to the conclusions.[274] Two possible representations of SAXS data are the Pair distribution function (a) and the Kratky plot (b). a) Pair distance distribution functions represent the distribution of the intra-atomic distances (r) in the particle, extracted from the scattering curve I(q), which in turn represents its Fourier transform. Where P(r) goes to zero is the maximum linear extent (Dmax), or roughly the diameter of the scattering species. P(r) is indicative of the particle shape: globular particles yield bell-shaped profiles (green) with a maximum at approximately Dmax/2, while multi-domain particles (red) often yield profiles with multiple shoulders and oscillations corresponding to intra and inter-subunit distances. Asterisks indicate the corresponding Dmax.
b) Kratky plot calculation provides information as to the conformation of the molecule and whether the molecule is globular, extended, or highly flexible. Proteins with highly extended conformations are characterized by large average sizes compared to globular proteins with a tightly packed core. In the case of a globular protein, the Kratky plot appears as a bell-shaped Gaussian (blue). Deviations from this behaviour suggested a higher level of flexibility. Indeed, the Kratky plot will not converge to the q axis if the protein has a pronounced flexibility (green, orange). Proteins with multiple domains could have additional peak(s) (or shoulders) at low q but the Kratky plot will still converge to the q axis at high q (red). Adapted from Jackson *et al*., 2015[274] (a) and from https://www-ssrl.slac.stanford.edu[275] (b).

2.16 Molecular modelling and SAXS data fitting

To investigate K48-diUb and K48-triUb binding mode to Jos, preliminary complexes models were created using rigid body modelling software available in the ATSAS package: CORAL[276] and SASREF.[277] For Jos-diUb complex unrestrained modelling was performed using the NMR-based structure of Jos (PDB: 1yzb)[7] and the crystallographic structure of Lys48-linked diUb in the open conformation (PDB: 3aul).[278] For restrained modelling, constraints among Ub binding sites 1 and 2 of Jos, and diUb were imposed based on PDB: 2jri. This derives from a HADDOCK structure of Jos in the presence of two monomeric Ub.[279] For Jos-triUb complex modelling, firstly monoUb structure PDB: 1UBQ[31] was used to produce a Lys48-triUb chain by CORAL, imposing as restraints the G76-Lys48 linkage and considering the last five residues of each monomeric Ub as dummy residues. Then, the CORAL-derived triUb structure was used as entry PDB for further unrestrained Jos-triUb complex models. DENSS (DENsity from Solution Scattering, https://tdgrant.com/) was used as an algorithm for calculating 3D particle electron densities directly from 1D solution scattering. The final figures were generated using UCSF Chimera 1.16.[280] Volume correlation was calculated according to DENSS manual instructions.

2.17 Crystal trials for Jos-triUb complex

Jos-triUb complex was prepared as described in 2.14. SEC fractions of the complex complex were pooled together and concentrated up to 11.5 mg/ml by Vivaspin™ 500 Centrifugal Concentrators (Sartorius). Crystallization trials were performed at the High-throughput crystallisation and ligand screening (HTX) facility at EMBL, in Grenoble. Access was granted by the Instruct-ERIC application. Six different 96 well kits were screened: Classics-Suite (QIAGEN), JCSG (Hampton Research, Cat. N° HR2-145), and PACT (Molecular Dimensions, Cat. N° MD1-36-FX) as standard screening kits; Midas-plus (Molecular Dimensions, Cat. N° MD1-107-FX), Morpheus (Molecular Dimensions, Cat. N° MD1-47-FX), and ProPlex (Molecular Dimensions, Cat. N° MD1-42-FX) as optimized screening for complexes. 2 µl of each sample were taken to run a quality control assay based on thermal shift assay. Crystallization set-up was performed in automated mode using a 1:1 sample: precipitant ratio (100 nl/ well) at RT.[281]

3 Results and Discussion

In the following sections, the results obtained in this work are described and discussed.

The first part of the project aimed at investigating the molecular mechanisms of Ataxin-3 aggregation, trying to elucidate its supramolecular assembly process *in vitro*. In particular, it was investigated the effect of temperature and buffer composition on different steps of Atx3 self-assembly, being two variables strongly influencing protein stability, folding, and aggregation propensity. This had the double aim of revealing new insights on Atx3 aggregation, and setup a robust and reproducible aggregation assay for screening anti-aggregating compounds.

Protein aggregation was followed by ThT-based fluorescence spectroscopy, and morphological analysis of final aggregates was performed by atomic force microscopy. Several variants of Atx3 were investigated, including Josephin domain, non-expanded Atx3-Q14, and two distinct expanded Atx-3 variants (i.e. the mild toxic Atx3-Q55, and the fully toxic Atx3-Q70). This was possible thanks to the employment of DNA recombinant technology which allowed the in-house preparation of the recombinant proteins. The high purity grade of the samples, crucial for *in vitro* aggregation studies, was achieved through sequential steps of purification by chromatographic techniques.

Moreover, this PhD project also aimed at characterizing Ataxin-3 in complex with polyubiquitin chains, one of the recognized Atx3 partners,[5–7] and investigating the effect of these chains on Ataxin-3 aggregation. Indeed, evidence suggests an existing competition among proteins' normal function and their tendency to aggregate, thus suggesting that exploring interactions between neurodegeneration-related proteins and their natural binders might offer promising therapeutic perspectives.[8–10] To evaluate if the specific interaction among Ataxin-3 and poly-ubiquitin chains could affect the aberrant aggregation, may thus represent the first piece of the puzzle for the design of potential therapeutics against Spinocerebellar Ataxia type 3.

Here, Lysine48-linked di- and tri-Ub chains were prepared in-house by enzymatic reaction, and gold-standard Structural Biology techniques, such as Small Angle X ray Scattering and X ray crystallography, were employed to assess sample quality and to try to solve the structure of the complex. In parallel, the effect of di- and tri-Ub chains on Atx3 aggregation was investigated. Even in this case, the protein aggregation was followed by ThT-based fluorescence spectroscopy and morphological analysis of final aggregates was performed by atomic force microscopy.

3.1 New insights on Ataxin-3 aggregation

As detailed in the Introduction chapter, despite extensive exploration of Atx3 aggregation processes,[149,152,153,215,216,218,226–228,230,282–286] many aspects of the intricate pathway of Atx3 self-

assembly remain controversial, hampering the development of targeted therapeutic interventions and emphasizing the need for further investigation.

This study focussed on the effect of temperature and buffer composition on the Atx3 aggregation process, gaining new insights for delving into this complex mechanism and finding conditions that can modulate the pathology at the molecular level.

3.1.1 Atx3 aggregation at 50°C is polyQ-independent

In the present study, the aggregation kinetics of Atx3 variants with distinct polyQ tract length was investigated at 50°C using ThT-based assay, being ThT a fluorescent probe known to selectively bind amyloid structures.[259] In particular, the following variants were used: Atx3-Q14, Atx3-Q55, and Atx3-Q70, where the number associated with "Q" reflects the number of glutamine residues within the polyQ tract. The three variants were chosen to represent the non-pathological, mildly toxic, and fully toxic Atx3, respectively.[166,167,287–290]

The selected Atx3 variants contain three UIMs in the C-terminal region, as the isoform predominantly expressed in brain tissue,[291,292,] and were prepared recombinantly expressing them in *E.coli* cells. The expression vectors for Atx3-Q14 and Atx3-Q55 were already available, as well as optimized protocols for their expression and purification. The production of Atx3-Q70 required, instead, cloning work for producing the correspondent expression vector, and the setup of the expression and purification procedures.

The manipulation of the length of CAG repeats encoding for the Q70 tract was not an easy task, because of the hairpin structure formed by the repeats,[293] DNA polymerase slippage on repeated sequences,[294], and sequence instability in cells.[295] To face these issues, the expansion of the polyQ tract was achieved by the *in vitro* method known as Synthesis of Long Iterative Polynucleotide (SLIP), optimizing the approach used in Figura *et al.,* 2015.[253] This approach takes advantage of highly repeated CAG/CAA sequences tendency to form hairpin structures, and consists in repeated cycles of plasmid digestion with a restriction enzyme, a single PCR cycle, bacterial transformation, and subsequent colony screening to identify colonies that present an altered number of repeats.

In more detail, the cDNA encoding for Atx3-Q55 was firstly transferred into pGEM-T-Easy vector between *Bsm*BI and *Eco*O109I restriction sites, according to Figura *et al.* 2015 procedure.[253] Two aliquots of the starting construct were independently and parallelly digested, one with *Bsm*BI and one with *Eco*O109I restriction enzyme, allowing cutting of the construct upstream and downstream to CAG expansion, respectively. Then, the two digested products were processed together by a single cycle of PCR to allow the filling of gaps due to imperfect annealing of DNA. To ensure the removal of not annealed fragments of the specific sequence flanking the repeated tract, a polymerase with 3'

exonuclease activity was used. The PCR product was then used to transform competent *E.coli* cells, and contraction/expansion of CAG triplets was evaluated by colony screening (Fig. 3.1a). As a result of an extensive screening campaign, colonies expressing more expanded forms of Atx3 were identified (Fig. 3.1b), and the precise length of the polyQ tract was confirmed through Sanger sequencing (Annex I). The final cDNA encoding Atx3-Q70 was successfully transferred in pMAL expression vector (Fig. 3.1c).

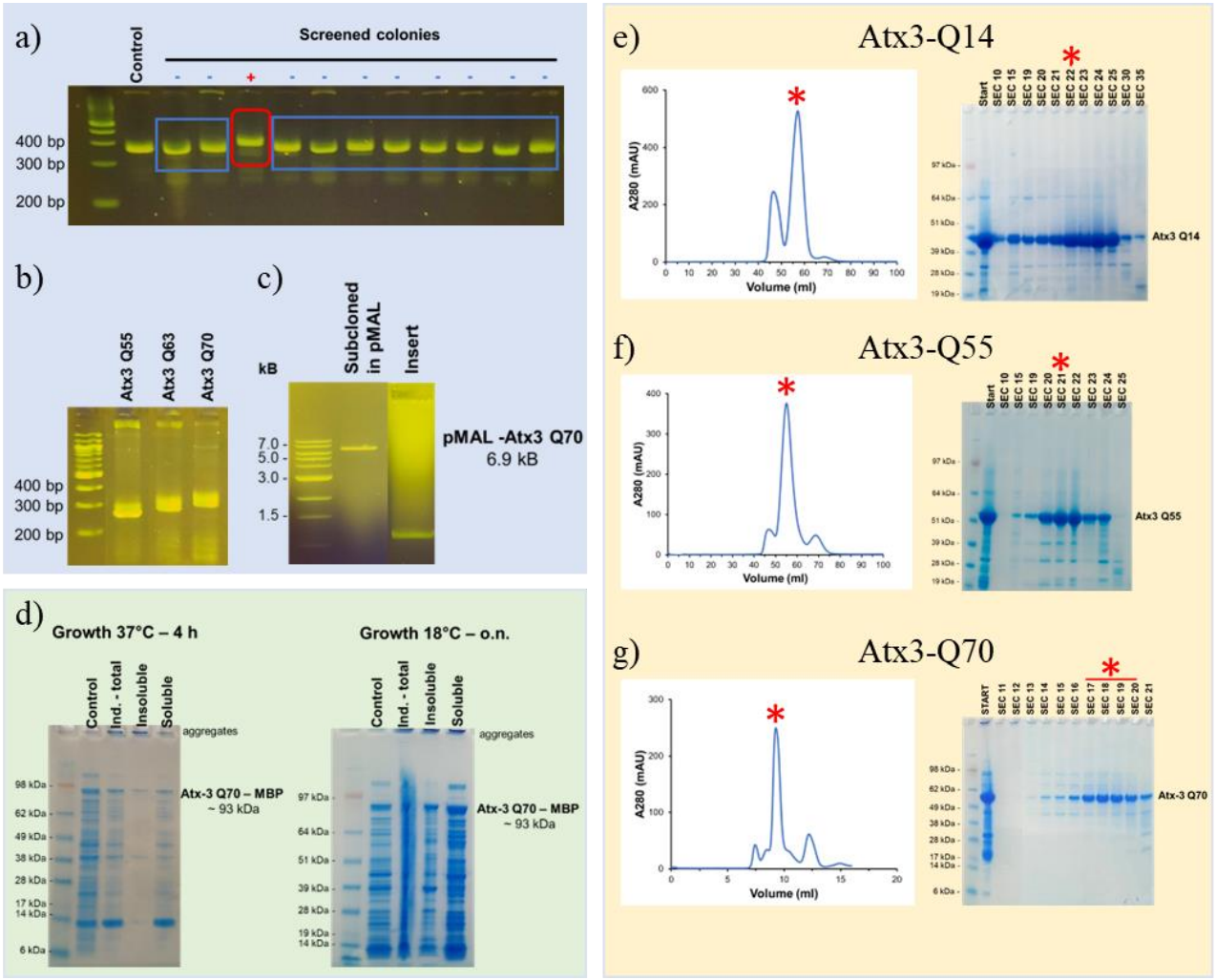


Figure 3.1 Production of recombinant Atx3 variants. Expression plasmids for Atx3-Q14 and Atx3-Q55 were already available in Dr Alfano's group (Fondazione Ri.MED). Plasmid for Atx3-Q70 was instead cloned within this project by the SLIP method, and here it is reported: a) 3% agarose gel of colony screening by PCR reveals CAG triplets' expansion; b) Agarose gel of polyQ expanded constructs. The Q55 starting construct is shown in lane 1. Q70 constructs (lane 3) were obtained by several cycles of expansion passing through an intermediate Q63 expanded form (lane 2); c) Agarose gel that shows the correct transferring of expanded Atx3-Q70 construct in pMAL expression vector.
Expression tests for pMAL-Atx3-Q70 were then conducted in *E.coli* Rosetta2 pLysS cells at two different temperatures. Evaluation was performed by SDS-PAGE analysis, here reported in d) IPTG induction for 4 h at 37°C, on the left and IPTG induction o.n. at 18°, on the right. Not induced cells were used as negative control (Control). Induced cells were lysed and analysed such as (Ind. - total) and after separation by centrifugation of Soluble and Insoluble fractions.
All Atx3 variants samples underwent SEC immediately before the aggregation assay, to avoid the impact of possible seeds of aggregation and to ensure sample quality. Here, SEC profiles and related SDS-PAGE are reported in e) Atx3-Q14; f) Atx3-Q55; g) Atx3-Q70. Asterisks correspond to the purest fractions selected for the ThT assay. SEC was performed by Hi load 16/600 75s for Atx3-Q14 and -Q55, and with Increase 10/300 75s for Atx3-Q70.

Once completed the molecular cloning for Atx3-Q70, all the three Atx3 variants were expressed and purified to proceed with ThT-assays. The preparation of recombinant Atx3-Q70 required optimization of the protocols already available for both Atx3-Q14 and Atx3-Q55. In particular, protein expression conducted at 18 °C overnight resulted in increased production of the protein in the soluble fraction of bacterial lysate with respect to expression at 37 °C (Fig. 3.1d). All the purification steps were performed at 4°C and in the presence of 5% glycerol to reduce protein aggregation. Moreover, due to the high contamination by nucleic acids, that likely remain attached to the protein because of the long positively charged polyQ tail, both DNase and RNase were used in the cells lysis step. The remaining nucleic acid contaminants were removed by an additional anionic exchange chromatography step with Capto Q column (Suppl. Fig. S4).

All protein samples (i.e. Atx3-Q14, Atx3-Q55, and Atx3-Q70) underwent Size Exclusion Chromatography (SEC) immediately before the ThT aggregation assay, to avoid the impact of possible seeds of aggregation and to ensure sample quality (Fig. 3.1e-g).

The aggregation kinetics measurements were carried out in large volume (3ml) incubating 14 µM monomeric protein at 50°C in 20 mM Sodium Phosphate pH 6.5 supplemented with 1 mM TCEP. The three Atx3 variants resulted to exhibit aggregation with comparable kinetics despite the difference in polyQ length, suggesting that at 50°C the aggregation mechanism is mainly controlled by the solely Jos domain, while the contribution of polyQ length is abolished (Fig. 3.2). Moreover, the aggregation kinetics lack of a lag phase indicating a fast aggregation process mediated by heterogeneous nucleation.

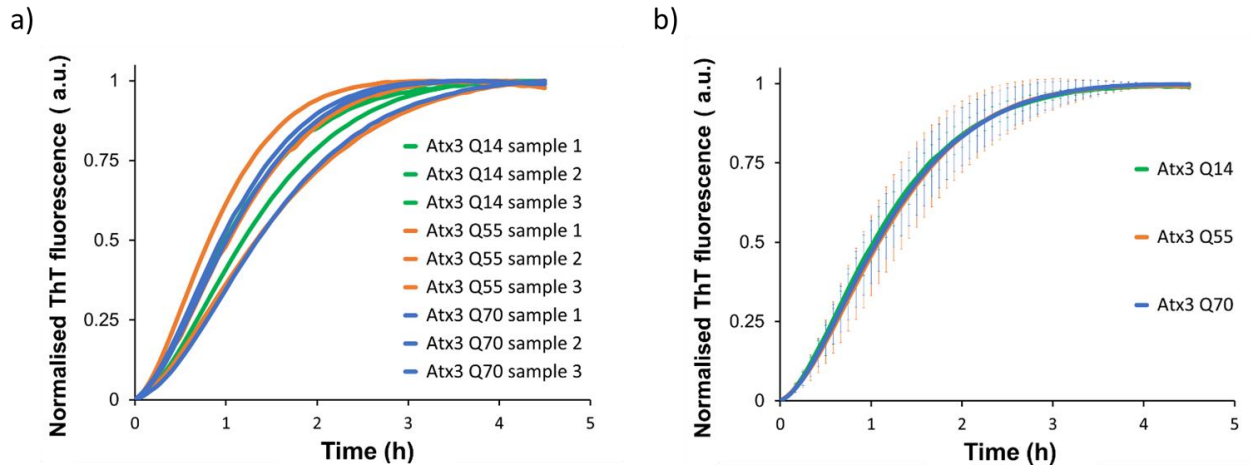


Figure 3.2 ThT aggregation assays of Atx3-Q14, -Q55, and -Q70. Time evolution of the ThT fluorescence intensity for 14 µM samples incubated at 50 °C for ~6 hours (spectra are acquired every 5 minutes in the detection range 435 – 700 nm, under λ_{exc}: 440 nm). a) Single curves for each experiment are reported. Each Atx3 form was tested in triplicate. ThT intensity at 480 nm is reported after normalisation to the maximum value and plotted against time. The slight level of variability among replicates is consistent with the use of in-house prepared protein batches, which may lead to minor events of stochastic nucleation. b) Each curve represents the mean of three independent experiments. Error bars represent the standard deviation of replicates.

These findings deviate significantly from documented results on Atx3 aggregation at 37°C.[149,152,153,226] Indeed, Atx3 fibrillogenic pathway has been characterized kinetically at 37°C as polymerization mediated by homogeneous nucleation, and consisting in a two-stages aggregation mechanism: the first stage is polyQ-independent and mediated by the Jos domain, while the second stage is mediated by the expanded polyQ stretch. Accordingly, deep differences in the aggregation kinetics at 37°C among non-expanded and expanded Atx3 have been reported.[152,153,217,225,226]

At 50 °C, the differences in the aggregation kinetics among non-expanded and expanded Atx3 are ironed out, and the aggregation process results are mediated by the solely Jos domain. Performing ThT-assays at 50°C can then result very useful when investigating the effect of anti-aggregating molecules designed to target the early phase of Atx3 self-assembly, which is mediated by Jos. Indeed, aggregation kinetics performed at 50°C reach completion in less than 5 hours for all three variants, while they last 35 hours at 37 °C when investigating non-expanded Atx3 variants.[152,153] This is because higher temperature reduces electrostatic repulsion and increases random Brownian motion and collision frequency.[296] Measurements at 50°C are thus much less time-consuming, resulting in very convenient experimental setup when a considerable number of potential anti-aggregating molecules needs to be screened.

3.1.2 At physiological temperature, Atx3 hydrolysis competes with aggregation of full-length Atx3

To be able to screen also molecules targeting the aggregation phases mediated by the polyQ trait, here ThT assays at 37°C were also attempted. Measurements were performed in small volume (200µl) using a CLARIOstar Plus microplate reader (BMG Labtech), incubating Atx3-Q14 samples at 37°C in the same buffer used for previous experiments at 50°C (i.e. 20 mM Sodium Phosphate buffer, pH 6.5, supplemented with 1 mM TCEP).

Surprisingly, the aggregation kinetics profile showed profound differences compared to what was already reported in the literature at the same temperature for Atx3.[152,153] In particular, after 160 h of continuous incubation at 37°C, ThT fluorescence value remains quite low, never reaching the plateau phase during the observed period (Fig. 3.3a). In contrast, published data report that, in similar experimental conditions, non-expanded Atx3 shows a sigmoidal aggregation kinetics profile typical of a nucleated growth polymerization mechanism, with a plateau phase reached in about 35 h.[152,153]

Atx3-Q14 stability and its progressive supramolecular assembly were then evaluated by SDS-PAGE analysis in a time-course experiment in batch at both 37°C and 50°C. More specifically, 14 µM protein stock was aliquoted in two PCR tubes, fulfilling them up to 330 µl and sealing such to prevent evaporation. The two aliquots were incubated in parallel in static thermo-blocks, one set at 37°C and

one at 50°C. 5 µl samples were then taken from each testing tube at different timepoints (i.e. 2 h, 6 h, 12 h, 24 h, 48 h, and 72 h), and run on SDS-PAGE in reducing conditions (Fig. 3.3b).

After 2 hours of incubation at 50°C, corresponding to the elongation phase of the ThT kinetics previously shown in Fig. 3.4, high molecular weight and SDS-insoluble species were detected. On the other hand, SDS-PAGE bands of the sample incubated at 37°C revealed no sign of aggregates for all tested timepoints and the presence of a prevalent species at a molecular weight much lower than the expected one for monomeric Atx3-Q14 (17-28 kDa *vs* 42 kDa).

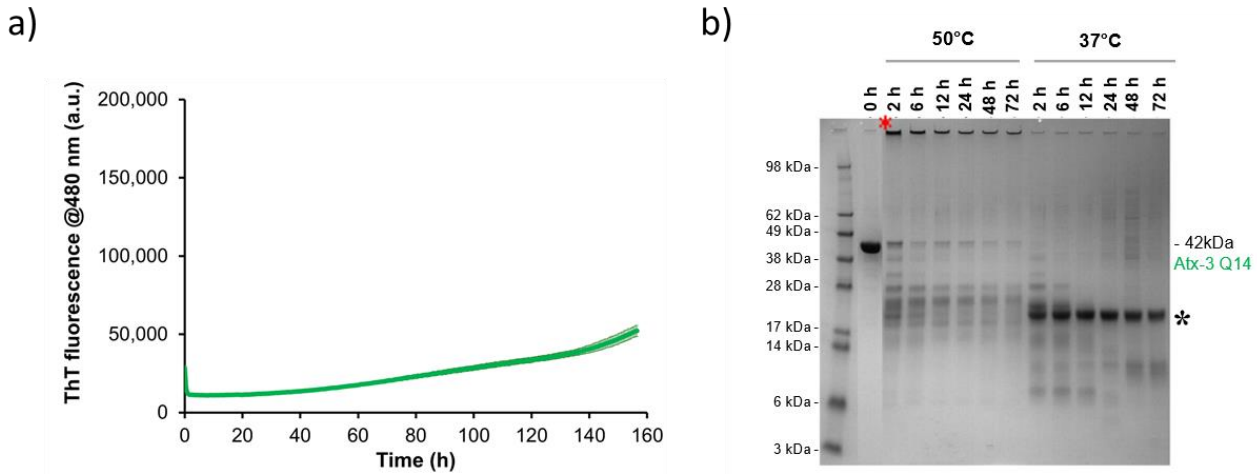


Figure 3.3 Atx3-Q14 aggregation kinetics at 37°C. a) Time evolution of the ThT fluorescence intensity at 480 nm for 14 µM Atx3-Q14 sample incubated at 37 °C for ~160 hours (spectra are acquired every 30 minutes at $\lambda_{em:}$ 480 nm, under λ_{exc}: 440 nm). Assays were performed in triplicate and the curve here reported represents the mean of the three independent experiments. Error bars represent the standard deviation of replicates. b) SDS-PAGE analysis of Atx3-Q14 stability overtime both 37°C and 50°C SDS-PAGE in reducing conditions reveals the formation of high molecular weight (MW) and SDS-insoluble aggregates for the sample incubated at 50°C, which remains trapped in the stacking part of the gel (red asterisk). For the sample incubated at 37°C, no sign of aggregates is visible, and instead, a prevalent band appears at a MW much lower than the one expected for monomeric Atx3-Q14 (17-28 kDa *vs* 42 kDa, black asterisk).

Intriguingly, the lower MW species observed on SDS-PAGE runs very similar to the solely Jos domain which has a MW of ~ 21 kDa, suggesting that at 37 °C Atx3-Q14 is subjected to a cleavage process and loses its intrinsically disordered C-terminal region. This agrees with previous findings showing that Atx3 has also autoproteolytic activity, sustained by the same residues responsible for its deubiquitinating activity.[297] The cleavage attack is reported to interest the carboxyl side of apolar, acidic, and polar uncharged residues, clustered in the unstructured C-terminal domain of the protein. The Josephin domain is thus preserved from auto-hydrolysis. [297]

According to aggregation kinetics data (Fig. 3.3a), the product resulting from cleavage at 37 °C was also able to slowly form aggregates, which were not revealed in the SDS-PAGE analysis likely because of their SDS-soluble nature. The kinetics profile of these aggregates resulted very similar to the one of the isolated Jos domain in the same conditions (Fig. 3.4), supporting the hypothesis that the two species share most of their amino acid sequence. However, the literature reports a sigmoidal

aggregation kinetics profile typical of a nucleated growth polymerization mechanism also for the Jos domain and other Atx3 variants missing the polyQ tail.[152,153,298] It remains then puzzling why a similar profile was not observed here.

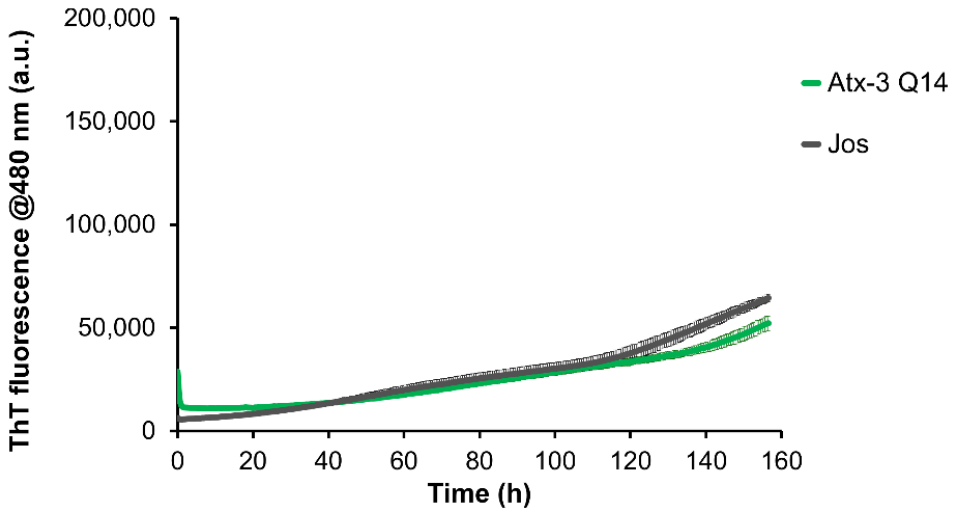


Figure 3.4 Comparison of aggregation kinetics at 37°C of isolated Jos and Atx3-Q14. Time evolution of the ThT fluorescence intensity at 480 nm for both Atx3-Q14 and Jos samples incubated at 37 °C for ~160 hours. Isolated Jos and Atx3-Q14 show very similar aggregation profiles at 37 °C, further confirming that at this temperature Atx3 undergoes to a hydrolysis process that leads to a cleaved product mostly overlapping with the Jos domain. Experiments were performed in triplicates and the curve represents the mean of replicates. The standard deviation is reported as error bar.

Altogether, the data here presented show that at the physiological temperature of 37 °C, Atx3 hydrolysis competes with aggregation of Atx3 full length. The hydrolysis leads to a domain of similar size to Jos, which is able to form SDS-soluble aggregates. However, the kinetics aggregation profile showed by this domain contrasts significantly with previous data on aggregation at 37°C of Jos and other Atx3 variants missing the polyQ trait, highlighting the need for further investigation. On the other hand, at 50°C the formation of oligomers might outpace hydrolysis events, potentially hindering Atx3 degradation through fibrillation.

3.1.3 TCEP both favours Atx3 hydrolysis at 37 °C and inhibits amyloid fibrils formation

To rationalize the findings reported in the previous paragraph, was performed a comparison among the composition of the buffer used in this PhD project until this stage (i.e. 20 mM NaPh, 1 mM TCEP), and that of buffers used in previous studies on Atx3 aggregation (Table 3.1). Indeed, experimental conditions, including buffer composition, are known to affect protein aggregation behaviour *in vitro*.[299–301] Accordingly, recent research published in 2022 by Figuereido *et al.*[298] shows that even small changes in buffer composition significantly modify Atx3 aggregation kinetics *in vitro*, altering both the oligomerization processes and the balance between ordered fibrillar structures and amorphous aggregation.

Reference	Protein variants	Buffer composition	Protein concentration	T	pH
226	Atx3-Q14	20 mM Hepes, 200 mM NaCl, 5 mM EDTA, 1 mM DTT, 1 mM PMSF	7 μM	37°C	7.5
152,153	Atx3-QHQ, Atx3-Q15, Atx3-Q64	100 mM Tris–HCl, 80 mM NaCl, 10% (v/v) glycerol, 5 mM EDTA, 15 mM βME, 2 mM PMSF	80, 60, 40, 30, 10 μM	37°C	7.4
149,231	Jos, Atx3-Q18	20 mM sodium phosphate, 10 mM TCEP, and 0.025% NaN_3	14 μM	50°C	6.5
230	Jos, Atx3-Q15, Atx3-Q64	100 mM Tris, 80 mM NaCl, 10% (v/v) glycerol, 5 mM EDTA, 15 mM βME, 2 mM PMSF	30 μM	37°C	7.5
229	Atx3-Q13	20 mM Sodium Phosphate, 150 mM NaCl, 1 mM DTT	10, 7, 5, 4, 2 μM	37°C	7.5
In this PhD project	Jos, Atx3-Q14, Atx3-Q55, Atx3-Q70	20 mM Sodium Phosphate, 1 mM TCEP	14 μM	37°C, 50°C	6.5

Table 3.1 Composition of buffers used in in vitro Atx3 aggregation experiments. Atx3-QHQ stands for Atx3 without polyQ trait.

The comparison resulted in high variability in the buffer composition among studies on Atx3 aggregation, with differences in salt concentration, pH, type of reducing agent (i.e. βME, DTT, or TCEP), and presence of additives such as glycerol, EDTA, and PMSF.

According to the findings in Figuereido *et al.*,[298] only differences regarding the type of reducing agent were worth taking into account to get clues about the Atx3 behaviour observed at 37 °C in this PhD work. Indeed, Figuereido *et al.* report that, in their assay conditions, varying the NaCl concentration between 50 and 200 mM, or the pH between 7.0 and 8.5, does not strongly affect Atx3 amyloid assembly, and neither does not EDTA and PMSF. In contrast, glycerol is reported to have a major role in decreasing ataxin-3 fibril assembly, even at concentrations as low as 0.5% (v/v).

No findings are reported about the effect of the type of reducing agent on Atx3 assembly.

Intriguingly, the reducing agent TCEP resulted in being used only in this PhD work and in previous work from Pastore's group performed at 50 °C.[149] Here, it was then investigated if TCEP might have a role in the hydrolysis of Atx3 observed at 37 °C, and consequently in the aggregation behaviour of the protein.

In more detail, firstly Atx3 stability was evaluated by SDS-PAGE in a time-course experiment in batch at 37°C in 20 mM Sodium Phosphate pH 6.5, either supplemented with 1 mM TCEP or 1 mM DTT. It resulted in a clear difference depending on the reducing agent present in the aggregation buffer. More specifically, in the presence of DTT, full-length Atx3 appears intact up to 97 hours,

whereas, following incubation with TCEP, it appears already cleaved in the first observed timepoint (Fig. 3.5a).

These results clearly suggest that the presence of TCEP correlates with Atx3 full-length degradation at physiological temperature, probably favouring its reported autocatalytic activity,[297] thus discouraging its use when studying the aggregation of Atx3 full length at 37 °C. On the other hand, Atx3 cleavage is inhibited when in the presence of DTT, suggesting that this condition is crucial to investigate aggregation of full-length Atx3 *in vitro* at 37 °C. Accordingly, the typical nucleation-mediated aggregation profile of Atx3 at 37 °C was obtained when repeating the ThT assay in the presence of DTT (Fig. 3.5b).

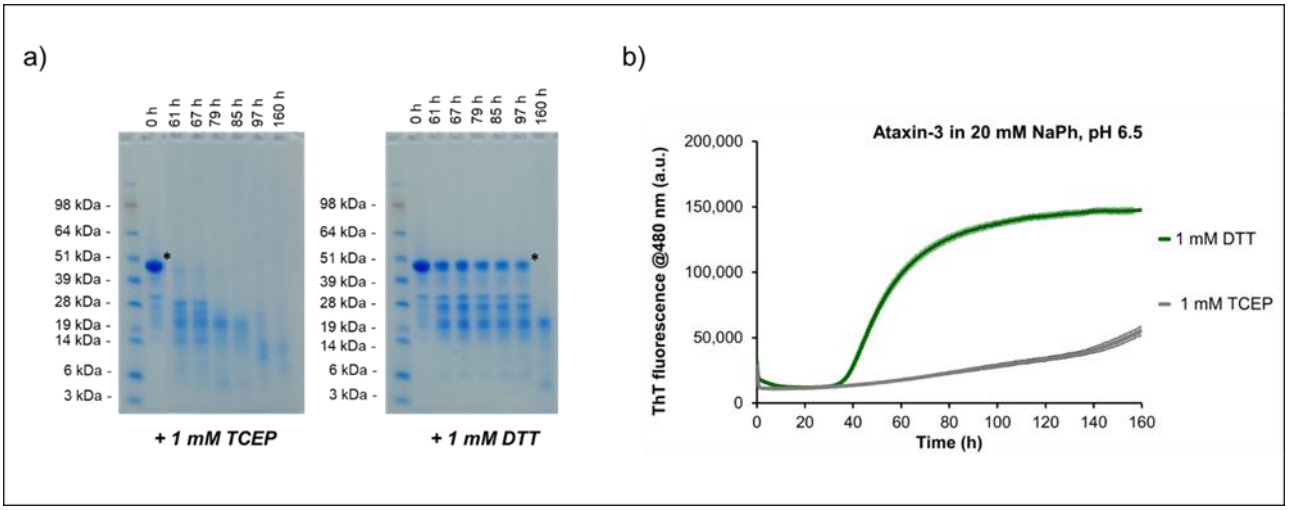


Figure 3.5 Effect of DTT and TCEP on Atx3 aggregation at 37°C. a) SDS-PAGE stability analysis of Atx3-Q14 in the presence of TCEP (left) and DTT (right). In the presence of 1 mM DTT, it is possible to notice integer bands corresponding to Atx3 full length up to 97 h; in the presence of TCEP, Atx3 is already degraded at the first observed timepoint. b) Time evolution of the ThT fluorescence intensity at 480 nm for Atx3-Q14 samples incubated at 37 °C for ~160 hours in 20 mM Sodium Phosphate pH 6.5, in the presence of 1 mM DTT (green curve) and 1 mM TCEP (grey curve). In the presence of DTT, Atx3 follows the typical nucleation-mediated aggregation mechanism previously reported in the literature. Experiments were performed in triplicates and the curves here reported represent the mean of replicates. The standard deviation is reported as error bar.

To understand if TCEP has also a role in Atx3 self-assembly, ThT assays on the Jos domain were performed at 37 °C in the presence of TCEP, in the presence of DTT, and with no reducing agent (Fig. 3.6). In this case, Jos was chosen as testing protein because of the hydrolysis activity observed for Atx3 full length when in presence of TCEP. Measurements were performed in two distinct buffer conditions (i.e. 20 mM NaPh pH 6.5 and 20 mM Hepes pH 7.5, 150 mM NaCl), which differ for the buffering species (NaPh *vs* Hepes), for pH (6.5 *vs* 7.5), and for ionic strength (0 mM NaCl *vs* 150 mM NaCl).

Interestingly, results showed significant differences depending only on the specific redox environment. In the presence of DTT, the aggregation kinetics displayed the expected sigmoidal profile for aggregation mediated by homogeneous nucleation (Fig. 3.6, green curves). Conversely, when in the presence of TCEP, the aggregation kinetics was significantly slowed down and mediated

by heterogeneous nucleation, failing to reach a plateau even after 160 hours of incubation (Fig. 3.6, grey curves). Intermediate behaviour was observed when in the absence of reducing agents and was characterized by an exponential curve indicating an aggregation process mediated by heterogenous nucleation (Fig. 3.6, red curves). The general trend was maintained in both tested buffers, with expected differences linked only to distinct environments encountered by the dye, stemming from differences in ThT binding sites that might result in varying dye quantum yields.[302]

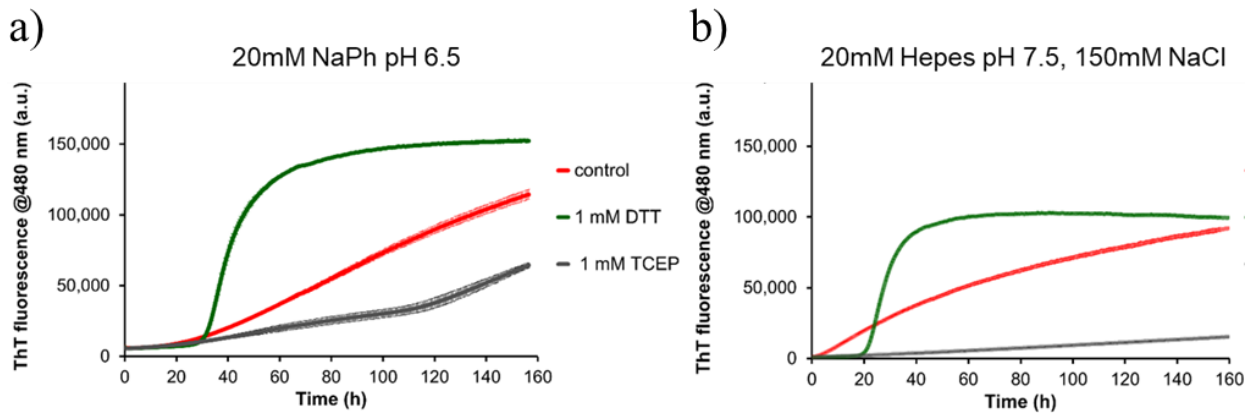


Figure 3.6 Jos aggregation at 37°C in different redox conditions. a) Time evolution of the ThT fluorescence intensity at 480 nm for Jos samples incubated at 37 °C for ~160 hours in 20 mM Sodium Phosphate, pH 6.5 in the absence of reducing agent (red curve), in the presence of 1 mM DTT (green curve) and in presence of 1 mM TCEP (grey curve). b) Time evolution of the ThT fluorescence intensity at 480 nm for Jos samples incubated at 37 °C for ~160 hours in 20 mM Hepes, pH 7.5, 150 mM NaCl in the absence of reducing agent (red curve), in the presence of 1 mM DTT (green curve) and in presence 1 mM TCEP (grey curve). Jos aggregation follows a nucleated growth polymerization mechanism when in the presence of DTT, with a shorter lag phase in Hepes buffer as previously reported by Figuereido *et al.*[298] Heterogeneous nucleation is instead observed both when in the presence of TCEP and when in the absence of reducing agents. All the conditions were tested in triplicates. Each curve represents the mean of replicates, and standard deviation represents the error bar.

The differences in aggregation profile revealed by the ThT-assay, resulted well correlated with differences in the morphology of the aggregates (Fig. 3.7). Indeed, investigation of the kinetics endpoints by Atomic Force Microscopy (AFM) revealed a homogeneous distribution of curvilinear wormlike fibrils ~60-100 nm length for samples incubated in presence of DTT (Fig. 3.7c). In contrast, fibrils were not observed in samples incubated with TCEP, which instead resulted characterized by the presence of monomeric protein and dispersed amorphous aggregates (Fig. 3.7d). An intermediate behaviour was again observed for samples incubated with no reducing agent, in which both fibrils and some shorter, less mature aggregates, were observed (Fig. 3.7b).

The fibrils morphology observed in the presence of DTT is consistent with previous studies. For instance, Ellisdon *et al.*[152] report curvilinear oligomers ~100 nm length, detected by Transmission Electron Microscopy (TEM), for a variant of Atx3 devoid of polyQ trait incubated at 37°C in 100 mM Tris pH 7.4, 80 mM NaCl, 10% glycerol, 2 mM PMSF, 5mM EDTA, and 15 mM βME. Similar species are reported by Figuereido *et al.*[298] and Silva *et al.*[303] for non-expanded Atx3 incubated at 37

°C in 20 mM Hepes pH 7.5, 150 mM NaCl, 1mM DTT, and in 20 mM NaPh pH 7.5, 150 mM NaCl, 1mM DTT, respectively.

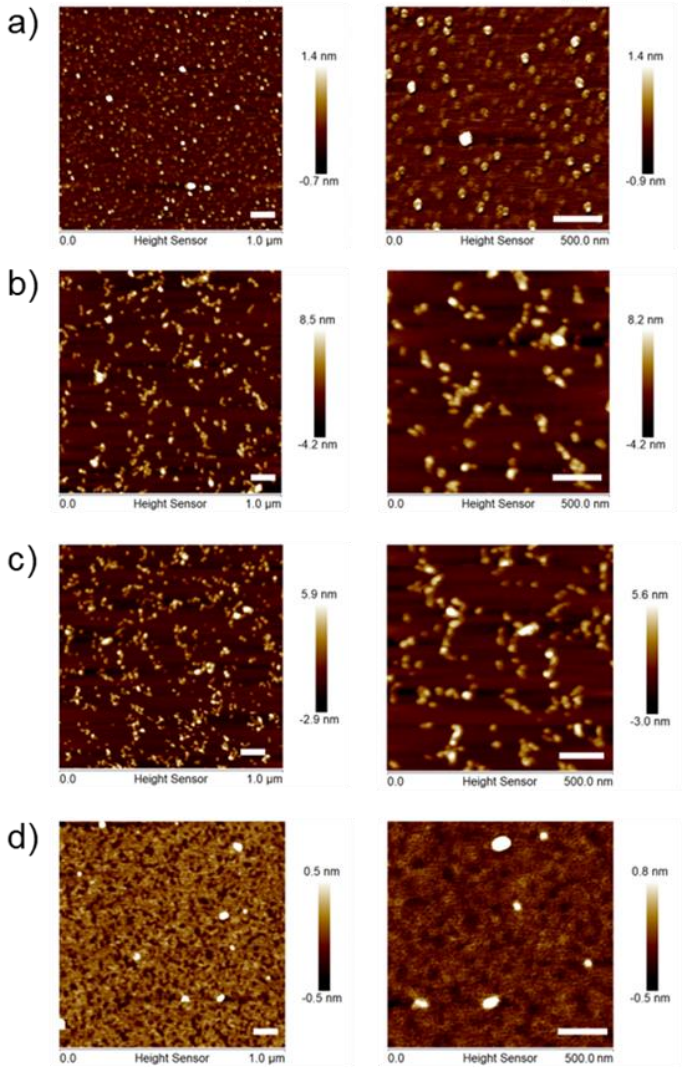


Figure 3.7 Reducing agents affect Jos aggregates morphology. The morphology of Jos aggregates, obtained by incubating the protein at 37 °C in different redox environments, was investigated by AFM. Samples were diluted and deposited on mica previously treated with $MgCl_2$, to favour negative-charged aggregates adhesion on the substrate. Samples were then rinsed with ultrapure water and gently dried flushing gaseous N_2. Images were acquired in the air in tapping mode by a Bruker FastScan AFM. AFM height image 1 µm x 1 µm (left), and AFM height image 0.5 µm x 0.5 µm (right) are reported for a) monomeric Jos, and aggregation endpoints of Jos incubated at 37 °C b) in the absence of reducing agent; c) in presence of 1 mM DTT; and d) in presence of 1 mM TCEP. Scale bar: 100 nm.
Monomeric Jos consists of globular particles of 3 nm height (a), while Jos samples after incubation with no reducing agent result heterogeneous, with both wormlike fibrils (60-70 nm length) and some shorter, less mature aggregates (b). Samples which followed incubation in the presence of 1 mM DTT, present instead a homogeneous distribution of tiny curvilinear wormlike fibrils, 3 nm height and 60-100 nm length (c). A different scenario is observed for samples incubated in the presence of TCEP, which consist of monomeric Jos and amorphous aggregates (d).

As shown, the use of DTT or TCEP, despite both being commonly used reducing agents, proved to have a strong impact on the aggregation of Atx3, changing the kinetics profile and morphology of the final aggregates. Therefore, these data strengthen the importance of careful selection of each buffer component to perform a robust *in vitro* aggregation assay, since they may ultimately influence protein stability and aggregation propensity.

3.1.4 Atx3 fibrils formation is favoured by nuclei covalently linked through disulfide bridges

The different effects of DTT and TCEP observed on Atx3 aggregation, directed the investigation towards the role of disulfide bridges in the self-assembly process of the protein. Indeed, both compounds are notoriously used as reducers of disulfide bridges formed among thiol side chains of cysteine residues in proteins. However, they are chemically very different as well as their redox mechanism (Fig. 3.8). DTT and TCEP can therefore exhibit different redox efficacy depending on the features of the protein under study and on the experimental conditions of pH and temperature.[304] Moreover, DTT has a very short half-life in aqueous solution (e.g. around 10 h at pH 7.5 and around 1.5 h at pH 8.5),[305] therefore it must be considered if it undergoes modifications or degradation during prolonged experiments such as the ThT-assays performed at 37 °C in this PhD work.

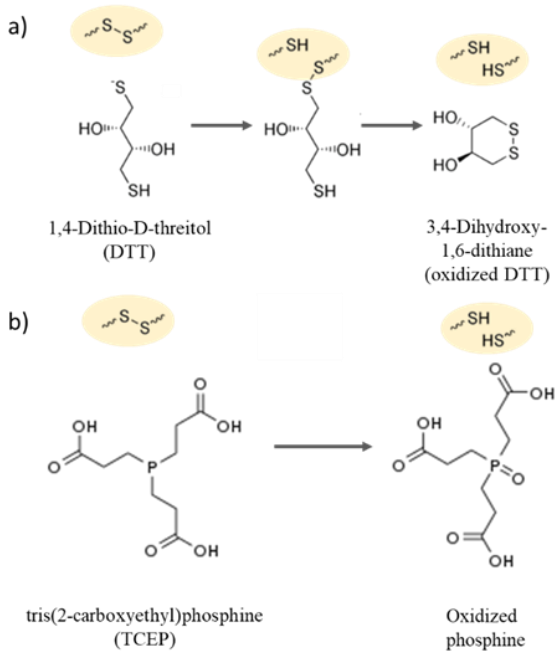


Figure 3.8 Structure and redox mechanism of DTT and TCEP. a) DTT, also known as Cleland's reagent, is a reducing agent based on thiols, which allows easy and specific reduction of intra- and inter-molecular disulfide bonds in proteins. The reaction proceeds by two sequential thiol-disulfide exchanges achieved by means of a SN_2 nucleophilic substitution. It therefore requires deprotonation of the thiol group of DTT (-SH → $-S^-$). The conversion of the thiol group in thiolate is pH-dependent, thus pH is a crucial variable when performing cleavage of disulfide bonds mediated by DTT and other thiols. Once the thiolate is formed, the nucleophilic attack occurs on the most electrophilic sulphur atom involved in the protein disulfide bond, forming an intermediate mixed-disulfide species. The reaction then proceeds thanks to the high reactivity of the second thiol of DTT, which leads to the formation of a stable six-membered ring with an internal disulfide bond (3,4-dihydroxy-1,6-dithiane, fully oxidized DTT).[306–308] In very rare cases, a DTT adduct may be formed (i.e. the two sulphur atoms of DTT may form disulfide bonds to different sulphur atoms in the protein). In such cases, DTT cannot cyclize since it has no remaining free thiols. b) TCEP is a reducing agent based on phosphines. The nucleophilicity of the phosphine in the reduction process is crucial since, as in the case of thiol-based reducing agents, the reaction involves a SN2 mechanism. The reaction proceeds with an initial nucleophilic attack of the phosphorus on one of the sulphurs of the S–S bond, with the formation of thio-phosphonium carbocation and the release of the first sulfhydryl. The phosphonium ion then undergoes hydrolysis to afford the phosphine oxide and the release of the second sulfhydryl. This second step is irreversible because of the stability of the oxidized phosphine.[309–311] TCEP is commercially available as hydrochloride (TCEP-HCl), which has multiple advantages over other disulfide-reducing agents. It is crystalline, solid, non-volatile, odourless, and water-soluble, and it is not susceptible to air oxidation. Moreover, it has a greater reducing capacity than DTT, the reduction reaction is faster, and it can be run in a wide pH range. However, TCEP-HCl is quite expensive for routine use in protein preparation, making DTT still the most extensively used disulfide reducing agent.

Firstly, the compound's stability at 37 °C was assessed by monitoring changes overtime in their 1D ^{1}H NMR spectra (Fig. 3.9). Samples were tested at a concentration of 1mM, which corresponds to the concentration of the reducing agent used in the ThT assays, and were subjected to continuous incubation at 37 °C. NMR spectra were recorded in both Hepes pH 7.5 and NaPh pH 6.5, using an 800 MHz Bruker spectrometer. Analysis of the NMR spectra acquired for TCEP revealed no changes in the structure of the compound, according to its well-known stability (Fig. 3.9a). As expected, equivalent analysis for DTT revealed, instead, a progressive disappearance of the NMR peaks related to the reduced form of the compound, and a concomitant appearance of the NMR peaks characteristic of the 6-membered ring of oxidized DTT (Fig. 3.9b). The irreversible oxidation was observed already after 6 hours of incubation at 37 °C and reached completion in between 120-138 h.

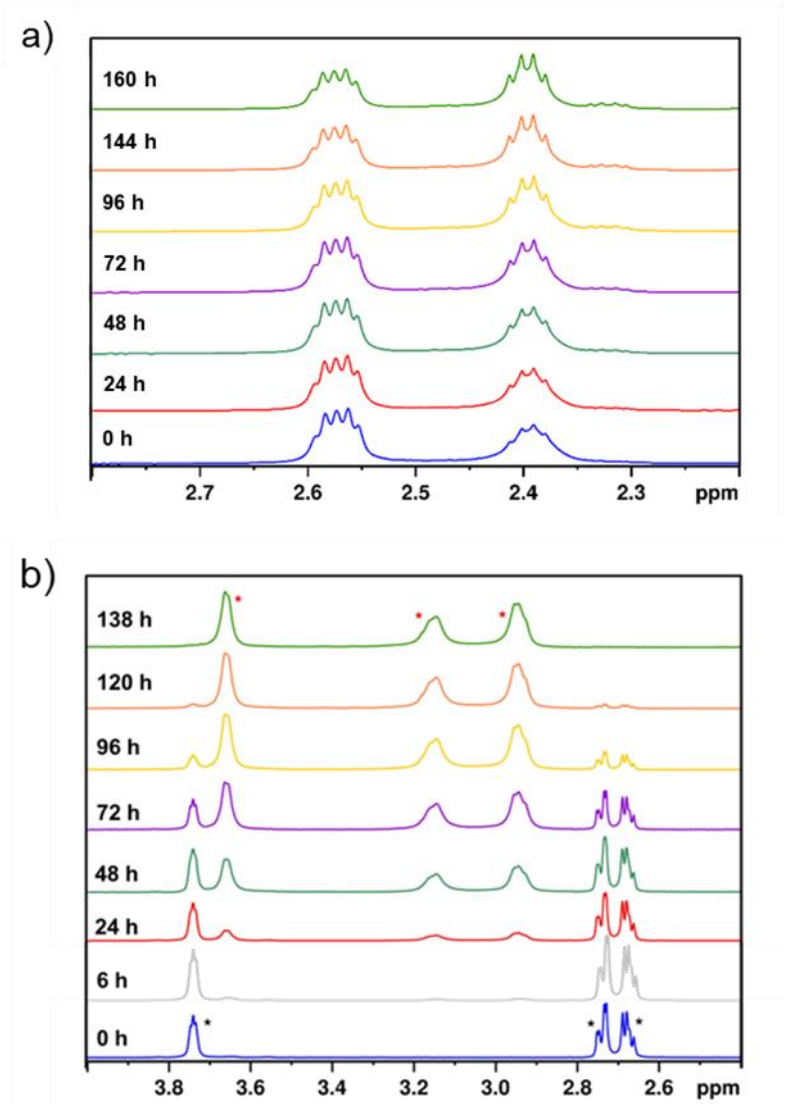


Figure 3.9 Stability at 37°C of TCEP and DTT assessed by monitoring changes overtime in the 1D ^{1}H NMR spectra of the two compounds. a) 1D ^{1}H NMR spectra acquired overtime on 1 mM TCEP solution during continuous incubation of sample at 37°C. The compound is stable up to 160 h of incubation at 37 °C. b) 1D ^{1}H NMR spectra acquired overtime on 1 mM DTT solution during continuous incubation of sample at 37°C. Irreversible oxidation is observed already after 6 hours of incubation at 37 °C and reaches completion in between 120-138 h. Black asterisks indicate proton NMR peaks of reduced DTT (3.74 ppm, 2.76 – 2.63 ppm), while red asterisks indicate proton NMR peaks of oxidized DTT (3.66 ppm, 3.15 ppm, 2.95 ppm).

The time window during which stability of reduced DTT at 37 °C was guaranteed, resulted much shorter than the 160 h of ThT kinetics assays employed to investigate Atx3 aggregation. This result dictated to investigation of whether oxidized DTT could have a role in the formation of Atx3 fibrils at 37 °C. To this purpose, aggregation kinetics was monitored by ThT-assay for both Jos and Atx3-Q14 samples incubated at 37 °C both in the presence of freshly prepared DTT and in the presence of fully oxidized DTT (Fig. 3.10a). The lag phase of the aggregation process was nearly abolished when in the presence of oxidized DTT, while persisting for long periods when incubating the proteins with freshly prepared DTT, varying linearly with the concentration of the compound (Fig. 3.10b).
Despite this difference, AFM experiments performed on the aggregation endpoints revealed that Atx3 aggregates maintain the same morphology of curvilinear wormlike fibrils (3 nm height, 60-100 nm length) both when adding in the aggregation buffer freshly prepared DTT and when adding oxidized DTT (Fig. 3.10c), suggesting that the two tested samples follow the same aggregation pathway in the two conditions.

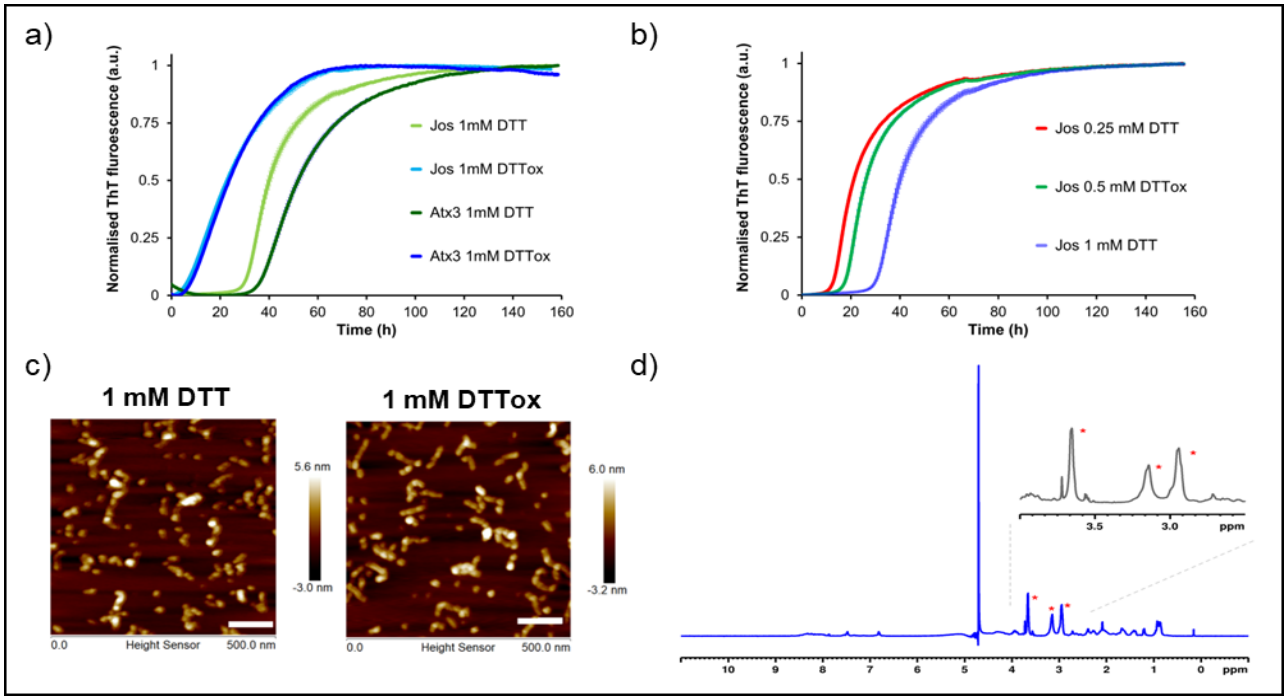


Figure 3.10 Effect of oxidized DTT on Atx3 aggregation. a) ThT assays of Jos and Atx3-Q14 in 20 mM Sodium Phosphate, pH 6.5 in the presence of 1 mM of freshly prepared DTT (green curves), and 1 mM of oxidized DTT (blue curves). Each curve represents the mean of replicates normalized to the maximum value reached. b) ThT assays of Jos in the presence of 0. 25 mM (red curve), 0.5 mM (green curve), and 1 mM DTT (blue curve). Each curve represents the mean of replicates normalized to the maximum value reached. c) Morphological analysis by AFM of the aggregation endpoints after 160 h incubation at 37°C in the presence of 1 mM DTT (on the left) and 1 mM DTTox (on the right). AFM height image 0.5 µm x 0.5 µm. Scale bar: 100 nm. d) 1D ^{1}H NMR spectrum of the Jos sample incubated in the presence of 1 mM DTT after incubation at 37°C. Red asterisks indicate the proton NMR peaks of oxidized DTT.

Altogether, these results strongly suggest that oxidized DTT somehow speeds up Atx3 aggregation. On the other hand, when incubating the sample in the presence of freshly prepared DTT, the latter gradually gets oxidized during the aggregation lag phase, and only when in its oxidized form is able

to affect the aggregation process. Consistently, 1D ^{1}H NMR spectrum of the kinetics endpoint of Atx3-Q14 incubated with freshly prepared DTT, revealed the complete disappearance of reduced DTT in favour of its fully oxidized form (Fig. 3.10d).

DTT loses then its reducing reactivity during Atx3 aggregation kinetics experiments, with the consequence that the formation of both intra- and inter-molecular Cys-Cys disulfide bridges may occur. Indeed, the only Jos domain contains four cysteine residues, including the catalytic Cys-14, which may be involved in the formation of disulfide bonds in a suboptimal reducing environment. This may lead to the formation of covalently linked small oligomers during the lag phase of the aggregation process, which then act as nuclei for the formation of mature fibrils.

This hypothesis was verified by testing Jos samples by non-reducing SDS-PAGE after 160 h incubation at 37 °C in i) the absence of reducing agents, and in the presence of ii) freshly prepared DTT, iii) oxidized DTT, and iv) TCEP (Fig. 3.11). For the two conditions with DTT, it resulted in a marked presence of high MW species trapped in the stacking region of the polyacrylamide gel, which correlates with supramolecular assemblies linked by disulfide bridges. In the case of TCEP, Jos samples run as monomers consistently with the stability of the reducing agent which prevents from cysteines oxidation. When in the absence of reducing agents, both supramolecular assemblies linked by disulfide bridges and monomeric protein are visible as bands of the SDS-PAGE. Moreover, unpublished data kindly shared by Dr Sandra Macedo-Ribeiro - recognized international expert on Atx3 aggregation – reveal that Atx3 aggregation strongly decreases in any reducing condition when all cysteines in the protein sequence are substituted by serines (data not shown).

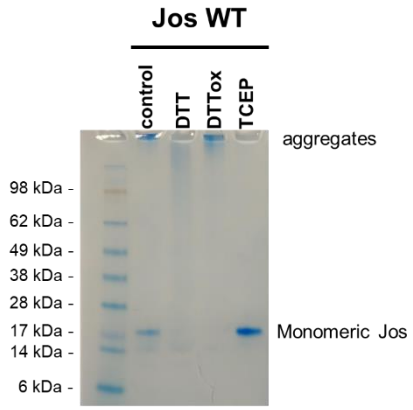


Figure 3.11 SDS-PAGE analysis of Jos aggregation endpoints. Jos samples were tested by non-reducing SDS-PAGE after 160 h incubation at 37 °C in the absence of reducing agents (control), in the presence of freshly prepared DTT (DTT), in presence of oxidized DTT (DTTox), and in the presence of TCEP (TCEP). For the two conditions with DTT, it resulted in a marked presence of high MW species trapped in the stacking region of the polyacrylamide gel, which correlates with supramolecular assemblies linked by disulfide bridges.

In conclusion, all evidence collected here strongly supports that Jos aggregation is mediated by the formation of inter-molecular disulfide bridges. All four cysteine residues within Jos sequence can

potentially form intermolecular bridges capable of crosslinking monomers, thereby promoting the formation of aggregation nuclei which then mediate the formation of mature fibrils.

This mechanism might be directly implicated in the formation of pathological aggregates in SCA3, as similarly reported for infectious from prion protein,[312] tau pathologies,[313] and superoxide dismutase SOD1 aggregates associated with amyotrophic lateral sclerosis.[314,315] The same mechanism has been reported to be relevant also to the formation of inclusion bodies by the neuropathological protein TDP-43, which is also associated with amyotrophic lateral sclerosis and frontotemporal lobar degeneration. The thiol-disulfide exchange mediated by oxidation in the RRM2 domain of TDP-43 favours the formation of disordered RRM2 dimers, suggesting that disulfide bridges shift the equilibrium to species that serve as nuclei of aggregation.[316] A similar behaviour may take place for Atx3 aggregation. However, while RRM2 of TDP-43 contains only 2 cysteines, the solely Jos domain contains four cysteine residues, all potentially able to form intermolecular disulfide bonds: C14, belonging to the catalytic triad, C18, C114 and C172 (Fig. 3.12a).

To discern if a specific couple of cysteines is involved in Jos aggregation, here four point mutants, where each cysteine is replaced by an alanine, were produced: Jos-C14A, -C18A, -C114A and -C172A. All the mutants maintain the overall Jos folding according to circular dichroism (CD) analysis (Fig. 3.12b).

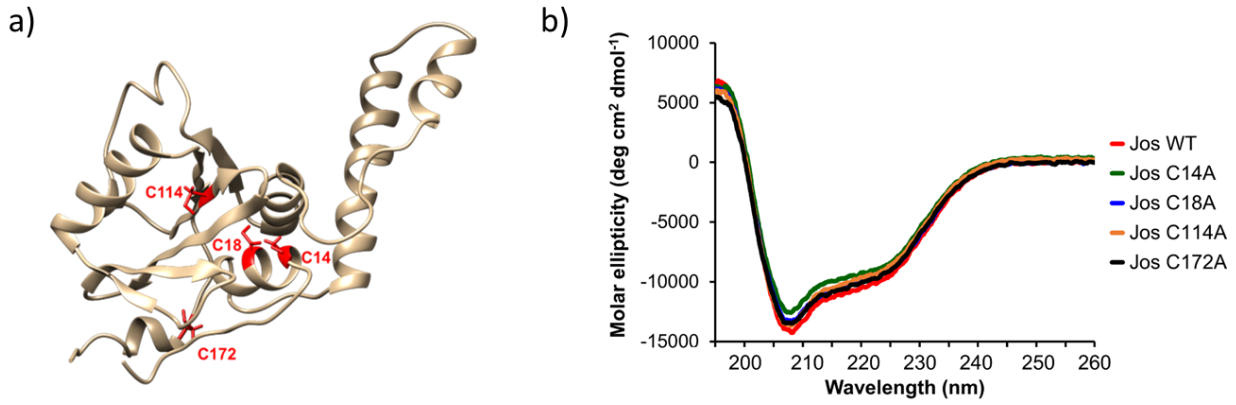


Figure 3.12 Point mutation of cysteine residues in Jos domain. To discern if a specific couple of cysteines is involved in Jos aggregation, four point mutants of Jos were produced. a) Jos NMR structure (PDB: 1yzb) with C14, C18, C114 and C172 highlighted in red. b) Circular dichroism analysis of the new Jos mutants. CD spectra well overlap for all the four mutants revealing no significative changes in secondary structure with respect to the wild type protein.

Aggregation kinetics at 37 °C of the different Jos mutants were then tested by ThT assay. The investigation was carried out in i) 20 mM Hepes pH 7.5, 150 mM NaCl, and in the same buffer supplemented with ii) freshly prepared DTT, and iii) TCEP (Fig. 3.13). All four mutants resulted to follow a similar aggregation pathway to Jos wild type, independently of the mutated cysteine (Fig. 3.13a). Consistently, non-reducing SDS-PAGE performed on the kinetics endpoints revealed the same trend already observed for Jos wild type (Fig. 3.13b).

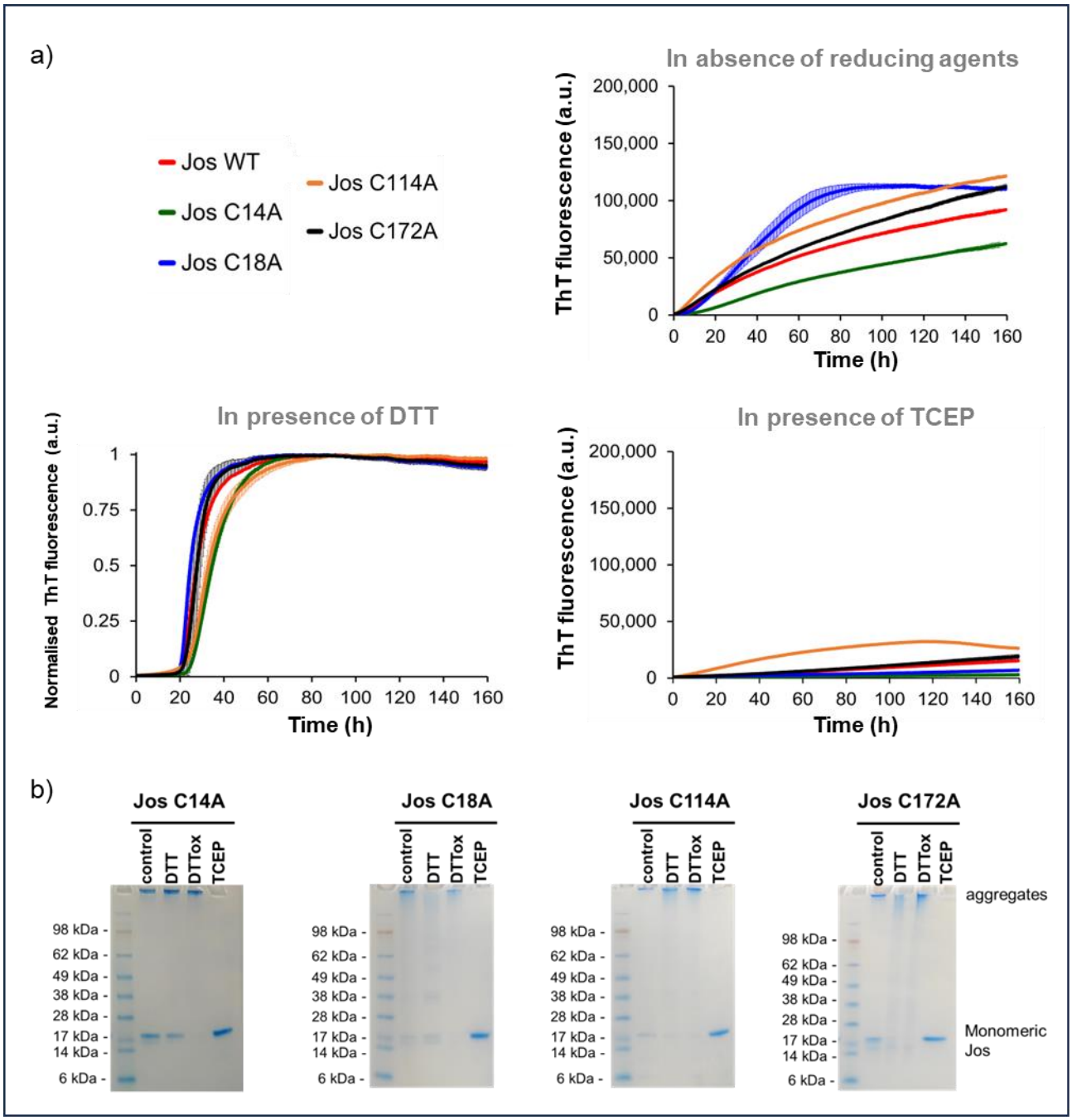


Figure 3.13 Aggregation of Jos mutants. a) Time evolution of the ThT fluorescence intensity at 480 nm for Jos point mutants incubated at 37 °C for ~160 hours in i) 20 mM Hepes pH 7.5, 150 mM NaCl, and in the same buffer supplemented with ii) freshly prepared DTT and iii) TCEP. All the conditions were tested in triplicate and mean values are reported. Error bars are equivalent to the standard deviation of replicates. b) Non reducing SDS-PAGE of Jos mutants final aggregates in i) 20 mM Hepes pH 7.5, 150 mM NaCl (control), and in the same buffer supplemented with ii) 1mM DTT, iii) 1mM DTTox and iv) 1mM TCEP.

Based on the investigation of cysteine point mutants, the Atx3 aggregation process does not seem to involve a main culprit cysteines pair. Additional investigation on cysteine double mutants, maybe useful in discerning if specific cysteine residues in Jos sequence are more significantly involved in the protein aggregation.

3.1.5 Oxidized DTT weakly binds both monomeric and oligomeric Jos

Evidence collected in this project supports that Jos aggregation *in vitro* is mediated by the formation of inter-molecular disulfide bridges. This is prevented when in the presence of TCEP in the testing buffer, being TCEP a very stable molecule which well performs its reducing activity all over the long incubation period required for Atx3 aggregation assay at 37 °C. On the other hand, the formation of small oligomers covalently linked by disulfide bonds occurs when the testing buffer is supplemented with DTT, being DTT very prone to oxidation which impairs its reducing activity.

Here, the binding affinity of monomeric Jos for DTT, TCEP or oxidized DTT was also investigated by NMR Chemical Shift Perturbation (NMR-CSP) and Saturation Transfer Difference NMR (STD-NMR) experiments.

To perform NMR-CSP analysis, the ^{15}N-labeled protein was titrated with a progressively higher amount of unlabelled compound and chemical shift perturbations at each titration point were monitored by the acquisition of two-dimensional (2D) Heteronuclear Single Quantum Correlation (HSQC) spectrum. The analysis resulted in no CSP for any Jos amino acid residue, indicating that the amide backbone of the protein is not interested in binding events with any tested compound (Fig. 3.14).

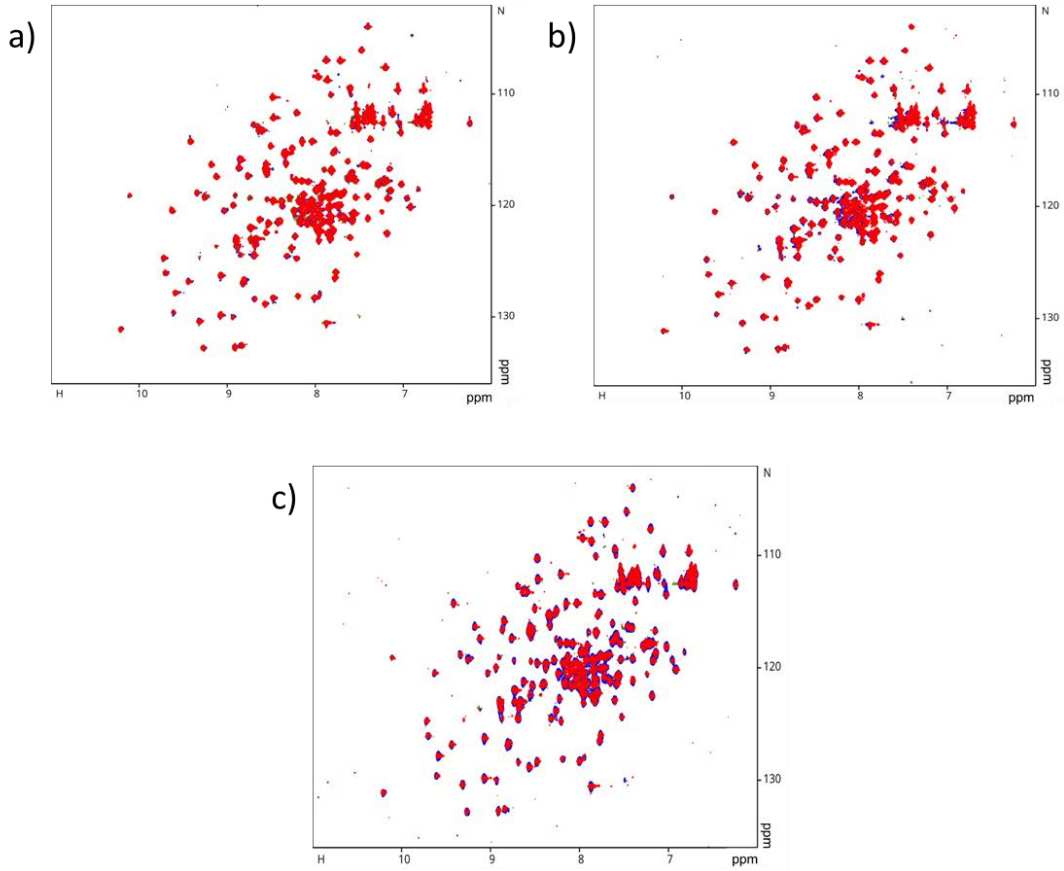


Figure 3.14 ^{15}N-HSQC spectra for CSP analysis. Overlap of the HSQC spectra acquired on ^{15}N-Jos in the absence of the compound (in blue), and in the presence of the compound in a molar ratio protein: compound of 1: 1.5 (in red). No significant CSP (>0.1ppm) were found, suggesting that none of the compounds interacts with the amide backbone of the protein. a) Jos/DTT; b) Jos/oxidized DTT; c) Jos/TCEP.

The binding affinity of monomeric Jos for DTT, TCEP or oxidized DTT was then investigated by STD-NMR experiments at 37 °C. This is because STD-NMR experiments, in contrast to CSP-NMR, are designed to reveal binding events not necessarily interesting to the amide backbone of the protein. Results from STD-NMR confirmed that both DTT and TCEP do not bind to monomeric Jos, while in the case of oxidized DTT a low but non-null STD signal is present (Fig. 3.15). This suggests that oxidized DTT not only allow the formation of inter-molecular disulfide bonds because of its missing reducing activity, but might also trigger aggregates formation by weak and likely non-specific interaction with Jos side chains.

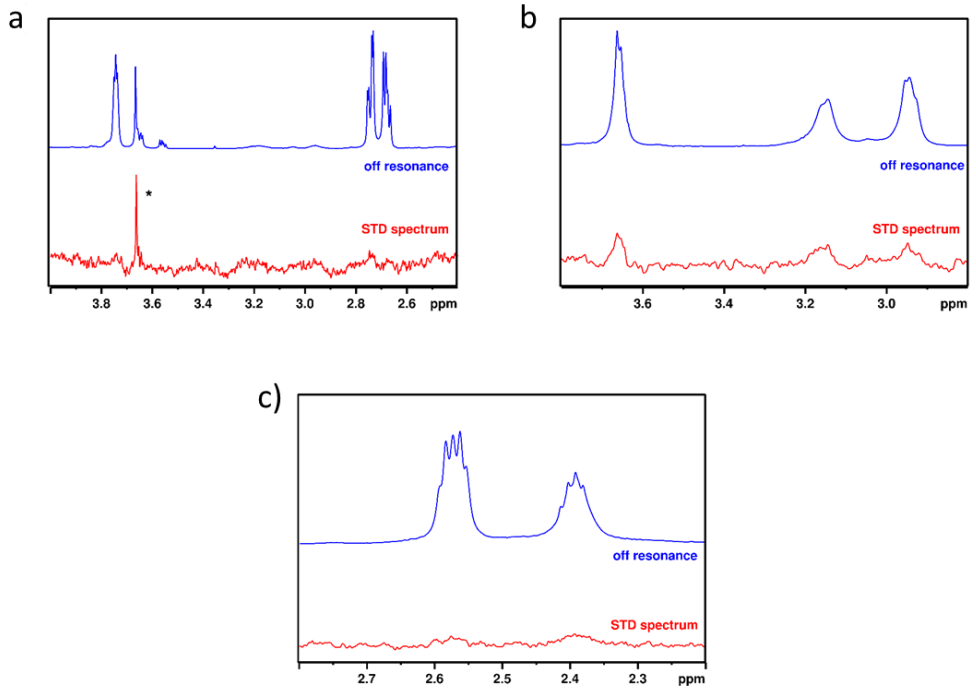


Figure 3.15 STD-NMR experiments to assess the binding affinity of monomeric Jos for DTT, TCEP or oxidized DTT. STD-NMR experiment is carried out by recording a control spectrum of the compound, called *off resonance* spectrum. A second spectrum, called *on resonance* spectrum, is then recorded, in which the protein is saturated by selective irradiation. The subtraction of on resonance spectrum from that off resonance yields the so-called STD spectrum, consisting of peaks of the compound protons, whose intensity is therefore the difference in the intensities among off and on resonance spectra. A finite value for this difference reflects the spatial proximity of the compound with protein, i.e. bound ligand, while this value vanishes for unbound compounds. Here off resonance (in blue) and STD spectra (in red) in the presence of Jos are reported for a) DTT; b) oxidized DTT; and c) TCEP. The asterisk in a) indicates an unidentified buffer contaminant present in the DTT sample. The presence of peaks with residual intensity in the STD spectrum for oxidized DTT indicates a binding event among the compound and the protein.

STD-NMR experiments were also acquired overtime such to assess variation in STD signals intensity in function of time, and to test whether oxidized DTT is able to bind also Jos oligomeric species that progressively appear during incubation at 37°C. A progressive decrease in the protein signals was observed overtime, consistent with the gradual aggregation of Jos, along with an increase in the STD signals up to 48 h (Fig. 3.16). This suggests that oxidized DTT might have more affinity for Jos oligomers than for the monomeric protein. However, further investigations are necessary to establish if the increased intensity of the STD signals is indicative of enhanced affinity, rather than an effect

of spin diffusion due to the increased size of the protein, and consequential improved magnetization transfer.

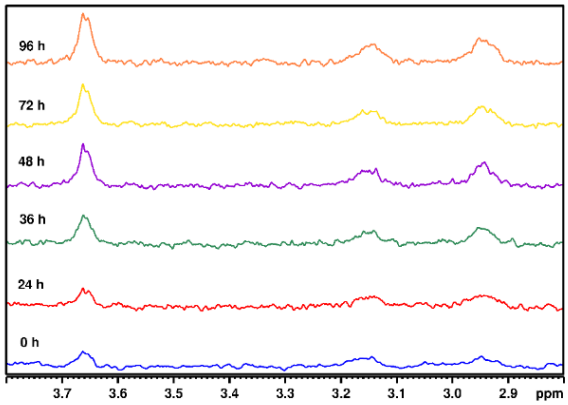


Figure 3.16 STD-NMR spectra of oxidized DTT in the presence of Jos. STD-NMR experiments were acquired in the function of time at 37 °C to monitor changes in the intensity of the STD signals in the presence of progressively larger Jos aggregates.

3.2 Characterization of Jos in complex with polyUb chains

In the present work, it was investigated the potential protective role of polyUb, one of the known Atx3 partners, against Atx3 aberrant aggregation. Indeed, evidence shows that the presence of the natural interactors could represent a protective constraint for protein structure, suggesting that physiological interaction with natural partners and aggregation tendency are strictly connected and could represent two competitive processes.[8–10] Moreover, it has been demonstrated that some Atx3 residues involved in processing polyUb chains, belong to the same hydrophobic patch involved in aggregation, suggesting that interaction among Atx3 and polyUb could stabilize the native state of Atx3, thereby reducing its aberrant supramolecular assembly.[231]

An example of partner-orphan mediated aggregation is represented by annexin A11 (ANXA11), an exponent of the family of Ca^{2+} binding annexins, involved as TDP43, SOD and FUS in ALS.[317] Under normal physiological conditions, ANXA11 interacts with calcyclin through a calcium-regulated mechanism. In the presence of calcium, the N-terminus, which is an intrinsically disordered region rich in Gly, Pro and Tyr, become available to interact with calcyclin forming a helical structure. On the contrary, in the absence of calcium, the N-terminal domain maintains its low-complexity structure, folds back and interacts with the C-terminal, inhibiting the binding with the natural partner. Some ALS-related mutations in ANXA11 interesting key aa in the N-terminal region (i.e. G38R, D40G), impair the proper binding with calcyclin. This results in exposing regions that are normally protected, and in increasing protein propensity to form pathological aggregates.[317]

Another example comes from TDP-43, another ALS-related protein, for which pathological aggregation also appears to be inhibited by natural interactors.[318,319] TDP-43 works as RNA binding protein and it usually localizes in the nucleus, but in pathological conditions, it changes subcellular

localization forming cytoplasmic inclusions.[320–323] Selected sequence-specific RNA aptamers with high binding affinity significantly affect the aggregation rate, suggesting that tight and specific RNA-TDP-43 interactions could work as chaperones. Otherwise, pathological mutations that affect this protective interaction leave the protein exposed to an aberrant self-assembly mechanism.

Partner-orphan mediated aggregation has been observed also for tau, a protein correlated with AD and physiologically involved in microtubules stabilization and transport along axons. Tau hyperphosphorylation induces dissociation of tau from tubulin, causing the loss of the microtubule-stabilizing function, possibly contributing to neuronal degeneration.[324,325] When dissociated from microtubules, tau can acquire various conformations, bringing together the N- and C-terminal portions of the protein,[326] and exposing the hexapeptide motifs of microtubule-binding domain, which have a propensity to form β-sheets[327]. Furthermore, several tau missense mutations that impair microtubules interactions lead to neurodegeneration.[328]

These findings strengthen the idea that the physiological function of a protein can actively compete with the formation of abnormal aggregates. Therefore, partner-orphan model of aggregation could be very useful for drug design aimed at targeting the region of the aggregation-prone protein involved in the interaction with its natural partners. Accordingly, structure characterization of Jos in complex with Lys48-linked di- or tri-Ub (K48-diUb and K48-triUb, respectively) was attempted to get molecular details of the interaction interface. This is mainly because this information offers valuable clues for designing synthetic peptides targeting Jos binding sites for polyUb that could serve as effective drugs for treating SCA3 pathology.

3.2.1 K48-triUb decreases Jos aggregation at 50 °C

Here, the effect of K48-triUb on Jos aggregation was investigated at 50°C. The Jos point mutant C14A was used as testing protein to impair the proteolytic activity of the protein towards the K48-triUb substrate. Aggregation kinetics experiments were performed in large volume (3ml) at 50 °C by ThT-assay for both monomeric Jos and Jos in complex with K48-triUb. Interestingly, results showed that K48-triUb markedly slows down the aggregation kinetics, supporting the hypothesis of interference of polyUb chains on Jos aggregation pathway (Fig. 3.17).

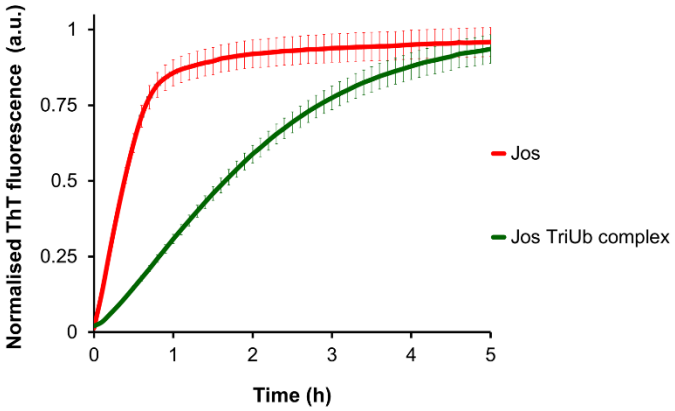


Figure 3.17 Aggregation kinetics at 50 °C of apo Jos and Jos in complex with K48-triUb. Time evolution of the ThT fluorescence intensity at 480 nm for Jos samples incubated at 50 °C for ~5 hours in 20 mM Sodium Phosphate pH 6.5 supplemented with 1 mM TCEP. Red curve refers to Jos in its apo form, while green curve refers to Jos in complex with K48-triUb. ThT intensity is reported after normalisation to the maximum value reached. Error bars correspond to 5% of each measurement.

Based on these encouraging preliminary data, polyUb impact on Jos aggregation was also investigated at physiological temperature. In particular, it was compared the aggregation kinetics of Jos in its apo form with the aggregation kinetics of both Jos/K48-diUb and Jos/K48-triUb complexes (Suppl. Fig. S6). Aggregation kinetics experiments were performed by ThT assays in small volume (200µl) in 20 mM Hepes pH 7.5, 150 mM NaCl and 1 mM DTT, according to Figueiredo *et al.*[298] and the insights on Jos aggregation obtained in the first part of the project (section 3.1).

In contrast with results obtained at 50 °C, results at 37 °C showed that complexes follow similar aggregation kinetics of isolated Jos, consisting in a nucleation-mediated polymerization (Fig. 3.18).

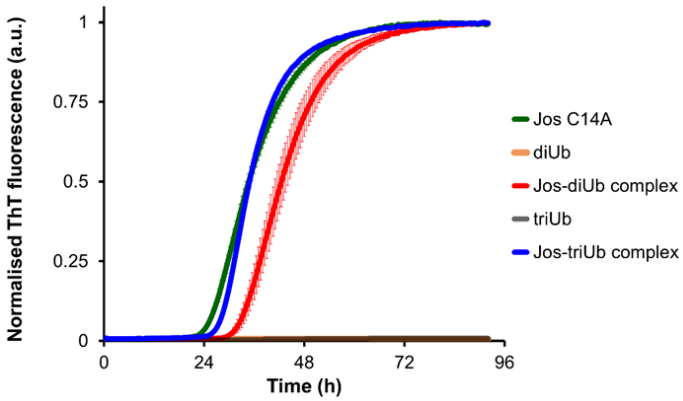


Figure 3.18 Aggregation kinetics at 37°C of Jos-polyUb complexes. Time evolution of the ThT fluorescence intensity at 480 nm for Jos samples incubated at 37 °C for ~96 hours in 20 mM Hepes pH 7.5, 150 mM NaCl, 1 mM DTT. Green curve refers to Jos in its apo form, while blue and red curves refer to Jos/K48-triUb and Jos/K48-diUb, respectively. ThT-assay was also performed for isolated K48-triUb and K48-diUb, orange and grey curves, respectively. Each curve represents the mean of triplicates. Fluorescence values are normalized at maximum value. Error bars correspond to the standard deviation of replicates.

To deepen the meaning of these results, the ThT-assay endpoints were analysed by AFM. Morphological analysis confirmed that both isolated K48-diUb and K48-triUb do not form aggregates after incubation at 37°C. Only spherical species with a size ranging in the interval 0.5 – 2 nm height were detected, compatible with the expected size of monomeric forms (Fig. 3.19a). On the other hand, in the case of complexes samples, it was possible to detect fibrils comparable in size to the ones formed by isolated Jos, and a large amount of small species on the background (Fig. 3.19b-c).

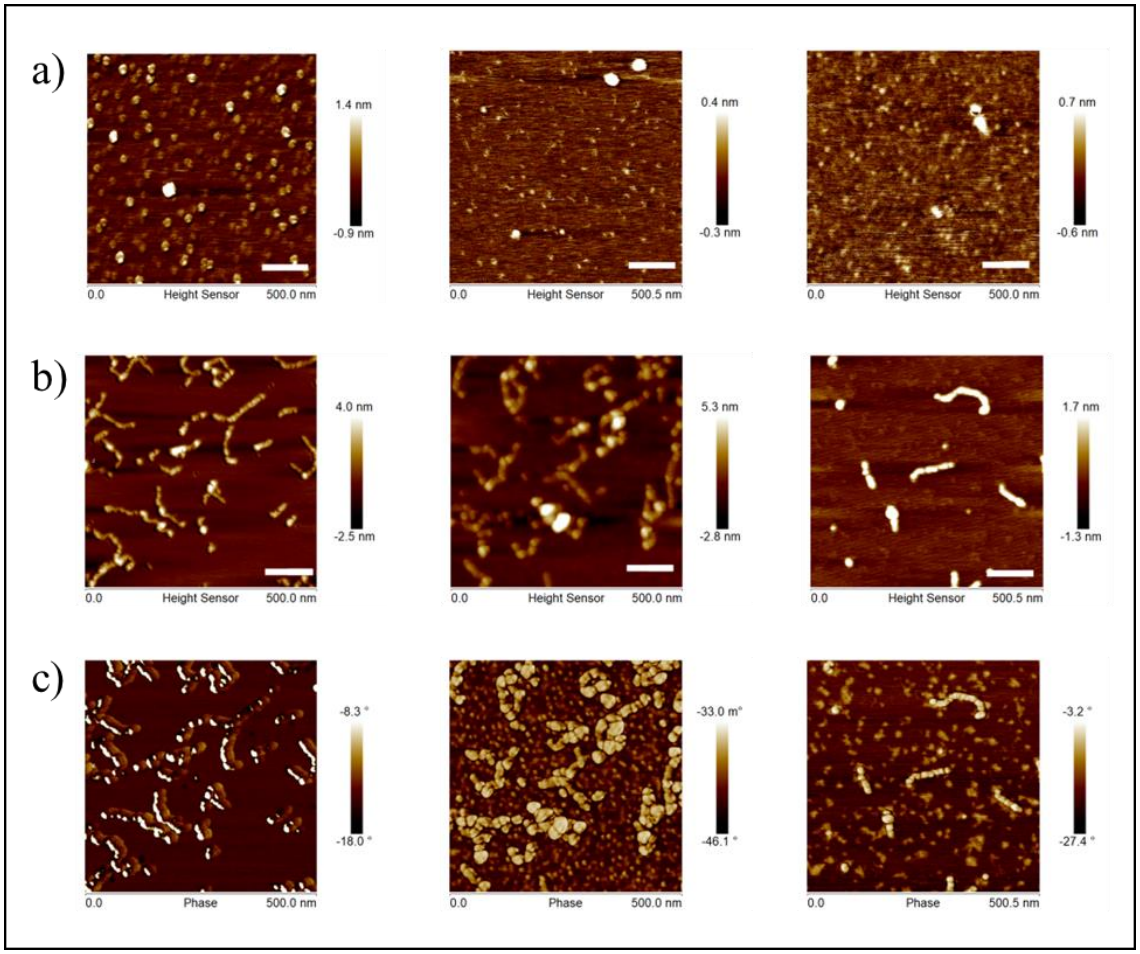


Figure 3.19 AFM images of ThT-assay endpoints of Jos/K48-diUb and Jos/K48-triUb complexes. a) AFM height images 0.5µm x 0.5µm for freshly purified Jos (left), and ThT-assay endpoints of isolated K48-diUb (middle), and isolated K48-triUb (right). b) AFM height images 0.5µm x 0.5µm for ThT-assay endpoints of Jos (left), Jos/K48-diUb complex (middle), and Jos/K48-triUb complex (right). c) AFM phase images 0.5µm x 0.5µm for ThT-assay endpoints of Jos (left), Jos/K48-diUb complex (middle), and Jos/K48-triUb complex (right). Scale bar: 100 nm. Morphological analysis confirmed that both isolated K48-diUb and K48-triUb do not form aggregates after incubation at 37°C. On the other hand, in the case of complexes samples fibrils comparable in size to the ones formed by isolated Jos are detectable. Complexes samples also contain a considerable number of small species on the background, which very likely correspond to monomeric polyUb chains. These can be particularly appreciated in the AFM phase images in c).

Interestingly, the small species observed on the background in both Jos/K48-triUb and Jos/K48-diUb samples had an estimated size close to the one of isolated K48-triUb and K48-diUb, respectively, suggesting a possible dissociation of these species from the complex. Analytical SEC was thus performed on ThT-assay endpoints of Jos/K48-triUb and Jos/K48-diUb samples to assess if complex dissociation occurred (Fig. 3.20a). Analysis of the SEC fractions by SDS-PAGE confirmed that ThT-

assay endpoints of Jos/K48-triUb and Jos/K48-diUb contain SDS-soluble aggregates formed by only Jos, while K48-Ub chains eluate as free species (Fig. 3.20b).

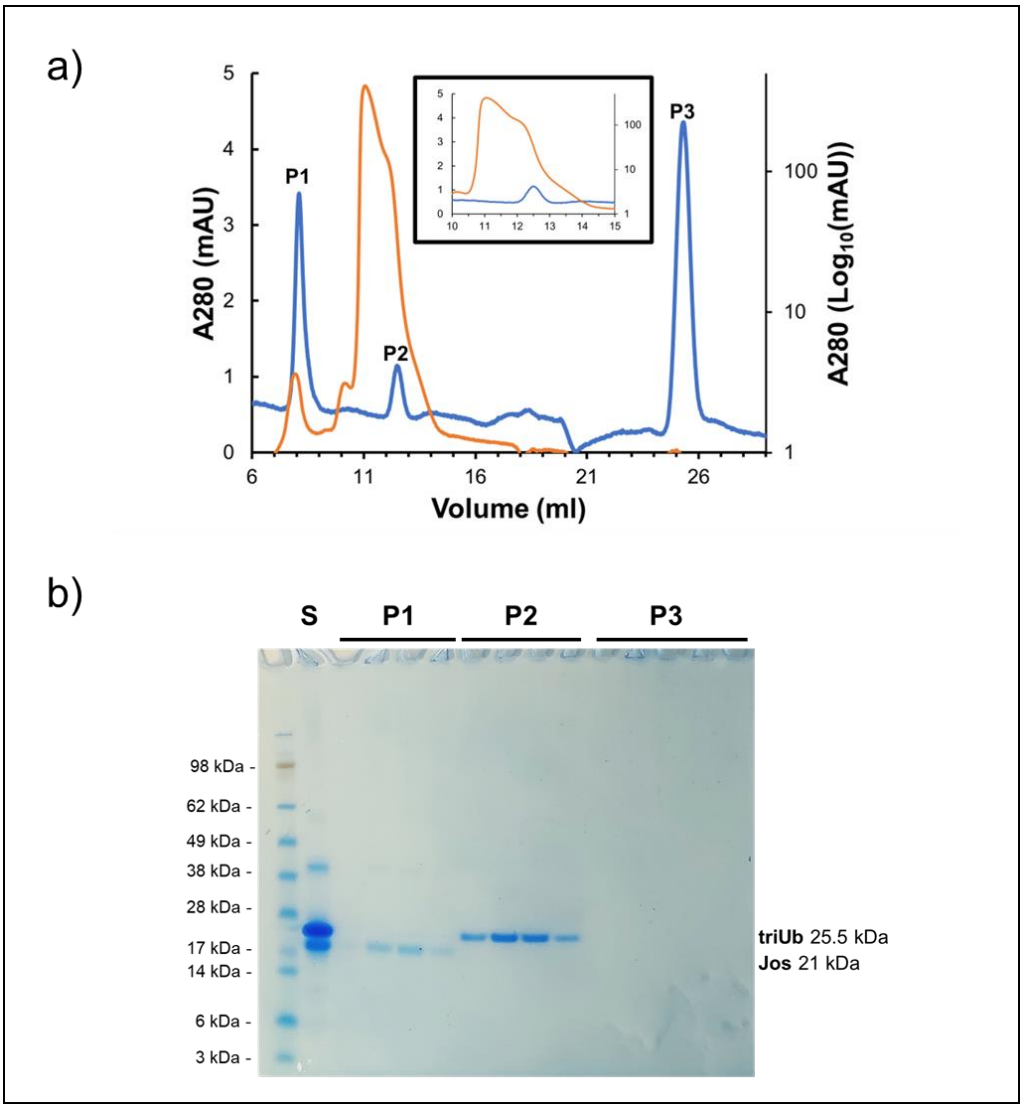


Figure 3.20 SEC analysis of ThT-assay endpoint of Jos/K48-triUb complex. a) SEC profiles of ThT-assay endpoint of Jos/K48-triUb complex (in blue), and freshly prepared complex (in orange). Retention volumes for Jos/K48-triUb complex and free K48-triUb correspond to 11ml and 12.5, respectively. b) SDS-PAGE analysis of SEC fractions of Jos/K48-triUb ThT-assay endpoint. S corresponds to the pre-SEC sample; P1-P3 are related to the indicated peaks in the SEC profile. Results show that P1 contains isolated Jos, P2 contains free K48-triUb and P3 presumably corresponds to the DTT added to the sample to perform the ThT-assay.

Altogether, results show that at 37 °C and a concentration of 14 μM, Jos/K48-triUb dissociates in favour of Jos aggregation. Indeed, the Jos binding site for K48-triUb involves the hydrophobic patch of Jos identified as aggregation prone[231] that, when exposed, allows interaction among monomers with consequent formation of oligomeric species. According to the new insights obtained in this work (section 3.1), these oligomers are stabilized by disulfide bonds thus impairing the re-association of the complex with K48-triUb in favour of Jos self-assembly. Further analysis should be carried out using a higher concentration of complex or cross-linking triUb and Jos, thus creating a more stable complex.

3.2.2 Structural characterization of Jos/K48-diUb and Jos/K48-triUb

Structural characterization of Jos/K48-diUb and Jos/K48-triUb complexes was firstly performed by Small Angle X ray Scattering (SAXS), being this a quick method to characterize protein samples through the extraction of several overall parameters related to the size, oligomeric state, and overall shape of the molecule.[330,331]

SAXS experiments were performed on 5 mg/ml complex sample in TRIS buffer supplemented with 1% sucrose to reduce radiation damage. Given the high concentration of sample, the buffer was also supplemented with 2 mM TCEP to avoid disulfide bridges formation and to consequently reduce Jos aggregation propensity. Indeed, sample mono-dispersity is strictly required to avoid misinterpretation of SAXS data, since the presence of aggregates in solution could lead to a wrong estimation of SAXS-derived parameters, such as radius of gyration (R_g). In addition, samples were subjected to size-exclusion chromatography coupled SAXS (SEC-SAXS) to remove potential aggregated species and maximize sample purity (Fig. 3.21a).[329–332]

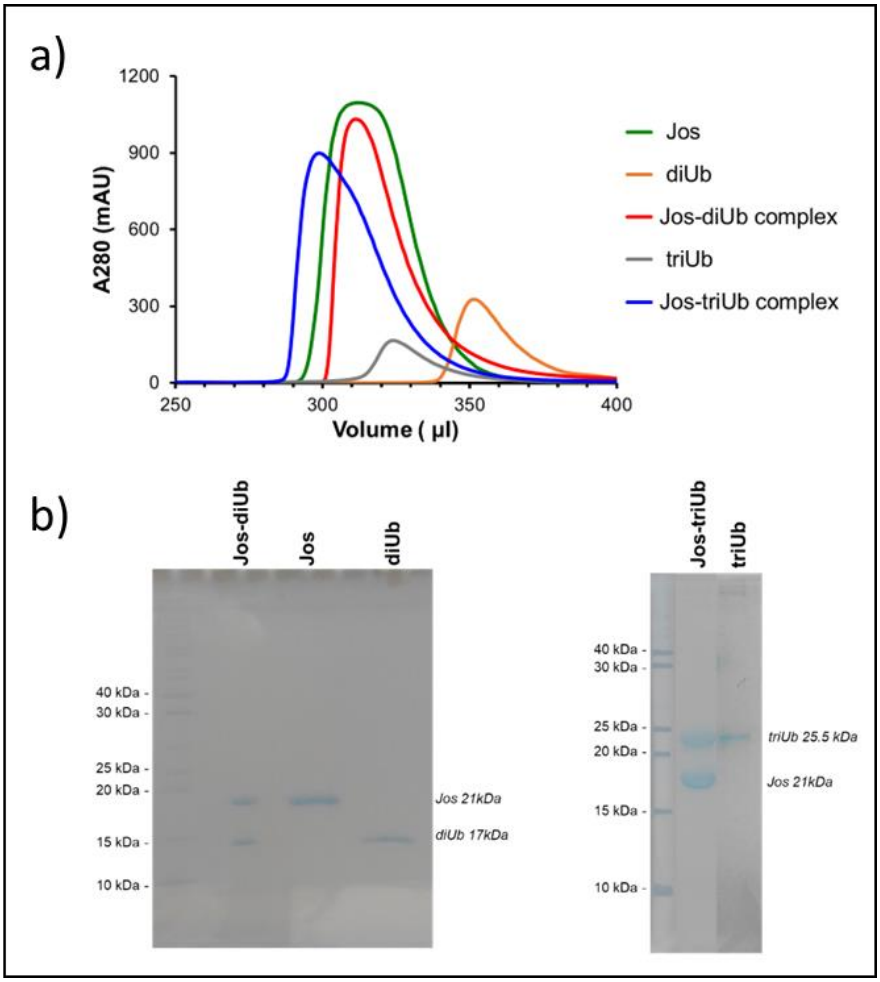


Figure 3.21 Preparation of SAXS samples. Samples for SAXS were subjected to size-exclusion chromatography coupled SAXS (SEC-SAXS) to remove potential aggregated species and maximize sample purity. a) SEC profiles for Jos C14A (in green), K48-diUb (in orange), K48-triUb (in grey), Jos/K48-diUb complex (in red), Jos/K48-triUb complex (in blue); b) SDS-PAGE analysis of samples before loading on SEC-SAXS.

SAXS data were collected for both Jos/K48-diUb and Jos/K48-triUb complexes, and isolated Jos, diUb and triUb as constituents of the complexes. The R_g and I(0) parameters were then derived by Guinier analysis of the SAXS radially averaged one-dimensional curves, and used for the scaling of both the Krakty and P(r) plots (Fig. 3.22).

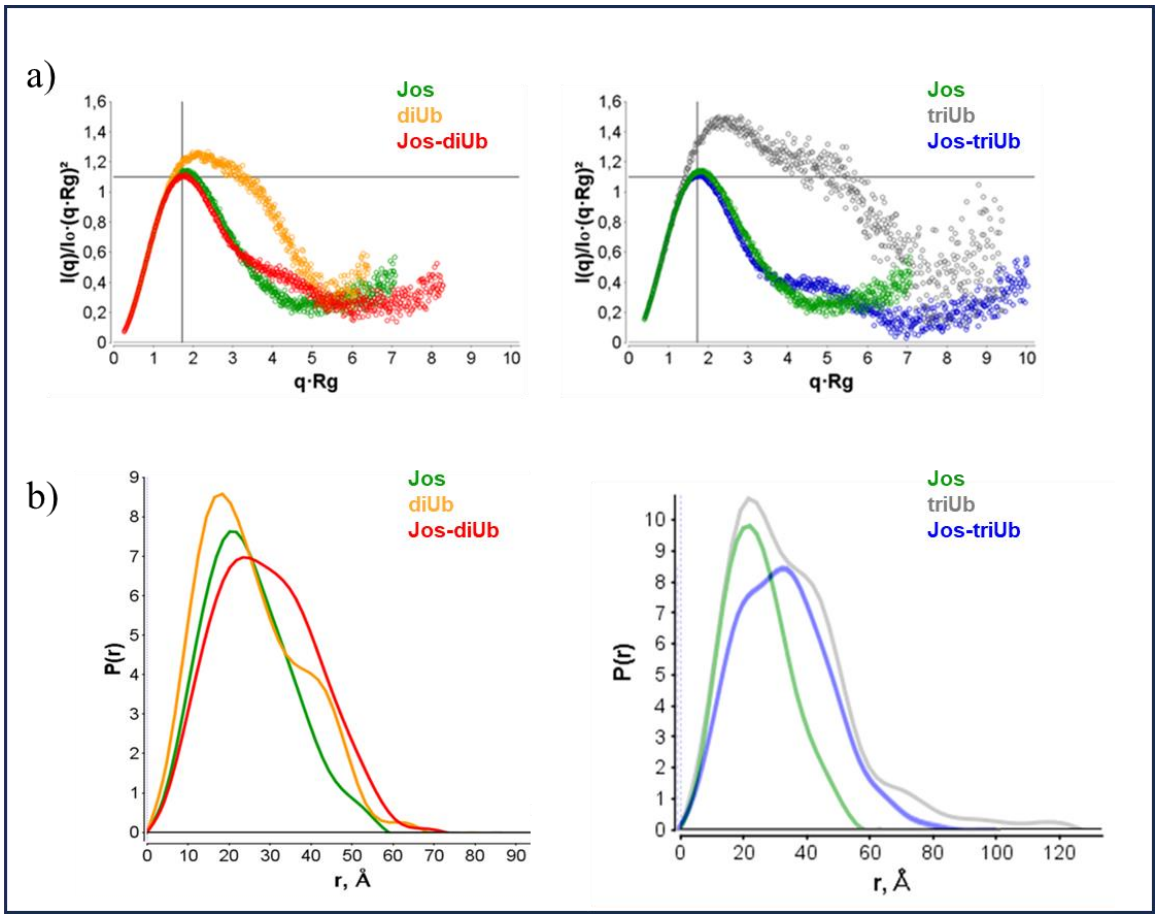


Figure 3.22 Kratky and P(r) plots derived from SAXS curves. Jos C14A (in green), diUb (in orange), triUb (in grey), Jos-diUb complex (in red), Jos-triUb complex (in blue). a) Guinier-based normalized Kratky plot. In order to compare different proteins, the dimensionless representation of the Krakty plot is used, where the scattering intensity is scaled by $(q \times R)^2 / I(0)$ and the scattering vector q is multiplied by R_g.[333] In dimensionless Kratky plots, globular proteins always have a maximum at $q \times R_g = 3^{1/2}$ and $(q \times R_g)^2 I(q)/I(0) = 1.104$. These coordinates are marked by the cross. b) P(r) plot of the different samples.

In the case of Jos, as expected for a globular protein, the dimensionless Kratky plot appears as a bell-shaped peak. Broadening of the bell-shaped curve and a shift of the maxima to larger $q \times R_g$ values were instead observed in the case of isolated K48-diUb and K48-triUb, suggesting a higher level of flexibility for these two molecules (Fig 3.22a).[334] In the case of the complexes, the curves resemble a multi-modal distribution, typical of multi-domain proteins. The main component is superimposed with the bell-shaped peak of the isolated Jos, while the shoulder indicates the presence of a flexible domain, constituted by the Ub chain partner.

These results were confirmed by the analysis of P(r) plots, where it is reported the distance distribution function (Fig. 3.22b). A P(r) bell-shaped profile, typical of globular particles, was

obtained for Jos, while profiles with multiple shoulders and oscillations, corresponding to intra and inter-subunit distances, were obtained for all the other samples under analysis. In particular, the complexes resulted in more extended than isolated Jos (>Dmax), but more compact with respect to the corresponding free Ub chain.

The size parameters calculated for all the samples resulted to be consistent with the expected values derived from the monomeric species and their estimated molecular masses (Table 3.2).

	MW (kDa)	R_g (Å)	Dmax (Å)	Volume (Å^3)
Monomeric Jos	21	19.37	63	33,359
diUb	17	20.42	68	23,763
triUb	25.5	29.28	127.5	90,356
Jos-diUb complex	38	22.65	77.5	56,877
Jos-triUb complex	46.5	25.56	80	84,818

Table 3.2 Size parameters calculated for both Jos/K48-diUb and Jos/K48-triUb complexes. R_g, Dmax and Volume values estimated by SAXS profiles using ScÅtter IV software. Samples are ordered in ascending order by expected MW. R_g is the root-mean-square distance of the macromolecular mass from its centre of mass. It is a measure of the overall size or spatial extent of the macromolecule, and it is useful for understanding the conformation, flexibility, and compactness of the macromolecular structure in solution. Dmax represents the largest distance between any two points of the particle, describing the size or maximum extent of a macromolecule or particle in solution.[333]

Altogether these results show that both K48-diUb and K48-triUb chains are stably attached to Jos, yielding complexes that are more tightly packed than the corresponding polyUb chains.

Based on the observation that Atx3 seems to work better with longer polyUb chains,[5,279] suggesting that the functional complex requires longer substrate chains, SAXS-derived structure modelling was performed only for Jos/K48-triUb complex. Specifically, both Jos ubiquitin binding site 1 (I77, L91, W120, F163) and site 2 (Y27, F28, W87), and Ub hydrophobic patch L8, I44, V70, present on each Ub subunit, were considered as potential interaction surfaces.

Structure modelling was attempted by a combination of CORAL approach, and DENsity from Solution Scattering (DENSS, *https://tdgrant.com/*), an algorithm for calculating 3D particle electron densities directly from 1D solution scattering data. In particular, CORAL is a method for rigid body modelling of multidomain protein complexes against multiple scattering curves, where also symmetry and contacts can be taken into account.

Unfortunately, in the case of K48-triUb, the only available experimental structure is reported for cyclic triUb (PDB: 7CAP),[335] wherein the potential interaction surfaces are not solvent-exposed and do not align with the experimental SAXS profile (Fig. 3.23a). Therefore, as entry substructure for complex modelling, it was used a linear K48-triUb obtained by CORAL using three subunits of Ub (PDB 1UBQ)[31] and K48-triUb SAXS profile (Fig. 3.23b).

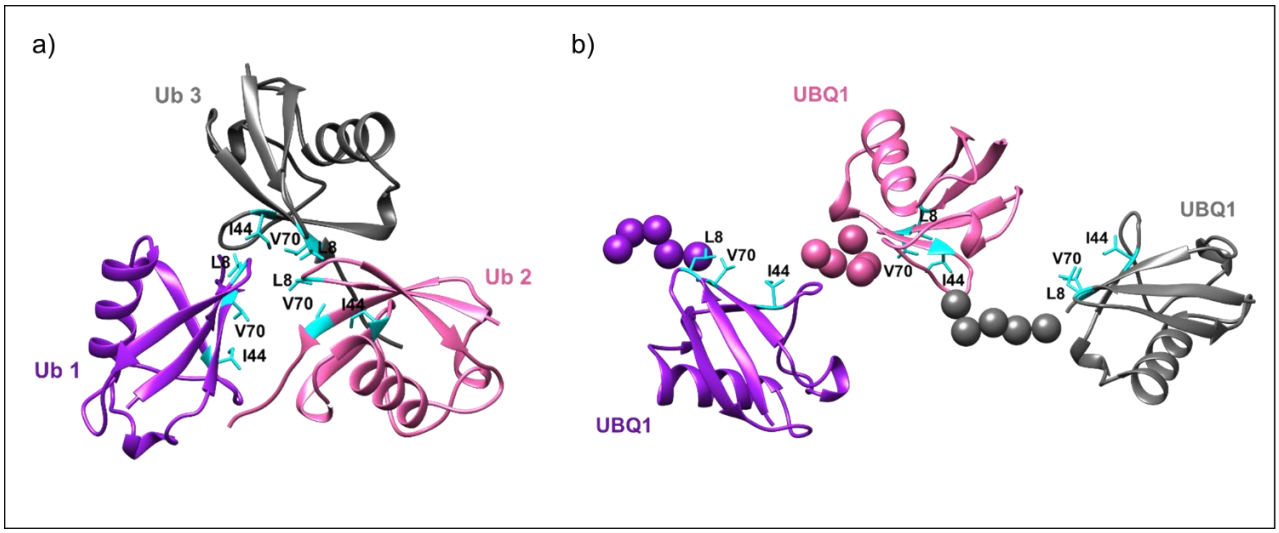


Figure 3.23 Modelling of linear K48-triUb. a) Crystallographic structure of K48-triUb in cyclic conformation (PDB: 7CAP). Each Ub subunit is represented in a different colour (Ub1, purple; Ub 2, pink; Ub 3, grey). The hydrophobic patch L8-I44-V70 is in cyan and is not exposed to solvent. b) model for linear K48-triUb obtained by CORAL using three Ub subunits (PDB: 1UBQ). Final five residues of each subunit were treated as dummy residues (spheres) to increase the flexibility of the system. In cyan, L8-I44-V70 patch, that is more solvent-exposed than in cyclic K48-triUb.

The structural model obtained for linear K48-triUb was then used as entry structure for complex modelling together with Jos NMR structure (PDB 1yzb). Unrestrained approach reached convergence after 57 iterations (χ^2: 1.115), but multiple reconstructions (i.e. 4 CORAL runs with the same parameters) showed that unfortunately, this approach is not sufficient to solve the ambiguity of the system (Fig. 3.24a). Indeed, four statistically equivalent models were obtained, all correlating around 80% with the SAXS-derived electron density shape, each displaying K48-triUb in a different position (Fig. 3.24b).

Despite the diversity, all models indicate both Jos catalytic site and ubiquitin binding site 1 as preferred interaction surface (Fig. 3.24a).

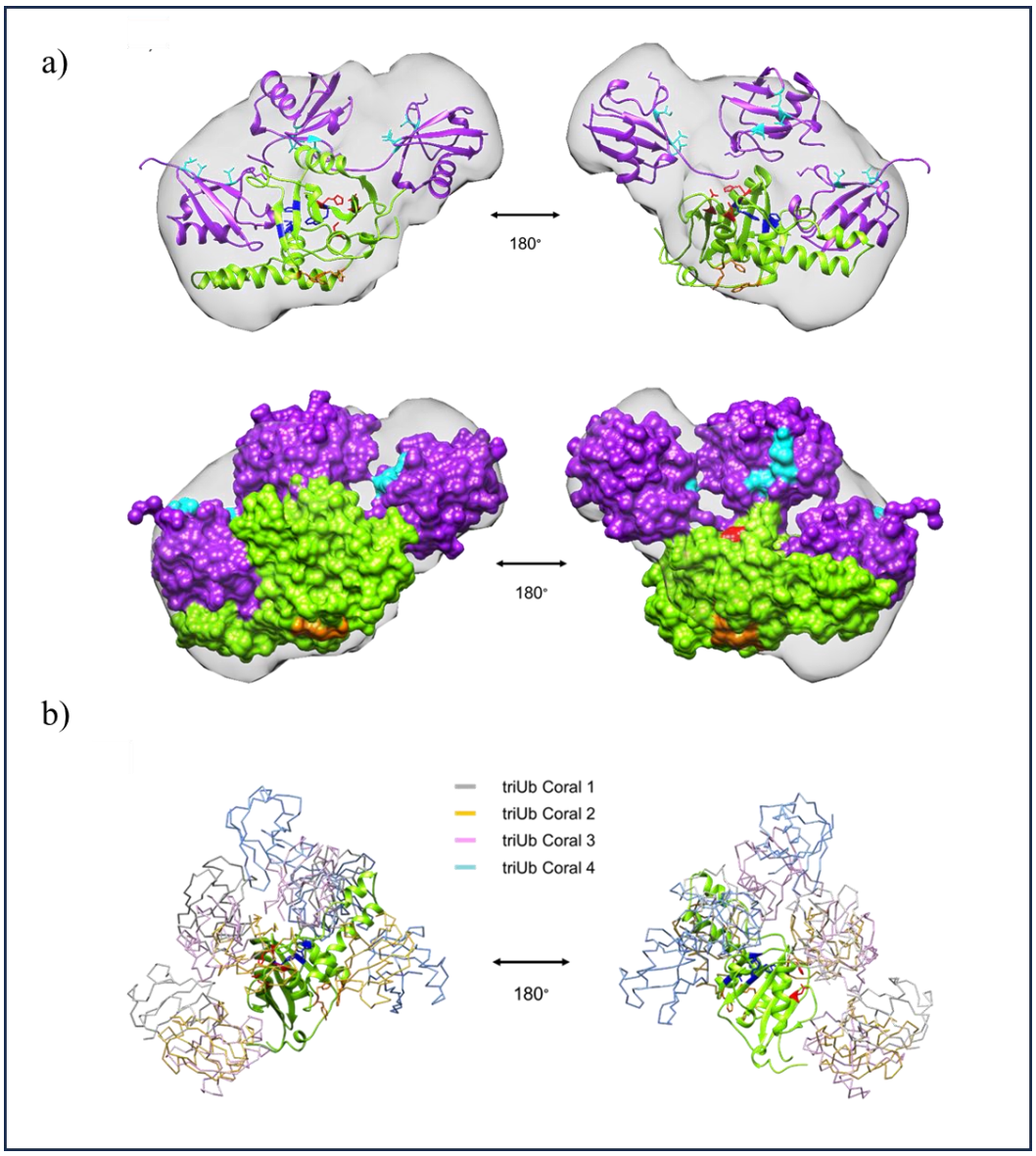


Figure 3.24 Structure modelling of Jos/K48-triUb complex. Exemplification CORAL run to obtain a rigid-body model of Jos/K48-triUb complex. a) in the first row, ribbons representation of one of the CORAL of Jos-triUb complex obtained with unrestrained modelling. Jos (PDB: 1yzb) is represented in green; triUb is represented in purple. Jos: catalytic site (C14-H119-N134), in red; Ub binding site 1 (I77, L91, W120, F163), in blue, Ub binding site 2 (Y27, F28, W87), in orange; triUb: L8-I44-V70, in cyan. In the second row, molecular surfaces representation, respecting the previous colour code. Dark grey shell is the electron density function extracted from Jos-triUb complex SAXS profile with DENSS algorithm. Shell volume is adjusted in Chimera 1.16, according to SAXS experimental volume: 84818 $Å^3$. b) ribbon representation of all the four Jos-triUb CORAL obtained. Jos of all the models is aligned and represented in green. In blue, pink, yellow and grey triUb of the different rigid body models show that triUb could be located in different positions with respect to Jos.

Although modelling from SAXS curves was not conclusive for determining the binding mode between K48-triUb and Jos, the results confirmed the formation of the complex and the sample quality, allowing to obtain access to the EMBL High-throughput crystallization and fragments screening (HTX) facility in Grenoble. A first wide crystallization screening was performed selecting both standard kits for primary screening (i.e. Classics-Suite, JCSG, PACT) and kits specialized for

complexes (i.e., Midas-plus, Morpheus, ProPlex). Screenings were performed at RT, being the Tm of the complex 55°C, as assessed by thermal shift assay.

Surprisingly, kits specialized for complex formation gave poorer results than conventional kits, which instead led to the formation of microcrystals (Fig. 3.25b). These will be added to refinement screens or to sparse matrix screens, to increase the probability of obtaining better crystallization hits.

Based on the results of this screening, therefore, it will be necessary to proceed with additional trials, using a higher sample concentration and focusing on the chemical conditions (buffer composition and precipitant percentages) that allowed the identification of promising traces of crystallization.

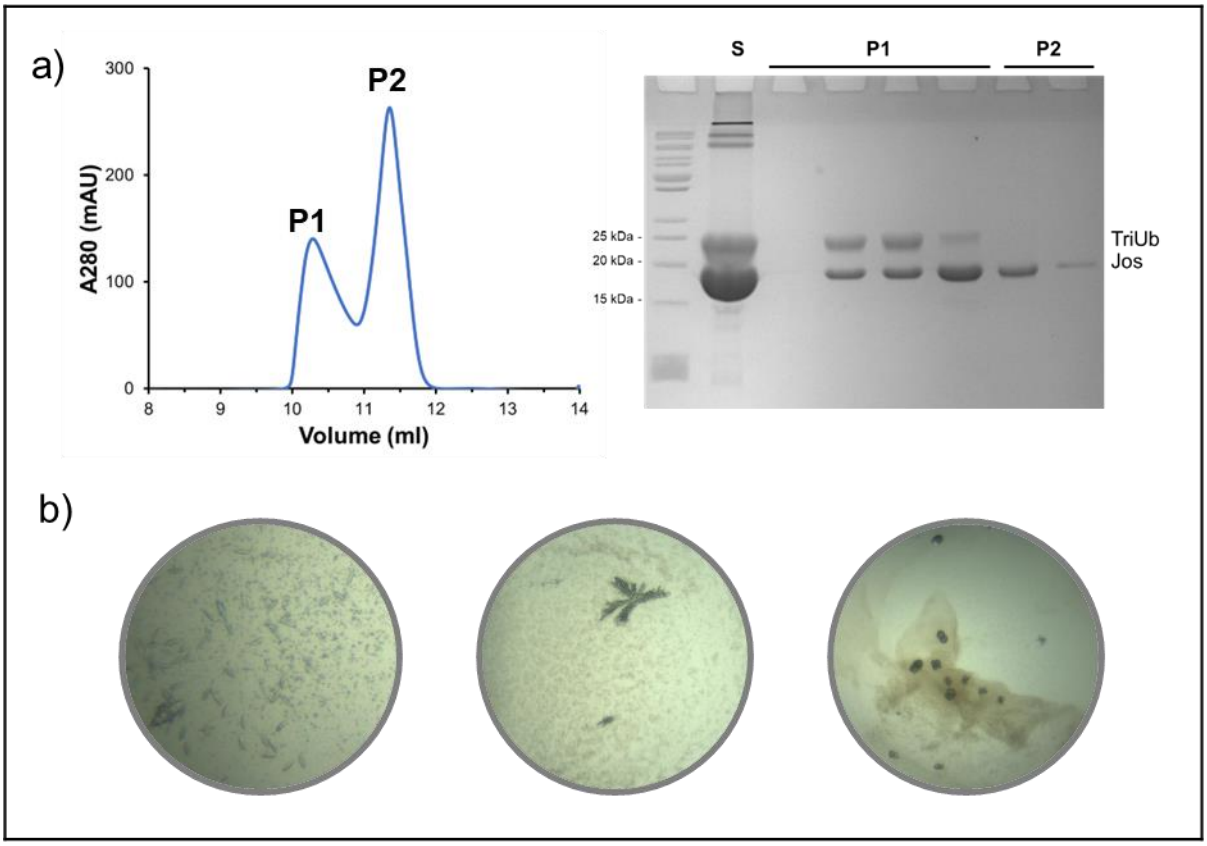


Figure 3.25 Jos-triUb crystallization trial. a) SEC profile and correspondent SDS-PAGE of Jos-triUb sample. Since preliminary screenings had indicated the dissociation of the complex in favour of polyUb crystal formation, in this case, the complex was prepared by incubating Jos: triUb in 2: 1 molar ratio, thus maintaining an excess of Jos. Subsequently, SEC with Increase 10/300 75s in 20 mM Sodium Phosphate, pH 6.5 and 1 mM TCEP was performed, and the fractions containing the pure complex (P1) were pooled down together and concentrated to produce a crystallization sample of 11.5 mg/ml in 100 µL. b) Some examples of crystal hints obtained by this first crystallization trials, after 20 days. On the left, sea urchin micro-crystals (size ~ 20 µm) in 0.01 M nickel chloride, 0.1 M Tris pH 8.5, 1 M lithium sulphate; in the middle, an irregular single crystal (size ~ 230 µm) in 0.2 M magnesium chloride, 0.1 M Hepes sodium salt pH 7.5, 30% isopropane; on the right micro-crystals (size ~ 20 µm) in 40% glycerol ethoxylate; phase separation is visible on the background.

4 Conclusions

This PhD thesis was aimed at investigating the molecular mechanisms of Ataxin-3 aggregation, to gain new insights into its complex and still controversial supramolecular assembly process. Moreover, the potential protective role of polyUb chains against Atx3 aggregation was evaluated.

Notably, Atx3 exhibits a strong inclination to form oligomeric structures, and even slight adjustments to *in vitro* experimental conditions have been reported to alter its amyloid assembly kinetics.[229,303] Tight control of the multiple variables and aggregation buffer components is thus critical to the development of reproducible assays to monitor protein aggregation and for further evaluation of potential anti-aggregating molecules.[298]

In this context, this project focused on the effect of temperature and buffer composition on different Atx3 aggregation steps, highlighting the role of reducing environment in the aggregation process. Aggregation kinetics experiments *in vitro* of non-expanded and expanded Atx3 variants revealed that at 37 °C Atx3 hydrolysis, presumably due to reported Atx3 autocatalytic activity,[297] competes with aggregation of Atx3 full length. The hydrolysis leads to a domain of similar size to Jos, which is able to form SDS-soluble aggregates. On the other hand, at 50°C the formation of oligomers outpaces hydrolysis events, hindering Atx3 degradation through fibrillation. Moreover, results revealed that Atx3 aggregation at 50 °C follows a polyQ-independent mechanism, thus representing a good strategy for vast screening of potential anti-aggregating molecules interfering solely with the Jos-dependent aggregation mechanism.

Intriguingly, Atx3 hydrolysis observed at 37 °C resulted to occur only when using TCEP as a reducing agent in the aggregation buffer, while resulting in inhibition when incubating the protein in the presence of DTT. This suggests that, although both TCEP and DTT are commonly used reducing agents for proteins, their well-known difference in stability[304] could expose Atx3 to a different reducing environment, influencing its autocatalytic activity, in particular in the case of long-duration experiments as aggregation kinetics at 37°C. Of note, from an extensive comparison of aggregation buffers composition used in previous studies on Atx3 aggregation, TCEP resulted in being used only in this PhD work and in previous work from Pastore's group performed at 50 °C.[149]

Morphological analysis of final aggregates through atomic force microscopy revealed that in the presence of DTT curvilinear short fibrils are formed. On the other hand, when in the presence of TCEP, Atx3 is less prone to aggregation and only a few amorphous aggregate species are formed.

Interestingly, stability tests performed on DTT through 1D ^{1}H NMR spectra revealed that the compound oxidizes within the lag phase of the Atx3 aggregation kinetics at 37 °C, forming 3,4-Dihydroxy-1,6-dithiane (here indicated as oxidized DTT, DTTox) which is devoid of reducing

activity. Accordingly, in aggregation kinetics experiments performed in the presence of DTTox, the lag phase was almost abolished, suggesting that Jos aggregation is triggered by oxidized DTT rather than its reduced form. This result led to investigating the involvement of intermolecular disulfide bridges in Jos supramolecular assembly. Indeed, cysteines could potentially form intermolecular bridges capable of crosslinking monomers, thereby promoting the formation of aggregation nuclei.[294-299] Contribution of disulfide bridges in Jos aggregation was assessed by analysis of the aggregation endpoints by SDS-PAGE in non-reducing conditions, which led to the conclusion that Jos self-assembly process is mediated by the formation of intermolecular disulfide bridges, that could stabilize aggregation intermediates, serving as nuclei for further aggregation steps, adding a new piece to the intricate Atx3 aggregation process.

Aggregation propensity was then investigated for four Jos point mutants, where each cysteine residue (C14, C18, C114, C172) was substituted by alanine, to assess if a specific cysteines pair could be involved in the formation of intermolecular disulfide bridges. This was not the case, as results revealed that all Jos cysteines can be equally involved in the formation of cross-linking disulfide bridges.

In this experimental work, it was also investigated the binding mode of Jos with K48-diUb and K48-triUb, and the effect of this interaction on Atx3 aggregation propensity. In particular, Jos C14A was used as testing protein, thus inactivating the protein catalytic site.[297] This was intended to inhibit the cleavage of Ub chains, allowing the study of the interaction among the protein and its natural partner. Due to the lack of available structures of Atx3-polyUb complexes in the Protein data bank (PDB), firstly it was verified the formation of stable complexes among Jos and polyUb chains by small angle X rays scattering, thus proceeding with preliminary structural characterization of Jos/K48-triUb complex by rigid body modelling. The results did not allow the univocal identification of interaction surface among K48-triUb and Jos, but allowed to obtain access to the EMBL High-throughput crystallisation and fragments screening (HTX) facility in Grenoble to setup crystallization trials.

In parallel, the effect of different K48-diUb and K48-triUb on isolated Jos aggregation was evaluated at both 37 °C and 50 °C. Surprisingly, results showed that at 50 °C K48-triUb markedly slows down Jos aggregation kinetics, supporting the hypothesis of interference of polyUb chains on Jos aggregation pathway, while at 37 °C K48-triUb dissociated from the complex leaving unbound Jos exposed to the solvent and free to form self-assemblies. This opens the way to further optimization of the presented experimental setup, using higher protein concentration or crosslinking polyUb and Jos, obtaining a covalently- bound and stable substrate.

5 Bibliography

1. Kovacs, G. G. Concepts and classification of neurodegenerative diseases. *Handb. Clin. Neurol.* **145**, 301–307 (2017).
2. Dugger, B. N. & Dickson, D. W. Pathology of Neurodegenerative Diseases. *Cold Spring Harb. Perspect. Biol.* **9**, a028035 (2017).
3. Matos, C. A., de Macedo-Ribeiro, S. & Carvalho, A. L. Polyglutamine diseases: the special case of ataxin-3 and Machado-Joseph disease. *Prog. Neurobiol.* **95**, 26–48 (2011).
4. McLoughlin, H. S., Moore, L. R. & Paulson, H. L. Pathogenesis of SCA3 and implications for other polyglutamine diseases. *Neurobiol. Dis.* **134**, 104635 (2020).
5. 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.* **283**, 26436–26443 (2008).
6. Burnett, B., Li, F. & Pittman, R. N. The polyglutamine neurodegenerative protein ataxin-3 binds polyubiquitylated proteins and has ubiquitin protease activity. *Hum. Mol. Genet.* **12**, 3195–3205 (2003).
7. 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.* **102**, 10493–10498 (2005).
8. Pastore, A. & Temussi, P. A. The two faces of Janus: functional interactions and protein aggregation. *Curr. Opin. Struct. Biol.* **22**, 30–37 (2012).
9. A, P. & P, T. Protein aggregation and misfolding: good or evil? *J. Phys. Condens. Matter Inst. Phys. J.* **24**, (2012).
10. Temussi, P. A., Tartaglia, G. G. & Pastore, A. The seesaw between normal function and protein aggregation: How functional interactions may increase protein solubility. *BioEssays News Rev. Mol. Cell. Dev. Biol.* **43**, e2100031 (2021).
11. Martin, J. B. Molecular basis of the neurodegenerative disorders. *N. Engl. J. Med.* **340**, 1970–1980 (1999).
12. Chiti, F. & Dobson, C. M. Protein Misfolding, Functional Amyloid, and Human Disease. *Annu. Rev. Biochem.* **75**, 333–366 (2006).
13. Vergara, L., Abid, K. & Soto, C. Protein Misfolding, a Common Mechanism in the Pathogenesis of Neurodegenerative Diseases. in 285–304 (2008). doi:10.1007/978-0-387-32671-9_12.
14. Takalo, M., Salminen, A., Soininen, H., Hiltunen, M. & Haapasalo, A. Protein aggregation and degradation mechanisms in neurodegenerative diseases. *Am. J. Neurodegener. Dis.* **2**, 1–14 (2013).
15. Devi, S., Chaturvedi, M., Fatima, S. & Priya, S. Environmental factors modulating protein conformations and their role in protein aggregation diseases. *Toxicology* **465**, 153049 (2022).
16. Sheikh, S., Safia, null, Haque, E. & Mir, S. S. Neurodegenerative Diseases: Multifactorial Conformational Diseases and Their Therapeutic Interventions. *J. Neurodegener. Dis.* **2013**, 563481 (2013).
17. Jellinger, K. A. & Attems, J. Challenges of multimorbidity of the aging brain: a critical update. *J. Neural Transm. Vienna Austria 1996* **122**, 505–521 (2015).
18. Kumar, V. *et al.* Protein aggregation and neurodegenerative diseases: From theory to therapy. *Eur. J. Med. Chem.* **124**, 1105–1120 (2016).
19. Chitnis, T. & Weiner, H. L. CNS inflammation and neurodegeneration. *J. Clin. Invest.* **127**, 3577–3587 (2017).
20. Chi, H., Chang, H.-Y. & Sang, T.-K. Neuronal Cell Death Mechanisms in Major Neurodegenerative Diseases. *Int. J. Mol. Sci.* **19**, 3082 (2018).
21. Thal, D. R., Rüb, U., Orantes, M. & Braak, H. Phases of A beta-deposition in the human brain and its relevance for the development of AD. *Neurology* **58**, 1791–1800 (2002).
22. Braak, H. & Braak, E. Neuropathological stageing of Alzheimer-related changes. *Acta Neuropathol. (Berl.)* **82**, 239–259 (1991).
23. Braak, H. *et al.* Staging of brain pathology related to sporadic Parkinson's disease. *Neurobiol. Aging* **24**, 197–211 (2003).
24. Brettschneider, J. *et al.* Stages of pTDP-43 pathology in amyotrophic lateral sclerosis. *Ann. Neurol.* **74**, 20–38 (2013).
25. Ross, C. A. Huntington's disease: new paths to pathogenesis. *Cell* **118**, 4–7 (2004).

26. Collinge, J. Prion diseases of humans and animals: their causes and molecular basis. *Annu. Rev. Neurosci.* **24**, 519–550 (2001).
27. Paulson, H. L. Protein fate in neurodegenerative proteinopathies: polyglutamine diseases join the (mis)fold. *Am. J. Hum. Genet.* **64**, 339–345 (1999).
28. Bourdenx, M. *et al.* Protein aggregation and neurodegeneration in prototypical neurodegenerative diseases: Examples of amyloidopathies, tauopathies and synucleinopathies. *Prog. Neurobiol.* **155**, 171–193 (2017).
29. Arar, S., Haque, M. A. & Kayed, R. Protein aggregation and neurodegenerative disease: Structural outlook for the novel therapeutics. *Proteins* (2023) doi:10.1002/prot.26561.
30. Vaquer-Alicea, J. & Diamond, M. I. Propagation of Protein Aggregation in Neurodegenerative Diseases. *Annu. Rev. Biochem.* **88**, 785–810 (2019).
31. Vijay-Kumar, S., Bugg, C. E. & Cook, W. J. Structure of ubiquitin refined at 1.8 A resolution. *J. Mol. Biol.* **194**, 531–544 (1987).
32. Tutar, Y. *et al.* Role of Protein Aggregation in Neurodegenerative Diseases. in *Neurodegenerative Diseases* (IntechOpen, 2013). doi:10.5772/54487.
33. Soto, C. Unfolding the role of protein misfolding in neurodegenerative diseases. *Nat. Rev. Neurosci.* **4**, 49–60 (2003).
34. Stefani, M. & Dobson, C. M. Protein aggregation and aggregate toxicity: new insights into protein folding, misfolding diseases and biological evolution. *J. Mol. Med.* **81**, 678–699 (2003).
35. Duran-Aniotz, C. *et al.* Amyloid pathology arrangements in Alzheimer's disease brains modulate in vivo seeding capability. *Acta Neuropathol. Commun.* **9**, 56 (2021).
36. Forno, L. S. Neuropathology of Parkinson's disease. *J. Neuropathol. Exp. Neurol.* **55**, 259–272 (1996).
37. Spillantini, M. G., Crowther, R. A., Jakes, R., Hasegawa, M. & Goedert, M. α-Synuclein in filamentous inclusions of Lewy bodies from Parkinson's disease and dementia with Lewy bodies. *Proc. Natl. Acad. Sci. U. S. A.* **95**, 6469–6473 (1998).
38. Outeiro, T. F. *et al.* Dementia with Lewy bodies: an update and outlook. *Mol. Neurodegener.* **14**, 5 (2019).
39. Mann, D. M. A. & Snowden, J. S. Frontotemporal lobar degeneration: Pathogenesis, pathology and pathways to phenotype. *Brain Pathol.* **27**, 723–736 (2017).
40. DiFiglia, M. *et al.* Aggregation of huntingtin in neuronal intranuclear inclusions and dystrophic neurites in brain. *Science* **277**, 1990–1993 (1997).
41. Bruijn, L. I. *et al.* Aggregation and motor neuron toxicity of an ALS-linked SOD1 mutant independent from wild-type SOD1. *Science* **281**, 1851–1854 (1998).
42. Paulson, H. L. *et al.* Machado-Joseph disease gene product is a cytoplasmic protein widely expressed in brain. *Ann. Neurol.* **41**, 453–462 (1997).
43. Duyckaerts, C., Dürr, A., Cancel, G. & Brice, A. Nuclear inclusions in spinocerebellar ataxia type 1. *Acta Neuropathol. (Berl.)* **97**, 201–207 (1999).
44. Cali, I. *et al.* Case Report: Histopathology and Prion Protein Molecular Properties in Inherited Prion Disease With a De Novo Seven-Octapeptide Repeat Insertion. *Front. Cell. Neurosci.* **14**, 150 (2020).
45. Baker, H. F., Ridley, R. M., Duchen, L. W., Crow, T. J. & Bruton, C. J. Evidence for the experimental transmission of cerebral beta-amyloidosis to primates. *Int. J. Exp. Pathol.* **74**, 441–454 (1993).
46. Kane, M. D. *et al.* Evidence for Seeding of β-Amyloid by Intracerebral Infusion of Alzheimer Brain Extracts in β-Amyloid Precursor Protein-Transgenic Mice. *J. Neurosci.* **20**, 3606–3611 (2000).
47. Meyer-Luehmann, M. *et al.* Exogenous Induction of Cerebral ß-Amyloidogenesis Is Governed by Agent and Host. *Science* **313**, 1781–1784 (2006).
48. Kordower, J. H., Chu, Y., Hauser, R. A., Freeman, T. B. & Olanow, C. W. Lewy body-like pathology in long-term embryonic nigral transplants in Parkinson's disease. *Nat. Med.* **14**, 504–506 (2008).
49. Hansen, C. *et al.* α-Synuclein propagates from mouse brain to grafted dopaminergic neurons and seeds aggregation in cultured human cells. *J. Clin. Invest.* **121**, 715–725 (2011).
50. Kordower, J. H. *et al.* Transfer of host-derived α synuclein to grafted dopaminergic neurons in rat. *Neurobiol. Dis.* **43**, 552–557 (2011).
51. Volpicelli-Daley, L. A. *et al.* Exogenous α-synuclein fibrils induce Lewy body pathology leading to synaptic dysfunction and neuron death. *Neuron* **72**, 57–71 (2011).
52. Walker, L. C. & Jucker, M. Neurodegenerative Diseases: Expanding the Prion Concept. *Annu. Rev. Neurosci.* **38**, 87–103 (2015).

53. Soto, C. & Pritzkow, S. Protein misfolding, aggregation, and conformational strains in neurodegenerative diseases. *Nat. Neurosci.* **21**, 1332–1340 (2018).
54. Walker, L., Stefanis, L. & Attems, J. Clinical and neuropathological differences between Parkinson's disease, Parkinson's disease dementia and dementia with Lewy bodies - current issues and future directions. *J. Neurochem.* **150**, 467–474 (2019).
55. Nigam, D., Singh, C. & Tiwari, U. Evaluation of in vitro study of antioxidant and antibacterial activities of methanolic seed extract of Sesamum indicum. *J. Pharmacogn. Phytochem.* **3**, (2014).
56. Taipa, R., Pinho, J. & Melo-Pires, M. Clinico-Pathological Correlations of the Most Common Neurodegenerative Dementias. *Front. Neurol.* **3**, 68 (2012).
57. Koga, S., Sekiya, H., Kondru, N., Ross, O. A. & Dickson, D. W. Neuropathology and molecular diagnosis of Synucleinopathies. *Mol. Neurodegener.* **16**, 83 (2021).
58. Dickson, D. W. & Lin, W.-L. Neuronal Inclusions. in *Encyclopedia of the Neurological Sciences (Second Edition)* (eds. Aminoff, M. J. & Daroff, R. B.) 441–455 (Academic Press, Oxford, 2014). doi:10.1016/B978-0-12-385157-4.00595-9.
59. Paulson, H. L. *et al.* Intranuclear inclusions of expanded polyglutamine protein in spinocerebellar ataxia type 3. *Neuron* **19**, 333–344 (1997).
60. Mok, T. H. & Mead, S. Prion diseases. *Medicine (Baltimore)* **45**, 674–677 (2017).
61. Herczenik, E. & Gebbink, M. F. B. G. Molecular and cellular aspects of protein misfolding and disease. *FASEB J.* **22**, 2115–2133 (2008).
62. Dill, K. A. & MacCallum, J. L. The protein-folding problem, 50 years on. *Science* **338**, 1042–1046 (2012).
63. Onuchic, J. N., Luthey-Schulten, Z. & Wolynes, P. G. Theory of protein folding: the energy landscape perspective. *Annu. Rev. Phys. Chem.* **48**, 545–600 (1997).
64. Balch, W. E., Morimoto, R. I., Dillin, A. & Kelly, J. W. Adapting proteostasis for disease intervention. *Science* **319**, 916–919 (2008).
65. Sweeney, P. *et al.* Protein misfolding in neurodegenerative diseases: implications and strategies. *Transl. Neurodegener.* **6**, 6 (2017).
66. Baldwin, A. J. *et al.* Metastability of native proteins and the phenomenon of amyloid formation. *J. Am. Chem. Soc.* **133**, 14160–14163 (2011).
67. Muntau, A. C., Leandro, J., Staudigl, M., Mayer, F. & Gersting, S. W. Innovative strategies to treat protein misfolding in inborn errors of metabolism: pharmacological chaperones and proteostasis regulators. *J. Inherit. Metab. Dis.* **37**, 505–523 (2014).
68. Broadley, S. A. & Hartl, F. U. The role of molecular chaperones in human misfolding diseases. *FEBS Lett.* **583**, 2647–2653 (2009).
69. Winklhofer, K. F., Tatzelt, J. & Haass, C. The two faces of protein misfolding: gain- and loss-of-function in neurodegenerative diseases. *EMBO J.* **27**, 336–349 (2008).
70. De Baets, G., Van Doorn, L., Rousseau, F. & Schymkowitz, J. Increased Aggregation Is More Frequently Associated to Human Disease-Associated Mutations Than to Neutral Polymorphisms. *PLoS Comput. Biol.* **11**, e1004374 (2015).
71. Chiti, F. & Dobson, C. M. Protein Misfolding, Amyloid Formation, and Human Disease: A Summary of Progress Over the Last Decade. *Annu. Rev. Biochem.* **86**, 27–68 (2017).
72. Candelise, N. *et al.* Protein Aggregation Landscape in Neurodegenerative Diseases: Clinical Relevance and Future Applications. *Int. J. Mol. Sci.* **22**, 6016 (2021).
73. Davis, J. K. & Sindi, S. S. A Study in Nucleated Polymerization Models of Protein Aggregation. *Appl. Math. Lett.* **40**, 97–101 (2015).
74. Morris, A. M., Watzky, M. A. & Finke, R. G. Protein aggregation kinetics, mechanism, and curve-fitting: a review of the literature. *Biochim. Biophys. Acta* **1794**, 375–397 (2009).
75. Chiti, F. & Dobson, C. M. Amyloid formation by globular proteins under native conditions. *Nat. Chem. Biol.* **5**, 15–22 (2009).
76. Kumar, S. & Udgaonkar, J. B. Structurally distinct amyloid protofibrils form on separate pathways of aggregation of a small protein. *Biochemistry* **48**, 6441–6449 (2009).
77. Vetri, V. & Foderà, V. The route to protein aggregate superstructures: Particulates and amyloid-like spherulites. *FEBS Lett.* **589**, 2448–2463 (2015).
78. Morell, M. *et al.* Inclusion bodies: specificity in their aggregation process and amyloid-like structure. *Biochim. Biophys. Acta* **1783**, 1815–1825 (2008).

79. Wu, J. *et al.* Identifying the role of co-aggregation of Alzheimer's amyloid-β with amorphous protein aggregates of non-amyloid proteins. *Cell Rep. Phys. Sci.* **3**, 101028 (2022).
80. Vahdat, A. *The Importance of Macrophages, Lipid Membranes and Seeding in Experimental AA Amyloidosis*. (2019). doi:10.3384/diss.diva-159658.
81. Arosio, P., Knowles, T. P. J. & Linse, S. On the lag phase in amyloid fibril formation. *Phys. Chem. Chem. Phys. PCCP* **17**, 7606–7618 (2015).
82. Naiki, H. *et al.* Establishment of a kinetic model of dialysis-related amyloid fibril extension in vitro. *Amyloid* **4**, 223–232 (1997).
83. Serio, T. R. *et al.* Nucleated conformational conversion and the replication of conformational information by a prion determinant. *Science* **289**, 1317–1321 (2000).
84. Uversky, V. N. *et al.* Biophysical properties of the synucleins and their propensities to fibrillate: inhibition of alpha-synuclein assembly by beta- and gamma-synucleins. *J. Biol. Chem.* **277**, 11970–11978 (2002).
85. Chiti, F., Stefani, M., Taddei, N., Ramponi, G. & Dobson, C. M. Rationalization of the effects of mutations on peptide andprotein aggregation rates. *Nature* **424**, 805–808 (2003).
86. Iannuzzi, C., Irace, G. & Sirangelo, I. The Effect of Glycosaminoglycans (GAGs) on Amyloid Aggregation and Toxicity. *Molecules* **20**, 2510–2528 (2015).
87. El-Agnaf, O. M. A., Walsh, D. M. & Allsop, D. Soluble oligomers for the diagnosis of neurodegenerative diseases. *Lancet Neurol.* **2**, 461–462 (2003).
88. Haass, C. & Selkoe, D. J. Soluble protein oligomers in neurodegeneration: lessons from the Alzheimer's amyloid beta-peptide. *Nat. Rev. Mol. Cell Biol.* **8**, 101–112 (2007).
89. Falsone, A. & Falsone, S. F. Legal but lethal: functional protein aggregation at the verge of toxicity. *Front. Cell. Neurosci.* **9**, 45 (2015).
90. Verma, M., Vats, A. & Taneja, V. Toxic species in amyloid disorders: Oligomers or mature fibrils. *Ann. Indian Acad. Neurol.* **18**, 138–145 (2015).
91. Arrasate, M., Mitra, S., Schweitzer, E. S., Segal, M. R. & Finkbeiner, S. Inclusion body formation reduces levels of mutant huntingtin and the risk of neuronal death. *Nature* **431**, 805–810 (2004).
92. Cummings, C. J. *et al.* Mutation of the E6-AP ubiquitin ligase reduces nuclear inclusion frequency while accelerating polyglutamine-induced pathology in SCA1 mice. *Neuron* **24**, 879–892 (1999).
93. Johnston, J. A., Ward, C. L. & Kopito, R. R. Aggresomes: a cellular response to misfolded proteins. *J. Cell Biol.* **143**, 1883–1898 (1998).
94. Taylor, J. P. *et al.* Aggresomes protect cells by enhancing the degradation of toxic polyglutamine-containing protein. *Hum. Mol. Genet.* **12**, 749–757 (2003).
95. Treusch, S., Cyr, D. M. & Lindquist, S. Amyloid deposits: protection against toxic protein species? *Cell Cycle Georget. Tex* **8**, 1668–1674 (2009).
96. Wu, J. W. *et al.* Fibrillar oligomers nucleate the oligomerization of monomeric amyloid beta but do not seed fibril formation. *J. Biol. Chem.* **285**, 6071–6079 (2010).
97. Wells, C., Brennan, S., Keon, M. & Ooi, L. The role of amyloid oligomers in neurodegenerative pathologies. *Int. J. Biol. Macromol.* **181**, 582–604 (2021).
98. Migliore, L. & Coppedè, F. Genetics, environmental factors and the emerging role of epigenetics in neurodegenerative diseases. *Mutat. Res.* **667**, 82–97 (2009).
99. Cacace, R., Sleegers, K. & Van Broeckhoven, C. Molecular genetics of early-onset Alzheimer's disease revisited. *Alzheimers Dement. J. Alzheimers Assoc.* **12**, 733–748 (2016).
100. Wingo, T. S., Lah, J. J., Levey, A. I. & Cutler, D. J. Autosomal recessive causes likely in early-onset Alzheimer disease. *Arch. Neurol.* **69**, 59–64 (2012).
101. Hoogmartens, J., Cacace, R. & Van Broeckhoven, C. Insight into the genetic etiology of Alzheimer's disease: A comprehensive review of the role of rare variants. *Alzheimers Dement. Amst. Neth.* **13**, e12155 (2021).
102. Lesage, S. & Brice, A. Parkinson's disease: from monogenic forms to genetic susceptibility factors. *Hum. Mol. Genet.* **18**, R48-59 (2009).
103. Bandres-Ciga, S., Diez-Fairen, M., Kim, J. J. & Singleton, A. B. Genetics of Parkinson's disease: An introspection of its journey towards precision medicine. *Neurobiol. Dis.* **137**, 104782 (2020).
104. Nalls, M. A. *et al.* Identification of novel risk loci, causal insights, and heritable risk for Parkinson's disease: a meta-genome wide association study. *Lancet Neurol.* **18**, 1091–1102 (2019).
105. Lambert, J. C. *et al.* Meta-analysis of 74,046 individuals identifies 11 new susceptibility loci for Alzheimer's disease. *Nat. Genet.* **45**, 1452–1458 (2013).

106. Ciurea, A. V. *et al.* Unraveling Molecular and Genetic Insights into Neurodegenerative Diseases: Advances in Understanding Alzheimer's, Parkinson's, and Huntington's Diseases and Amyotrophic Lateral Sclerosis. *Int. J. Mol. Sci.* **24**, 10809 (2023).
107. Polymeropoulos, M. H. *et al.* Mutation in the alpha-synuclein gene identified in families with Parkinson's disease. *Science* **276**, 2045–2047 (1997).
108. Petrucci, S., Ginevrino, M. & Valente, E. M. Phenotypic spectrum of alpha-synuclein mutations: New insights from patients and cellular models. *Parkinsonism Relat. Disord.* **22 Suppl 1**, S16-20 (2016).
109. Chartier-Harlin, M. C. *et al.* Early-onset Alzheimer's disease caused by mutations at codon 717 of the beta-amyloid precursor protein gene. *Nature* **353**, 844–846 (1991).
110. Goate, A. *et al.* Segregation of a missense mutation in the amyloid precursor protein gene with familial Alzheimer's disease. *Nature* **349**, 704–706 (1991).
111. Murrell, J., Farlow, M., Ghetti, B. & Benson, M. D. A mutation in the amyloid precursor protein associated with hereditary Alzheimer's disease. *Science* **254**, 97–99 (1991).
112. Sherrington, R. *et al.* Cloning of a gene bearing missense mutations in early-onset familial Alzheimer's disease. *Nature* **375**, 754–760 (1995).
113. Levy-Lahad, E. *et al.* Candidate gene for the chromosome 1 familial Alzheimer's disease locus. *Science* **269**, 973–977 (1995).
114. Rogaev, E. I. *et al.* Familial Alzheimer's disease in kindreds with missense mutations in a gene on chromosome 1 related to the Alzheimer's disease type 3 gene. *Nature* **376**, 775–778 (1995).
115. Chiò, A. *et al.* Global epidemiology of amyotrophic lateral sclerosis: a systematic review of the published literature. *Neuroepidemiology* **41**, 118–130 (2013).
116. Ratnavalli, E., Brayne, C., Dawson, K. & Hodges, J. R. The prevalence of frontotemporal dementia. *Neurology* **58**, 1615–1621 (2002).
117. Abramzon, Y. A., Fratta, P., Traynor, B. J. & Chia, R. The Overlapping Genetics of Amyotrophic Lateral Sclerosis and Frontotemporal Dementia. *Front. Neurosci.* **14**, 42 (2020).
118. Burrell, J. R., Kiernan, M. C., Vucic, S. & Hodges, J. R. Motor neuron dysfunction in frontotemporal dementia. *Brain J. Neurol.* **134**, 2582–2594 (2011).
119. Turner, M. R. *et al.* Controversies and priorities in amyotrophic lateral sclerosis. *Lancet Neurol.* **12**, 310–322 (2013).
120. Pihlstrøm, L., Wiethoff, S. & Houlden, H. Genetics of neurodegenerative diseases: an overview. *Handb. Clin. Neurol.* **145**, 309–323 (2017).
121. Andrew, S. E. *et al.* The relationship between trinucleotide (CAG) repeat length and clinical features of Huntington's disease. *Nat. Genet.* **4**, 398–403 (1993).
122. Margulis, B. A., Vigont, V., Lazarev, V. F., Kaznacheyeva, E. V. & Guzhova, I. V. Pharmacological protein targets in polyglutamine diseases: mutant polypeptides and their interactors. *FEBS Lett.* **587**, 1997–2007 (2013).
123. Gatchel, J. R. & Zoghbi, H. Y. Diseases of unstable repeat expansion: mechanisms and common principles. *Nat. Rev. Genet.* **6**, 743–755 (2005).
124. Klockgether, T., Mariotti, C. & Paulson, H. L. Spinocerebellar ataxia. *Nat. Rev. Dis. Primer* **5**, 24 (2019).
125. Paulson, H. L., Shakkottai, V. G., Clark, H. B. & Orr, H. T. Polyglutamine spinocerebellar ataxias - from genes to potential treatments. *Nat. Rev. Neurosci.* **18**, 613–626 (2017).
126. A novel gene containing a trinucleotide repeat that is expanded and unstable on Huntington's disease chromosomes. The Huntington's Disease Collaborative Research Group. *Cell* **72**, 971–983 (1993).
127. Nahhas, F. A., Garbern, J., Krajewski, K. M., Roa, B. B. & Feldman, G. L. Juvenile onset Huntington disease resulting from a very large maternal expansion. *Am. J. Med. Genet. A.* **137A**, 328–331 (2005).
128. Matilla, T. *et al.* Presymptomatic analysis of spinocerebellar ataxia type 1 (SCA1) via the expansion of the SCA1 CAG-repeat in a large pedigree displaying anticipation and parental male bias. *Hum. Mol. Genet.* **2**, 2123–2128 (1993).
129. Ranum, L. P. W. *et al.* Molecular and Clinical Correlations in Spinocerebellar Ataxia Type I: Evidence for Familial Effects on the Age at Onset. *Am. J. Hum. Genet.* **55**, 244–252 (1994).
130. Kawaguchi, Y. *et al.* CAG expansions in a novel gene for Machado-Joseph disease at chromosome 14q32.1. *Nat. Genet.* **8**, 221–228 (1994).
131. Padiath, Q. S., Srivastava, A. K., Roy, S., Jain, S. & Brahmachari, S. K. Identification of a novel 45 repeat unstable allele associated with a disease phenotype at the MJD1/SCA3 locus. *Am. J. Med. Genet. Part B Neuropsychiatr. Genet. Off. Publ. Int. Soc. Psychiatr. Genet.* **133B**, 124–126 (2005).

132. Pulst, S.-M. *et al.* Spinocerebellar ataxia type 2: polyQ repeat variation in the CACNA1A calcium channel modifies age of onset. *Brain J. Neurol.* **128**, 2297–2303 (2005).
133. Zhuchenko, O. *et al.* Autosomal dominant cerebellar ataxia (SCA6) associated with small polyglutamine expansions in the alpha 1A-voltage-dependent calcium channel. *Nat. Genet.* **15**, 62–69 (1997).
134. Benton, C. S. *et al.* Molecular and clinical studies in SCA-7 define a broad clinical spectrum and the infantile phenotype. *Neurology* **51**, 1081–1086 (1998).
135. David, G. *et al.* Cloning of the SCA7 gene reveals a highly unstable CAG repeat expansion. *Nat. Genet.* **17**, 65–70 (1997).
136. Koide, R. *et al.* A neurological disease caused by an expanded CAG trinucleotide repeat in the TATA-binding protein gene: a new polyglutamine disease? *Hum. Mol. Genet.* **8**, 2047–2053 (1999).
137. Nakamura, K. *et al.* SCA17, a novel autosomal dominant cerebellar ataxia caused by an expanded polyglutamine in TATA-binding protein. *Hum. Mol. Genet.* **10**, 1441–1448 (2001).
138. Komure, O. *et al.* DNA analysis in hereditary dentatorubral-pallidoluysian atrophy: correlation between CAG repeat length and phenotypic variation and the molecular basis of anticipation. *Neurology* **45**, 143–149 (1995).
139. La Spada, A. R., Wilson, E. M., Lubahn, D. B., Harding, A. E. & Fischbeck, K. H. Androgen receptor gene mutations in X-linked spinal and bulbar muscular atrophy. *Nature* **352**, 77–79 (1991).
140. Saunders, H. M. & Bottomley, S. P. Multi-domain misfolding: understanding the aggregation pathway of polyglutamine proteins. *Protein Eng. Des. Sel. PEDS* **22**, 447–451 (2009).
141. Robertson, A. L. & Bottomley, S. P. Towards the treatment of polyglutamine diseases: the modulatory role of protein context. *Curr. Med. Chem.* **17**, 3058–3068 (2010).
142. Fan, H.-C. *et al.* Polyglutamine (PolyQ) diseases: genetics to treatments. *Cell Transplant.* **23**, 441–458 (2014).
143. Zoghbi, H. Y. & Orr, H. T. Glutamine repeats and neurodegeneration. *Annu. Rev. Neurosci.* **23**, 217–247 (2000).
144. Yushchenko, T., Deuerling, E. & Hauser, K. Insights into the Aggregation Mechanism of PolyQ Proteins with Different Glutamine Repeat Lengths. *Biophys. J.* **114**, 1847–1857 (2018).
145. Ross, C. A., Poirier, M. A., Wanker, E. E. & Amzel, M. Polyglutamine fibrillogenesis: The pathway unfolds. *Proc. Natl. Acad. Sci. U. S. A.* **100**, 1–3 (2003).
146. Schmitz-Hübsch, T. *et al.* Spinocerebellar ataxia types 1, 2, 3, and 6: disease severity and nonataxia symptoms. *Neurology* **71**, 982–989 (2008).
147. Maciel, P. *et al.* Correlation between CAG repeat length and clinical features in Machado-Joseph disease. *Am. J. Hum. Genet.* **57**, 54–61 (1995).
148. Eisenberg, D. & Jucker, M. The Amyloid State of Proteins in Human Diseases. *Cell* **148**, 1188–1203 (2012).
149. Masino, L. *et al.* The Josephin domain determines the morphological and mechanical properties of ataxin-3 fibrils. *Biophys. J.* **100**, 2033–2042 (2011).
150. Kopito, R. R. Aggresomes, inclusion bodies and protein aggregation. *Trends Cell Biol.* **10**, 524–530 (2000).
151. Shao, J. & Diamond, M. I. Polyglutamine diseases: emerging concepts in pathogenesis and therapy. *Hum. Mol. Genet.* **16 Spec No. 2**, R115-123 (2007).
152. Ellisdon, A. M., Thomas, B. & Bottomley, S. P. The two-stage pathway of ataxin-3 fibrillogenesis involves a polyglutamine-independent step. *J. Biol. Chem.* **281**, 16888–16896 (2006).
153. Ellisdon, A. M., Pearce, M. C. & Bottomley, S. P. Mechanisms of ataxin-3 misfolding and fibril formation: kinetic analysis of a disease-associated polyglutamine protein. *J. Mol. Biol.* **368**, 595–605 (2007).
154. Haacke, A. *et al.* Proteolytic cleavage of polyglutamine-expanded ataxin-3 is critical for aggregation and sequestration of non-expanded ataxin-3. *Hum. Mol. Genet.* **15**, 555–568 (2006).
155. Berke, S. J. S., Schmied, F. A. F., Brunt, E. R., Ellerby, L. M. & Paulson, H. L. Caspase-mediated proteolysis of the polyglutamine disease protein ataxin-3. *J. Neurochem.* **89**, 908–918 (2004).
156. Wellington, C. L. *et al.* Caspase cleavage of gene products associated with triplet expansion disorders generates truncated fragments containing the polyglutamine tract. *J. Biol. Chem.* **273**, 9158–9167 (1998).
157. Weber, J. J. *et al.* A combinatorial approach to identify calpain cleavage sites in the Machado-Joseph disease protein ataxin-3. *Brain J. Neurol.* **140**, 1280–1299 (2017).

158. Goti, D. *et al.* A mutant ataxin-3 putative-cleavage fragment in brains of Machado-Joseph disease patients and transgenic mice is cytotoxic above a critical concentration. *J. Neurosci. Off. J. Soc. Neurosci.* **24**, 10266–10279 (2004).
159. Liman, J. *et al.* CDK5 protects from caspase-induced Ataxin-3 cleavage and neurodegeneration. *J. Neurochem.* **129**, 1013–1023 (2014).
160. Araujo, M. *et al.* Trends in the Epidemiology of Spinocerebellar Ataxia Type 3 / Machado-Joseph Disease in the Azores Islands, Portugal. in (2016).
161. Jardim, L. B. *et al.* A survey of spinocerebellar ataxia in South Brazil - 66 new cases with Machado-Joseph disease, SCA7, SCA8, or unidentified disease-causing mutations. *J. Neurol.* **248**, 870–876 (2001).
162. Vale, J. *et al.* Autosomal dominant cerebellar ataxia: frequency analysis and clinical characterization of 45 families from Portugal. *Eur. J. Neurol.* **17**, 124–128 (2010).
163. Maruyama, H. *et al.* Molecular features of the CAG repeats and clinical manifestation of Machado-Joseph disease. *Hum. Mol. Genet.* **4**, 807–812 (1995).
164. Nakano, K. K., Dawson, D. M. & Spence, A. Machado disease. A hereditary ataxia in Portuguese emigrants to Massachusetts. *Neurology* **22**, 49–55 (1972).
165. Haberhausen, G., Damian, M. S., Leweke, F. & Müller, U. Spinocerebellar ataxia, type 3 (SCA3) is genetically identical to Machado-Joseph disease (MJD). *J. Neurol. Sci.* **132**, 71–75 (1995).
166. Dürr, A. *et al.* Spinocerebellar ataxia 3 and Machado-Joseph disease: clinical, molecular, and neuropathological features. *Ann. Neurol.* **39**, 490–499 (1996).
167. Paulson, H. Machado-Joseph disease/spinocerebellar ataxia type 3. *Handb. Clin. Neurol.* **103**, 437–449 (2012).
168. Ashizawa, T., Öz, G. & Paulson, H. L. Spinocerebellar ataxias: prospects and challenges for therapy development. *Nat. Rev. Neurol.* **14**, 590–605 (2018).
169. Riess, O., Rüb, U., Pastore, A., Bauer, P. & Schöls, L. SCA3: neurological features, pathogenesis and animal models. *Cerebellum Lond. Engl.* **7**, 125–137 (2008).
170. Rosenberg, R. N. Machado-Joseph disease: an autosomal dominant motor system degeneration. *Mov. Disord. Off. J. Mov. Disord. Soc.* **7**, 193–203 (1992).
171. Soong, B., Cheng, C., Liu, R. & Shan, D. Machado-Joseph disease: clinical, molecular, and metabolic characterization in Chinese kindreds. *Ann. Neurol.* **41**, 446–452 (1997).
172. Sudarsky, L. & Coutinho, P. Machado-Joseph disease. *Clin. Neurosci. N. Y. N* **3**, 17–22 (1995).
173. Rüb, U. *et al.* Clinical features, neurogenetics and neuropathology of the polyglutamine spinocerebellar ataxias type 1, 2, 3, 6 and 7. *Prog. Neurobiol.* **104**, 38–66 (2013).
174. Matos, C. A., de Almeida, L. P. & Nóbrega, C. Machado-Joseph disease/spinocerebellar ataxia type 3: lessons from disease pathogenesis and clues into therapy. *J. Neurochem.* **148**, 8–28 (2019).
175. Wan, N., Chen, Z., Wan, L., Tang, B. & Jiang, H. MR Imaging of SCA3/MJD. *Front. Neurosci.* **14**, (2020).
176. Putka, A. F., Mato, J. P. & McLoughlin, H. S. Myelinating Glia: Potential Therapeutic Targets in Polyglutamine Spinocerebellar Ataxias. *Cells* **12**, 601 (2023).
177. 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. Zurich Switz.* **8**, 669–679 (1998).
178. Sittler, A. *et al.* Deregulation of autophagy in postmortem brains of Machado-Joseph disease patients. *Neuropathol. Off. J. Jpn. Soc. Neuropathol.* **38**, 113–124 (2018).
179. Uchihara, T. *et al.* Non-expanded polyglutamine proteins in intranuclear inclusions of hereditary ataxias--triple-labeling immunofluorescence study. *Acta Neuropathol. (Berl.)* **102**, 149–152 (2001).
180. Yamada, M., Hayashi, S., Tsuji, S. & Takahashi, H. Involvement of the cerebral cortex and autonomic ganglia in Machado-Joseph disease. *Acta Neuropathol. (Berl.)* **101**, 140–144 (2001).
181. Hayashi, M., Kobayashi, K. & Furuta, H. Immunohistochemical study of neuronal intranuclear and cytoplasmic inclusions in Machado-Joseph disease. *Psychiatry Clin. Neurosci.* **57**, 205–213 (2003).
182. Seidel, K. *et al.* Axonal inclusions in spinocerebellar ataxia type 3. *Acta Neuropathol. (Berl.)* **120**, 449–460 (2010).
183. Seidel, K. *et al.* On the distribution of intranuclear and cytoplasmic aggregates in the brainstem of patients with spinocerebellar ataxia type 2 and 3. *Brain Pathol. Zurich Switz.* **27**, 345–355 (2017).
184. Costa, M. do C. & Paulson, H. L. Toward understanding Machado-Joseph disease. *Prog. Neurobiol.* **97**, 239–257 (2012).

185. Ichikawa, Y. *et al.* The genomic structure and expression of MJD, the Machado-Joseph disease gene. *J. Hum. Genet.* **46**, 413–422 (2001).
186. Bettencourt, C. & Lima, M. Machado-Joseph Disease: from first descriptions to new perspectives. *Orphanet J. Rare Dis.* **6**, 35 (2011).
187. Evers, M. M., Toonen, L. J. A. & van Roon-Mom, W. M. C. Ataxin-3 protein and RNA toxicity in spinocerebellar ataxia type 3: current insights and emerging therapeutic strategies. *Mol. Neurobiol.* **49**, 1513–1531 (2014).
188. Carvalho, A. L., Silva, A. & Macedo-Ribeiro, S. Polyglutamine-Independent Features in Ataxin-3 Aggregation and Pathogenesis of Machado-Joseph Disease. *Adv. Exp. Med. Biol.* **1049**, 275–288 (2018).
189. Zhemkov, V. A., Kulminskaya, A. A., Bezprozvanny, I. B. & Kim, M. The 2.2-Angstrom resolution crystal structure of the carboxy-terminal region of ataxin-3. *FEBS Open Bio* **6**, 168–178 (2016).
190. Kim, M. W., Chelliah, Y., Kim, S. W., Otwinowski, Z. & Bezprozvanny, I. Secondary structure of Huntingtin amino-terminal region. *Struct. Lond. Engl. 1993* **17**, 1205–1212 (2009).
191. 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.* **102**, 12700–12705 (2005).
192. Chai, Y., Berke, S. S., Cohen, R. E. & Paulson, H. L. Poly-ubiquitin binding by the polyglutamine disease protein ataxin-3 links its normal function to protein surveillance pathways. *J. Biol. Chem.* **279**, 3605–3611 (2004).
193. Satoh, T. *et al.* Mode of substrate recognition by the Josephin domain of ataxin-3, which has an endo-type deubiquitinase activity. *FEBS Lett.* **588**, 4422–4430 (2014).
194. 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* **5**, e13202 (2010).
195. Antony, P. M. A. *et al.* Identification and functional dissection of localization signals within ataxin-3. *Neurobiol. Dis.* **36**, 280–292 (2009).
196. Macedo-Ribeiro, S., Cortes, L., Maciel, P. & Carvalho, A. L. Nucleocytoplasmic shuttling activity of ataxin-3. *PloS One* **4**, e5834 (2009).
197. Uhlén, M. *et al.* Proteomics. Tissue-based map of the human proteome. *Science* **347**, 1260419 (2015).
198. Schmitt, I. *et al.* Inactivation of the mouse Atxn3 (ataxin-3) gene increases protein ubiquitination. *Biochem. Biophys. Res. Commun.* **362**, 734–739 (2007).
199. Komander, D. The emerging complexity of protein ubiquitination. *Biochem. Soc. Trans.* **37**, 937–953 (2009).
200. Wang, G., Sawai, N., Kotliarova, S., Kanazawa, I. & Nukina, N. Ataxin-3, the MJD1 gene product, interacts with the two human homologs of yeast DNA repair protein RAD23, HHR23A and HHR23B. *Hum. Mol. Genet.* **9**, 1795–1803 (2000).
201. Malisauskas, M. *et al.* Does the cytotoxic effect of transient amyloid oligomers from common equine lysozyme in vitro imply innate amyloid toxicity? *J. Biol. Chem.* **280**, 6269–6275 (2005).
202. Kayed, R. *et al.* Common structure of soluble amyloid oligomers implies common mechanism of pathogenesis. *Science* **300**, 486–489 (2003).
203. Nicastro, G. *et al.* Josephin domain of ataxin-3 contains two distinct ubiquitin-binding sites. *Biopolymers* **91**, 1203–1214 (2009).
204. Mueller, T. *et al.* CK2-dependent phosphorylation determines cellular localization and stability of ataxin-3. *Hum. Mol. Genet.* **18**, 3334–3343 (2009).
205. Tao, R.-S., Fei, E.-K., Ying, Z., Wang, H.-F. & Wang, G.-H. Casein kinase 2 interacts with and phosphorylates ataxin-3. *Neurosci. Bull.* **24**, 271–277 (2008).
206. Hirabayashi, M. *et al.* VCP/p97 in abnormal protein aggregates, cytoplasmic vacuoles, and cell death, phenotypes relevant to neurodegeneration. *Cell Death Differ.* **8**, 977–984 (2001).
207. Boeddrich, A. *et al.* An arginine/lysine-rich motif is crucial for VCP/p97-mediated modulation of ataxin-3 fibrillogenesis. *EMBO J.* **25**, 1547–1558 (2006).
208. Ristic, G., Sutton, J. R., Libohova, K. & Todi, S. V. Toxicity and aggregation of the polyglutamine disease protein, ataxin-3 is regulated by its binding to VCP/p97 in Drosophila melanogaster. *Neurobiol. Dis.* **116**, 78–92 (2018).
209. Johnson, S. L. *et al.* Targeting the VCP-binding motif of ataxin-3 improves phenotypes in Drosophila models of Spinocerebellar Ataxia Type 3. *Neurobiol. Dis.* **160**, 105516 (2021).
210. Li, F., Macfarlan, T., Pittman, R. N. & Chakravarti, D. Ataxin-3 is a histone-binding protein with two independent transcriptional corepressor activities. *J. Biol. Chem.* **277**, 45004–45012 (2002).

211. Evert, B. O. *et al.* High level expression of expanded full-length ataxin-3 in vitro causes cell death and formation of intranuclear inclusions in neuronal cells. *Hum. Mol. Genet.* **8**, 1169–1176 (1999).
212. Burnett, B. G. & Pittman, R. N. The polyglutamine neurodegenerative protein ataxin 3 regulates aggresome formation. *Proc. Natl. Acad. Sci. U. S. A.* **102**, 4330–4335 (2005).
213. Mazzucchelli, S. *et al.* Proteomic and biochemical analyses unveil tight interaction of ataxin-3 with tubulin. *Int. J. Biochem. Cell Biol.* **41**, 2485–2492 (2009).
214. Rodrigues, A.-J. *et al.* Absence of ataxin-3 leads to cytoskeletal disorganization and increased cell death. *Biochim. Biophys. Acta* **1803**, 1154–1163 (2010).
215. Bevivino, A. E. & Loll, P. J. An expanded glutamine repeat destabilizes native ataxin-3 structure and mediates formation of parallel beta -fibrils. *Proc. Natl. Acad. Sci. U. S. A.* **98**, 11955–11960 (2001).
216. Masino, L. *et al.* Characterization of the structure and the amyloidogenic properties of the Josephin domain of the polyglutamine-containing protein ataxin-3. *J. Mol. Biol.* **344**, 1021–1035 (2004).
217. Chow, M. K. M., Mackay, J. P., Whisstock, J. C., Scanlon, M. J. & Bottomley, S. P. Structural and functional analysis of the Josephin domain of the polyglutamine protein ataxin-3. *Biochem. Biophys. Res. Commun.* **322**, 387–394 (2004).
218. Chow, M. K. M., Ellisdon, A. M., Cabrita, L. D. & Bottomley, S. P. Polyglutamine expansion in ataxin-3 does not affect protein stability: implications for misfolding and disease. *J. Biol. Chem.* **279**, 47643–47651 (2004).
219. Menon, R. P. & Pastore, A. Expansion of amino acid homo-sequences in proteins: insights into the role of amino acid homo-polymers and of the protein context in aggregation. *Cell. Mol. Life Sci. CMLS* **63**, 1677–1685 (2006).
220. Bhattacharyya, A. *et al.* Oligoproline effects on polyglutamine conformation and aggregation. *J. Mol. Biol.* **355**, 524–535 (2006).
221. Darnell, G., Orgel, J. P. R. O., Pahl, R. & Meredith, S. C. Flanking polyproline sequences inhibit beta-sheet structure in polyglutamine segments by inducing PPII-like helix structure. *J. Mol. Biol.* **374**, 688–704 (2007).
222. Duennwald, M. L., Jagadish, S., Muchowski, P. J. & Lindquist, S. Flanking sequences profoundly alter polyglutamine toxicity in yeast. *Proc. Natl. Acad. Sci. U. S. A.* **103**, 11045–11050 (2006).
223. de Chiara, C., Menon, R. P., Dal Piaz, F., Calder, L. & Pastore, A. Polyglutamine is not all: the functional role of the AXH domain in the ataxin-1 protein. *J. Mol. Biol.* **354**, 883–893 (2005).
224. Thakur, A. K. *et al.* Polyglutamine disruption of the huntingtin exon 1 N terminus triggers a complex aggregation mechanism. *Nat. Struct. Mol. Biol.* **16**, 380–389 (2009).
225. Chen, S., Berthelier, V., Hamilton, J. B., O'Nuallain, B. & Wetzel, R. Amyloid-like features of polyglutamine aggregates and their assembly kinetics. *Biochemistry* **41**, 7391–7399 (2002).
226. 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.* **353**, 642–654 (2005).
227. 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 MCP* **14**, 1241–1253 (2015).
228. Natalello, A. *et al.* A Major Role for Side-Chain Polyglutamine Hydrogen Bonding in Irreversible Ataxin-3 Aggregation. *PLoS ONE* **6**, e18789 (2011).
229. Silva, A. *et al.* Distribution of Amyloid-Like and Oligomeric Species from Protein Aggregation Kinetics. *Angew. Chem. Int. Ed Engl.* **56**, 14042–14045 (2017).
230. Lupton, C. J. *et al.* Enhanced molecular mobility of ordinarily structured regions drives polyglutamine disease. *J. Biol. Chem.* **290**, 24190–24200 (2015).
231. Masino, L., Nicastro, G., Calder, L., Vendruscolo, M. & Pastore, A. Functional interactions as a survival strategy against abnormal aggregation. *FASEB J. Off. Publ. Fed. Am. Soc. Exp. Biol.* **25**, 45–54 (2011).
232. Koch, P. *et al.* Excitation-induced ataxin-3 aggregation in neurons from patients with Machado-Joseph disease. *Nature* **480**, 543–546 (2011).
233. Takahashi, T. *et al.* Soluble polyglutamine oligomers formed prior to inclusion body formation are cytotoxic. *Hum. Mol. Genet.* **17**, 345–356 (2008).
234. Jonasz Jeremiasz Weber. Calpains in the Molecular Pathogenesis of Polyglutamine Disorders and their Potential as a Therapeutic Target. (2017) doi:10.13140/RG.2.2.23946.11208.
235. Ikeda, H. *et al.* Expanded polyglutamine in the Machado-Joseph disease protein induces cell death in vitro and in vivo. *Nat. Genet.* **13**, 196–202 (1996).

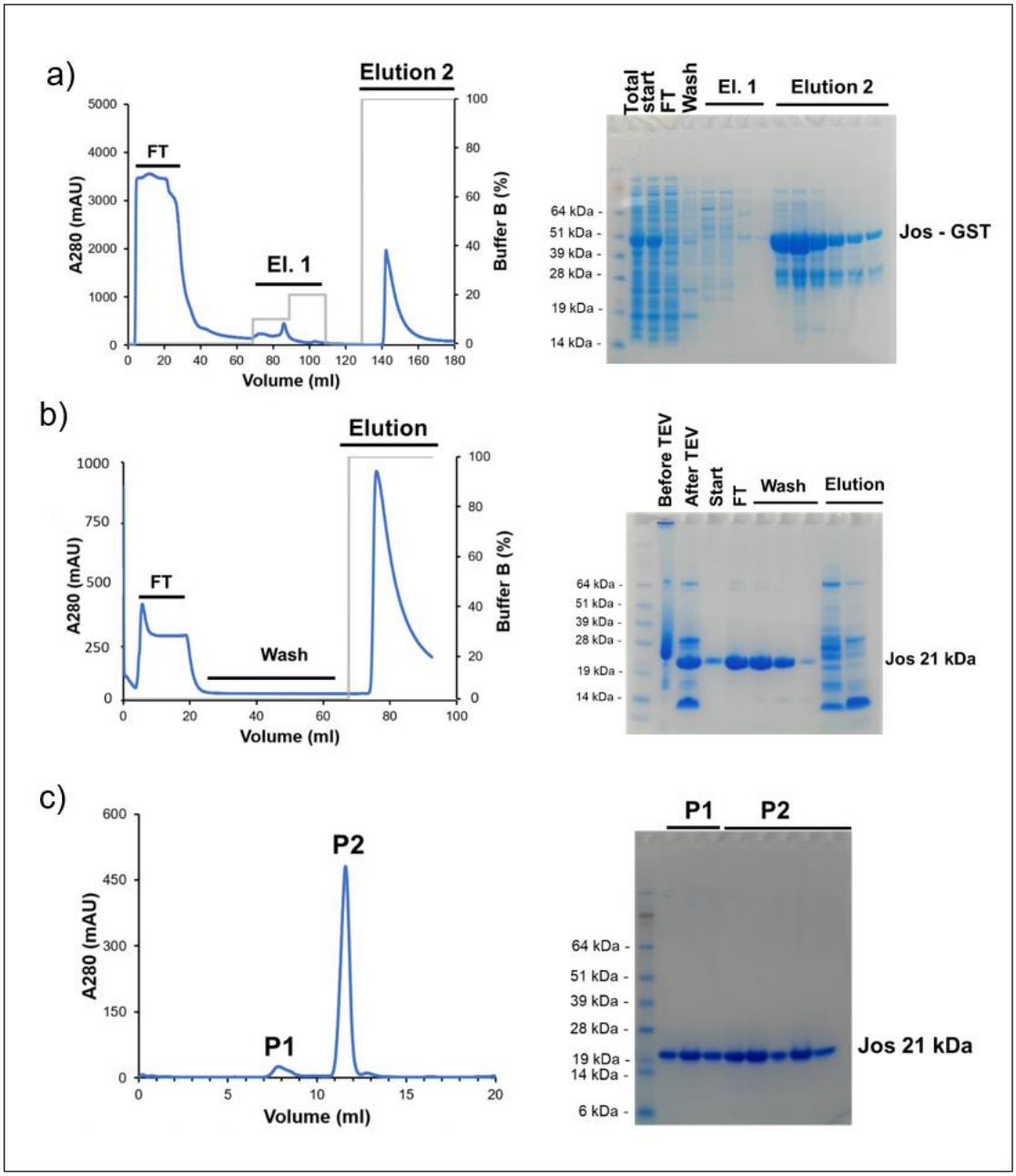
236. Kim, Y. J. *et al.* Caspase 3-cleaved N-terminal fragments of wild-type and mutant huntingtin are present in normal and Huntington's disease brains, associate with membranes, and undergo calpain-dependent proteolysis. *Proc. Natl. Acad. Sci. U. S. A.* **98**, 12784–12789 (2001).
237. Colomer Gould, V. F. *et al.* A Mutant Ataxin-3 Fragment Results From Processing at a Site N-terminal to Amino Acid 190 in Brain of Machado-Joseph Disease-Like Transgenic Mice. *Neurobiol. Dis.* **27**, 362–369 (2007).
238. Hoffner, G., Island, M.-L. & Djian, P. Purification of neuronal inclusions of patients with Huntington's disease reveals a broad range of N-terminal fragments of expanded huntingtin and insoluble polymers. *J. Neurochem.* **95**, 125–136 (2005).
239. Chen, Y.-S., Hong, Z.-X., Lin, S.-Z. & Harn, H.-J. Identifying Therapeutic Targets for Spinocerebellar Ataxia Type 3/Machado–Joseph Disease through Integration of Pathological Biomarkers and Therapeutic Strategies. *Int. J. Mol. Sci.* **21**, 3063 (2020).
240. Freeman, W. & Wszolek, Z. Botulinum toxin type A for treatment of spasticity in spinocerebellar ataxia type 3 (Machado-Joseph disease). *Mov. Disord. Off. J. Mov. Disord. Soc.* **20**, 644 (2005).
241. Diaféria, G. *et al.* Effect of speech therapy on quality of life in patients with spinocerebelar ataxia type 3. *Arq. Neuropsiquiatr.* **80**, 1017–1025 (2022).
242. Alves, S. *et al.* Allele-specific RNA silencing of mutant ataxin-3 mediates neuroprotection in a rat model of Machado-Joseph disease. *PloS One* **3**, e3341 (2008).
243. Menzies, F. M. *et al.* Puromycin-sensitive aminopeptidase protects against aggregation-prone proteins via autophagy. *Hum. Mol. Genet.* **19**, 4573–4586 (2010).
244. Jung, J., Xu, K., Lessing, D. & Bonini, N. M. Preventing Ataxin-3 protein cleavage mitigates degeneration in a Drosophila model of SCA3. *Hum. Mol. Genet.* **18**, 4843–4852 (2009).
245. Tarlac, V. & Storey, E. Role of proteolysis in polyglutamine disorders. *J. Neurosci. Res.* **74**, 406–416 (2003).
246. Moore, L. R. *et al.* Evaluation of Antisense Oligonucleotides Targeting ATXN3 in SCA3 Mouse Models. *Mol. Ther. Nucleic Acids* **7**, 200–210 (2017).
247. Menzies, F. M. *et al.* Calpain inhibition mediates autophagy-dependent protection against polyglutamine toxicity. *Cell Death Differ.* **22**, 433–444 (2015).
248. Watchon, M. *et al.* Calpain Inhibition Is Protective in Machado-Joseph Disease Zebrafish Due to Induction of Autophagy. *J. Neurosci. Off. J. Soc. Neurosci.* **37**, 7782–7794 (2017).
249. Fujikake, N. *et al.* Heat Shock Transcription Factor 1-activating Compounds Suppress Polyglutamine-induced Neurodegeneration through Induction of Multiple Molecular Chaperones. *J. Biol. Chem.* **283**, 26188–26197 (2008).
250. Bilen, J. & Bonini, N. M. Genome-wide screen for modifiers of ataxin-3 neurodegeneration in Drosophila. *PLoS Genet.* **3**, 1950–1964 (2007).
251. Thomas, M. *et al.* Pharmacologic and genetic inhibition of hsp90-dependent trafficking reduces aggregation and promotes degradation of the expanded glutamine androgen receptor without stress protein induction. *Hum. Mol. Genet.* **15**, 1876–1883 (2006).
252. Cushman-Nick, M., Bonini, N. M. & Shorter, J. Hsp104 Suppresses Polyglutamine-Induced Degeneration Post Onset in a Drosophila MJD/SCA3 Model. *PLOS Genet.* **9**, e1003781 (2013).
253. Figura, G., Koscianska, E. & Krzyzosiak, W. J. In Vitro Expansion of CAG, CAA, and Mixed CAG/CAA Repeats. *Int. J. Mol. Sci.* **16**, 18741–18751 (2015).
254. Takahashi, N., Sasagawa, N., Suzuki, K. & Ishiura, S. Synthesis of long trinucleotide repeats in vitro. *Neurosci. Lett.* **262**, 45–48 (1999).
255. Nicastro, G. *et al.* Assignment of the 1H, 13C, and 15N resonances of the Josephin domain of human ataxin-3. *J. Biomol. NMR* **30**, 457–458 (2004).
256. Groenning, M. Binding mode of Thioflavin T and other molecular probes in the context of amyloid fibrils-current status. *J. Chem. Biol.* **3**, 1–18 (2010).
257. Ban, T., Hamada, D., Hasegawa, K., Naiki, H. & Goto, Y. Direct observation of amyloid fibril growth monitored by thioflavin T fluorescence. *J. Biol. Chem.* **278**, 16462–16465 (2003).
258. Sabaté, R. & Saupe, S. J. Thioflavin T fluorescence anisotropy: an alternative technique for the study of amyloid aggregation. *Biochem. Biophys. Res. Commun.* **360**, 135–138 (2007).
259. Biancalana, M. & Koide, S. Molecular mechanism of Thioflavin-T binding to amyloid fibrils. *Biochim. Biophys. Acta* **1804**, 1405–1412 (2010).

260. Sulatskaya, A. I., Maskevich, A. A., Kuznetsova, I. M., Uversky, V. N. & Turoverov, K. K. Fluorescence Quantum Yield of Thioflavin T in Rigid Isotropic Solution and Incorporated into the Amyloid Fibrils. *PLoS ONE* **5**, e15385 (2010).
261. Foderà, V., Groenning M., Vetri V., Librizzi F., Spagnolo S., Conett C., Olsen L., van de Weert M., Leone M. Thioflavin T hydroxylation at basic pH and its effect on amyloid fibril detection. *J. Phys. Chem. B* **112**, 15174–15181 (2008).
262. Viegas, A., Manso, J., Nobrega, F. L. & Cabrita, E. J. Saturation-Transfer Difference (STD) NMR: A Simple and Fast Method for Ligand Screening and Characterization of Protein Binding. *J. Chem. Educ.* **88**, 990–994 (2011).
263. Krishnan, V. Ligand Screening by Saturation-Transfer Difference (STD) NMR Spectroscopy. *Curr. Anal. Chem.* **1**, 307–320 (2005).
264. Mureddu, L. & Vuister, G. W. Simple high-resolution NMR spectroscopy as a tool in molecular biology. *FEBS J.* **286**, 2035–2042 (2019).
265. Williamson, M. P. Using chemical shift perturbation to characterise ligand binding. *Prog. Nucl. Magn. Reson. Spectrosc.* **73**, 1–16 (2013).
266. Vranken, W. F. *et al.* The CCPN data model for NMR spectroscopy: development of a software pipeline. *Proteins* **59**, 687–696 (2005).
267. Pickart, C. M. Mechanisms underlying ubiquitination. *Annu. Rev. Biochem.* **70**, 503–533 (2001).
268. Pickart, C. M., Haldeman, M. T., Kasperek, E. M. & Chen, Z. Iodination of tyrosine 59 of ubiquitin selectively blocks ubiquitin's acceptor activity in diubiquitin synthesis catalyzed by E2(25K). *J. Biol. Chem.* **267**, 14418–14423 (1992).
269. Cook, W. J., Jeffrey, L. C., Carson, M., Chen, Z. & Pickart, C. M. Structure of a diubiquitin conjugate and a model for interaction with ubiquitin conjugating enzyme (E2). *J. Biol. Chem.* **267**, 16467–16471 (1992).
270. Faggiano, S., Alfano, C. & Pastore, A. The missing links to link ubiquitin: Methods for the enzymatic production of polyubiquitin chains. *Anal. Biochem.* **492**, 82–90 (2016).
271. Hershko, A. & Ciechanover, A. The ubiquitin system. *Annu. Rev. Biochem.* **67**, 425–479 (1998).
272. Incardona, M.-F. *et al.* EDNA: a framework for plugin-based applications applied to X-ray experiment online data analysis. *J. Synchrotron Radiat.* **16**, 872–879 (2009).
273. Kieffer, J. *et al.* New data analysis for BioSAXS at the ESRF. *J. Synchrotron Radiat.* **29**, 1318–1328 (2022).
274. Jackson, M. N. *et al.* An overview of selected current approaches to the characterization of aqueous inorganic clusters. *Dalton Trans. Camb. Engl. 2003* **44**, 16982–17006 (2015).
275. Data Analysis Primer | Small Angle X-ray Scattering/Diffraction. https://www-ssrl.slac.stanford.edu/smb-saxs/content/data-analysis-primer.
276. Petoukhov, M. V. *et al.* New developments in the ATSAS program package for small-angle scattering data analysis. *J. Appl. Crystallogr.* **45**, 342–350 (2012).
277. Petoukhov, M. V. & Svergun, D. I. Global Rigid Body Modeling of Macromolecular Complexes against Small-Angle Scattering Data. *Biophys. J.* **89**, 1237–1250 (2005).
278. Hirano, T. *et al.* Conformational Dynamics of Wild-type Lys-48-linked Diubiquitin in Solution. *J. Biol. Chem.* **286**, 37496–37502 (2011).
279. Nicastro, G. *et al.* Understanding the Role of the Josephin Domain in the PolyUb Binding and Cleavage Properties of Ataxin-3. *PLoS ONE* **5**, e12430 (2010).
280. Pettersen, E. F. *et al.* UCSF Chimera--a visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612 (2004).
281. Cornaciu, I. *et al.* The Automated Crystallography Pipelines at the EMBL HTX Facility in Grenoble. *J. Vis. Exp. JoVE* (2021) doi:10.3791/62491.
282. Scarff, C. A. *et al.* A tale of a tail: Structural insights into the conformational properties of the polyglutamine protein ataxin-3. *Int. J. Mass Spectrom.* **345–347**, 63–70 (2013).
283. Santambrogio, C. *et al.* The role of the central flexible region on the aggregation and conformational properties of human ataxin-3. *FEBS J.* **279**, 451–463 (2012).
284. Fujigasaki, H. *et al.* Ataxin-3 is translocated into the nucleus for the formation of intranuclear inclusions in normal and Machado-Joseph disease brains. *Exp. Neurol.* **165**, 248–256 (2000).
285. Marchal, S. *et al.* Structural instability and fibrillar aggregation of non-expanded human ataxin-3 revealed under high pressure and temperature. *J. Biol. Chem.* **278**, 31554–31563 (2003).

286. Shehi, E. *et al.* Temperature-dependent, irreversible formation of amyloid fibrils by a soluble human ataxin-3 carrying a moderately expanded polyglutamine stretch (Q36). *Biochemistry* **42**, 14626–14632 (2003).
287. Takiyama, Y., Sakoe, K., Nakano, I. & Nishizawa, M. Machado-Joseph disease: cerebellar ataxia and autonomic dysfunction in a patient with the shortest known expanded allele (56 CAG repeat units) of the MJD1 gene. *Neurology* **49**, 604–606 (1997).
288. van Alfen, N. *et al.* Intermediate CAG repeat lengths (53,54) for MJD/SCA3 are associated with an abnormal phenotype. *Ann. Neurol.* **49**, 805–807 (2001).
289. Cancel, G. *et al.* Marked Phenotypic Heterogeneity Associated with Expansion of a CAG Repeat Sequence at the Spinocerebellar Ataxia 3/Machado-Joseph Disease Locus. *Am. J. Hum. Genet.* **57**, 809–816 (1995).
290. Gu, W. *et al.* The shortest expanded allele of the MJD1 gene in a Chinese MJD kindred with autonomic dysfunction. *Eur. Neurol.* **52**, 107–111 (2004).
291. Harris, G. M., Dodelzon, K., Gong, L., Gonzalez-Alegre, P. & Paulson, H. L. Splice isoforms of the polyglutamine disease protein ataxin-3 exhibit similar enzymatic yet different aggregation properties. *PloS One* **5**, e13695 (2010).
292. Goto, J. *et al.* Machado-Joseph disease gene products carrying different carboxyl termini. *Neurosci. Res.* **28**, 373–377 (1997).
293. Usdin, K. The biological effects of simple tandem repeats: lessons from the repeat expansion diseases. *Genome Res.* **18**, 1011–1019 (2008).
294. Mirkin, S. M. DNA structures, repeat expansions and human hereditary disorders. *Curr. Opin. Struct. Biol.* **16**, 351–358 (2006).
295. Lin, Y., Hubert, L. & Wilson, J. H. TRANSCRIPTION DESTABILIZES TRIPLET REPEATS. *Mol. Carcinog.* **48**, 350–361 (2009).
296. Wang, L. *et al.* Effects of ionic strength and temperature on the aggregation and deposition of multi-walled carbon nanotubes. *J. Environ. Sci.* **51**, 248–255 (2017).
297. Mauri, P. L. *et al.* Ataxin-3 is subject to autolytic cleavage. *FEBS J.* **273**, 4277–4286 (2006).
298. Figueiredo, F. *et al.* A Robust Assay to Monitor Ataxin-3 Amyloid Fibril Assembly. *Cells* **11**, 1969 (2022).
299. Shiraki, K., Mimura, M., Nishinami, S. & Ura, T. Effect of additives on liquid droplets and aggregates of proteins. *Biophys. Rev.* **12**, 587–592 (2020).
300. Klement, K. *et al.* Effect of different salt ions on the propensity of aggregation and on the structure of Alzheimer's abeta(1-40) amyloid fibrils. *J. Mol. Biol.* **373**, 1321–1333 (2007).
301. Wang, W., Nema, S. & Teagarden, D. Protein aggregation—Pathways and influencing factors. *Int. J. Pharm.* **390**, 89–99 (2010).
302. Mikalauskaite, K., Ziaunys, M., Sneideris, T. & Smirnovas, V. Effect of Ionic Strength on Thioflavin-T Affinity to Amyloid Fibrils and Its Fluorescence Intensity. *Int. J. Mol. Sci.* **21**, 8916 (2020).
303. Silva, A., de Almeida, A. V. & Macedo-Ribeiro, S. Polyglutamine expansion diseases: More than simple repeats. *J. Struct. Biol.* **201**, 139–154 (2018).
304. Mthembu, S. N., Sharma, A., Albericio, F. & de la Torre, B. G. Breaking a Couple: Disulfide Reducing Agents. *Chembiochem Eur. J. Chem. Biol.* **21**, 1947–1954 (2020).
305. Stevens, R., Stevens, L. & Price, N. The stabilities of various thiol compounds used in protein purifications. *Biochem. Educ.* **11**, 70 (1983).
306. Bach, R. D., Dmitrenko, O. & Thorpe, C. Mechanism of thiolate-disulfide interchange reactions in biochemistry. *J. Org. Chem.* **73**, 12–21 (2008).
307. Benesch, R. E. & Benesch, R. The Acid Strength of the -SH Group in Cysteine and Related Compounds. *J. Am. Chem. Soc.* **77**, 5877–5881 (1955).
308. Freter, R., Pohl, E. R., Wilson, J. M. & Hupe, D. J. Role of the central thiol in determining rates of the thiol-disulfide interchange reaction. *J. Org. Chem.* **44**, 1771–1774 (1979).
309. Overman, L. E., Matzinger, D., O'Connor, E. M. & Overman, J. D. Nucleophilic cleavage of the sulfur-sulfur bond by phosphorus nucleophiles. Kinetic study of the reduction of aryl disulfides with triphenylphosphine and water. *J. Am. Chem. Soc.* **96**, 6081–6089 (1974).
310. Overman, L. E. & Petty, S. T. Nucleophilic cleavage of the sulfur-sulfur bond by phosphorus nucleophiles. III. Kinetic study of the reduction of a series of ethyl aryl disulfides with triphenylphosphine and water. *J. Org. Chem.* **40**, 2779–2782 (1975).

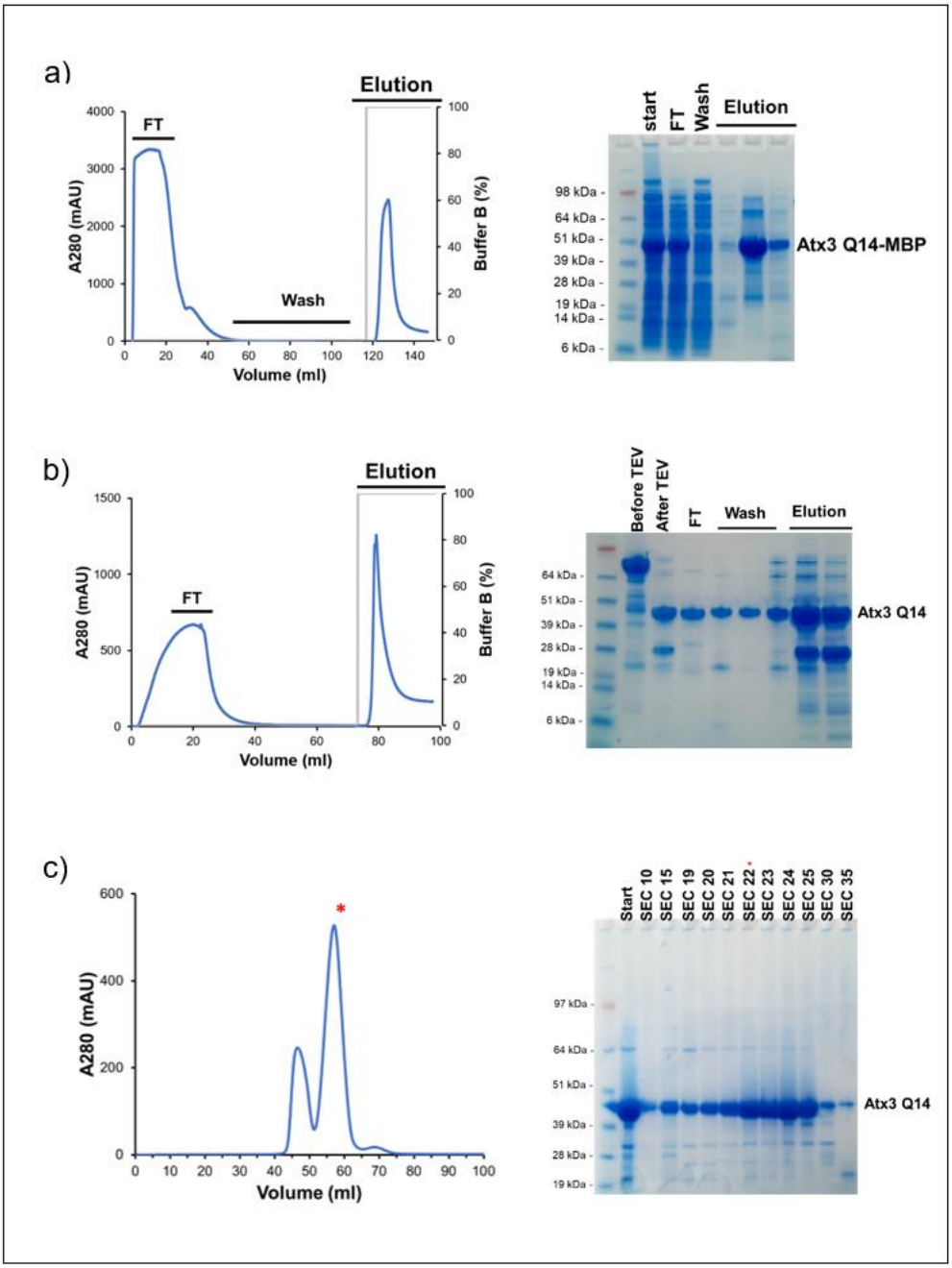
311. Overman, L. E. & O'Connor, E. M. Nucleophilic cleavage of the sulfur-sulfur bond by phosphorus nucleophiles. IV. Kinetic study of the reduction of alkyl disulfides with triphenylphosphine and water. *J. Am. Chem. Soc.* **98**, 771–775 (1976).
312. Lee, S. & Eisenberg, D. Seeded conversion of recombinant prion protein to a disulfide-bonded oligomer by a reduction-oxidation process. *Nat. Struct. Biol.* **10**, 725–730 (2003).
313. Kim, D. *et al.* Identification of disulfide cross-linked tau dimer responsible for tau propagation. *Sci. Rep.* **5**, 15231 (2015).
314. Furukawa, Y., Fu, R., Deng, H.-X., Siddique, T. & O'Halloran, T. V. Disulfide cross-linked protein represents a significant fraction of ALS-associated Cu, Zn-superoxide dismutase aggregates in spinal cords of model mice. *Proc. Natl. Acad. Sci. U. S. A.* **103**, 7148–7153 (2006).
315. Karch, C. M., Prudencio, M., Winkler, D. D., Hart, P. J. & Borchelt, D. R. Role of mutant SOD1 disulfide oxidation and aggregation in the pathogenesis of familial ALS. *Proc. Natl. Acad. Sci. U. S. A.* **106**, 7774–7779 (2009).
316. Rabdano, S. O. *et al.* Onset of disorder and protein aggregation due to oxidation-induced intermolecular disulfide bonds: case study of RRM2 domain from TDP-43. *Sci. Rep.* **7**, 11161 (2017).
317. Smith, B. N. *et al.* Mutations in the vesicular trafficking protein annexin A11 are associated with amyotrophic lateral sclerosis. *Sci. Transl. Med.* **9**, eaad9157 (2017).
318. Lukavsky, P. J. *et al.* Molecular basis of UG-rich RNA recognition by the human splicing factor TDP-43. *Nat. Struct. Mol. Biol.* **20**, 1443–1449 (2013).
319. Zacco, E. *et al.* RNA as a key factor in driving or preventing self-assembly of the TAR DNA-binding protein 43. *J. Mol. Biol.* **431**, 1671–1688 (2019).
320. Zhang, Y.-J. *et al.* The dual functions of the extreme N-terminus of TDP-43 in regulating its biological activity and inclusion formation. *Hum. Mol. Genet.* **22**, 3112–3122 (2013).
321. Kuo, P.-H., Doudeva, L. G., Wang, Y.-T., Shen, C.-K. J. & Yuan, H. S. Structural insights into TDP-43 in nucleic-acid binding and domain interactions. *Nucleic Acids Res.* **37**, 1799–1808 (2009).
322. Steinacker, P., Barschke, P. & Otto, M. Biomarkers for diseases with TDP-43 pathology. *Mol. Cell. Neurosci.* **97**, 43–59 (2019).
323. Jo, M. *et al.* The role of TDP-43 propagation in neurodegenerative diseases: integrating insights from clinical and experimental studies. *Exp. Mol. Med.* **52**, 1652–1662 (2020).
324. Alonso, A. C., Grundke-Iqbal, I. & Iqbal, K. Alzheimer's disease hyperphosphorylated tau sequesters normal tau into tangles of filaments and disassembles microtubules. *Nat. Med.* **2**, 783–787 (1996).
325. Zhang, B. *et al.* Microtubule-binding drugs offset tau sequestration by stabilizing microtubules and reversing fast axonal transport deficits in a tauopathy model. *Proc. Natl. Acad. Sci. U. S. A.* **102**, 227–231 (2005).
326. Jeganathan, S., von Bergen, M., Brutlach, H., Steinhoff, H.-J. & Mandelkow, E. Global hairpin folding of tau in solution. *Biochemistry* **45**, 2283–2293 (2006).
327. von Bergen, M. *et al.* Assembly of tau protein into Alzheimer paired helical filaments depends on a local sequence motif ((306)VQIVYK(311)) forming beta structure. *Proc. Natl. Acad. Sci. U. S. A.* **97**, 5129–5134 (2000).
328. Xia, Y. *et al.* Impaired tau-microtubule interactions are prevalent among pathogenic tau variants arising from missense mutations. *J. Biol. Chem.* **294**, 18488–18503 (2019).
329. Hutin, S., Brennich, M., Maillot, B. & Round, A. Online ion-exchange chromatography for small-angle X-ray scattering. *Acta Crystallogr. Sect. Struct. Biol.* **72**, 1090–1099 (2016).
330. Tully, M. D., Tarbouriech, N., Rambo, R. P. & Hutin, S. Analysis of SEC-SAXS data via EFA deconvolution and Scatter. *J. Vis. Exp. JoVE* (2021) doi:10.3791/61578.
331. Brennich, M. *et al.* Online data analysis at the ESRF bioSAXS beamline, BM29. *J. Appl. Crystallogr.* **49**, (2016).
332. Pernot, P. *et al.* Upgraded ESRF BM29 beamline for SAXS on macromolecules in solution. *J. Synchrotron Radiat.* **20**, 660–664 (2013).
333. Kikhney, A. G. & Svergun, D. I. A practical guide to small angle X-ray scattering (SAXS) of flexible and intrinsically disordered proteins. *FEBS Lett.* **589**, 2570–2577 (2015).
334. Mertens, H. D. T. & Svergun, D. I. Structural characterization of proteins and complexes using small-angle X-ray solution scattering. *J. Struct. Biol.* **172**, 128–141 (2010).
335. Hiranyakorn, M. *et al.* NMR Characterization of Conformational Interconversions of Lys48-Linked Ubiquitin Chains. *Int. J. Mol. Sci.* **21**, 5351 (2020).

6 Supplementary Information



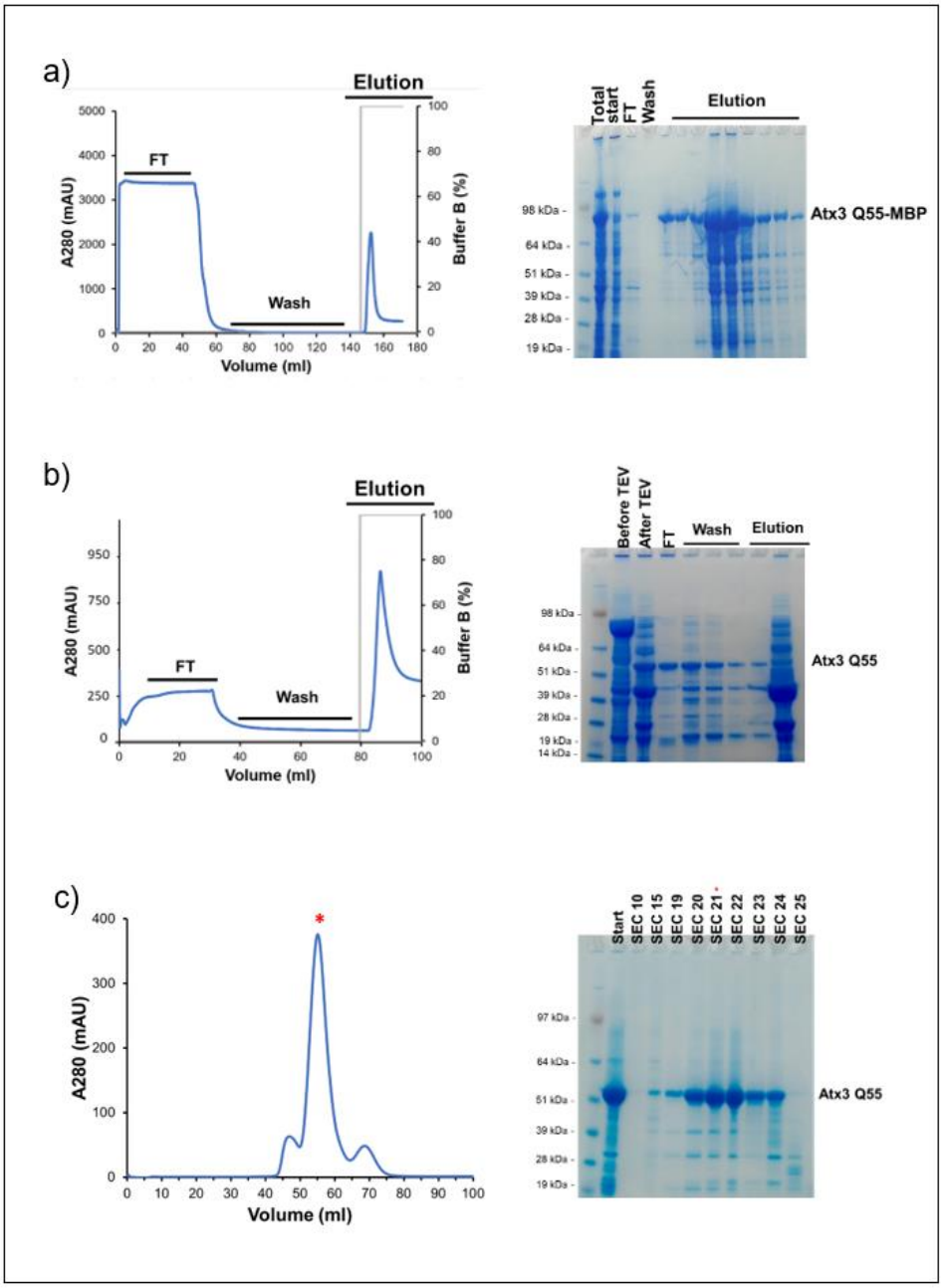
Supplementary Figure S1. Purification steps of recombinant Jos. On the left, chromatographic profile. Left axis indicates absorbance at 280 nm, while right axis indicates % elution buffer. On the right, SDS-PAGE analysis related to the fractions a) Affinity chromatography of the first IMAC step performed on cell lysate, to isolate recombinant tagged protein. Elution 1 (El.1) consists of two steo with 10% and 20% bufferB; elution 2 is a final isocratic step with 100% buffer B. b) Second IMAC step performed after tag removal. Before TEV and after TEV refers to tagged protein digestion by TEV. c) Size exclusion chromatography by Increase 10/300 75s. Peak 2 (P2, 11-12 ml) was selected for further experiments. GST-Jos: 49 kDa; Jos: 21 kDa.

Purification steps of recombinant Atx3-Q14



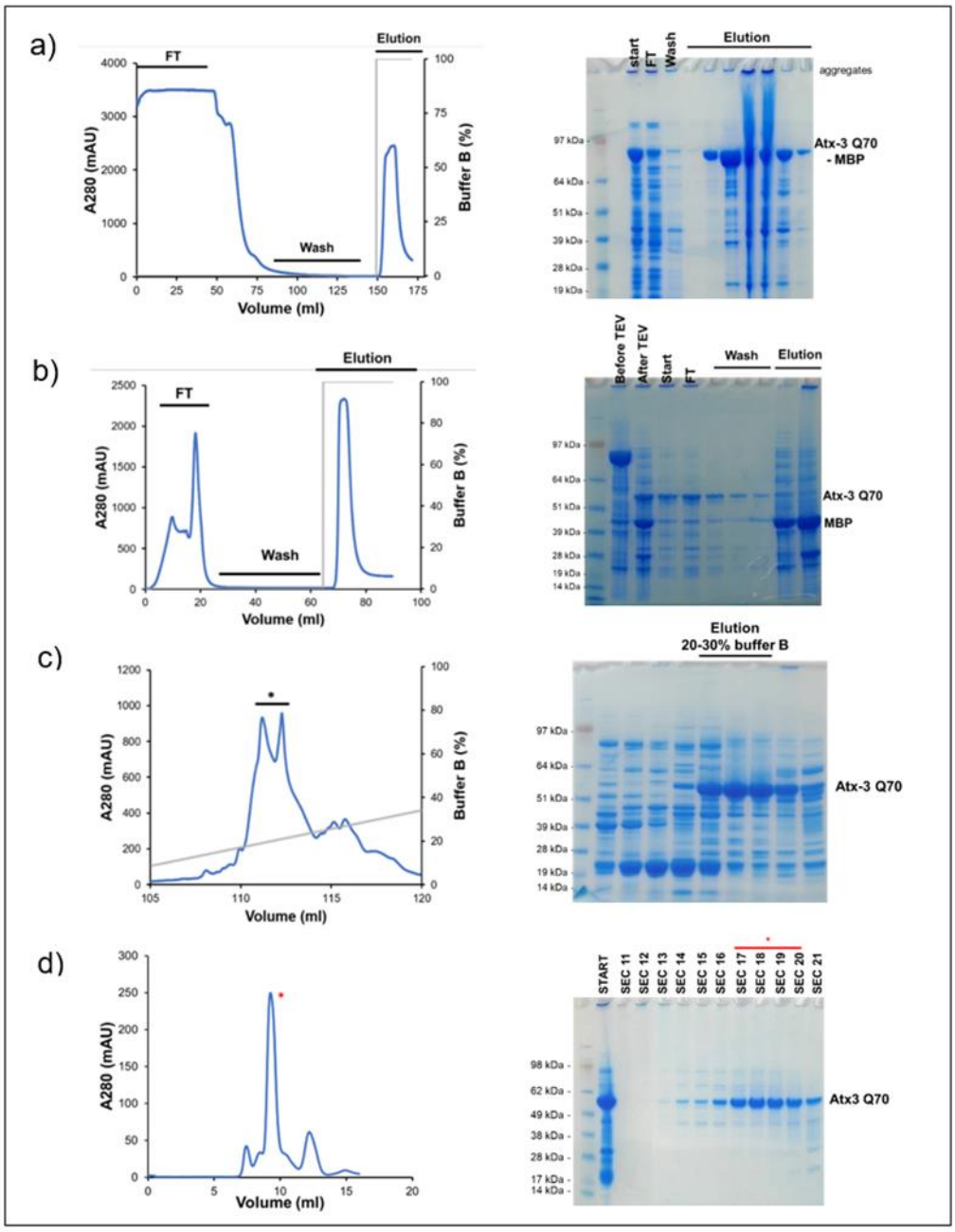
Supplementary Figure S2. Purification steps of recombinant Atx3-Q14. On the left, chromatographic profile. Left axis indicates absorbance at 280 nm, while right axis indicates % elution buffer. On the right, SDS-PAGE analysis related to the fractions a) Affinity chromatography of the first IMAC step performed on His-tagged protein. b) Second IMAC step performed after tag removal. Before TEV and after TEV refers to tagged protein digestion by TEV. c) Size exclusion chromatography by Hi load 16/600 75s. Monomeric Atx3-Q14 elutes in peak comprising 55-60 ml. MBP-Atx3-Q14: 85 kDa; Atx3-Q14: 41 kDa.

Purification steps of recombinant Atx3-Q55



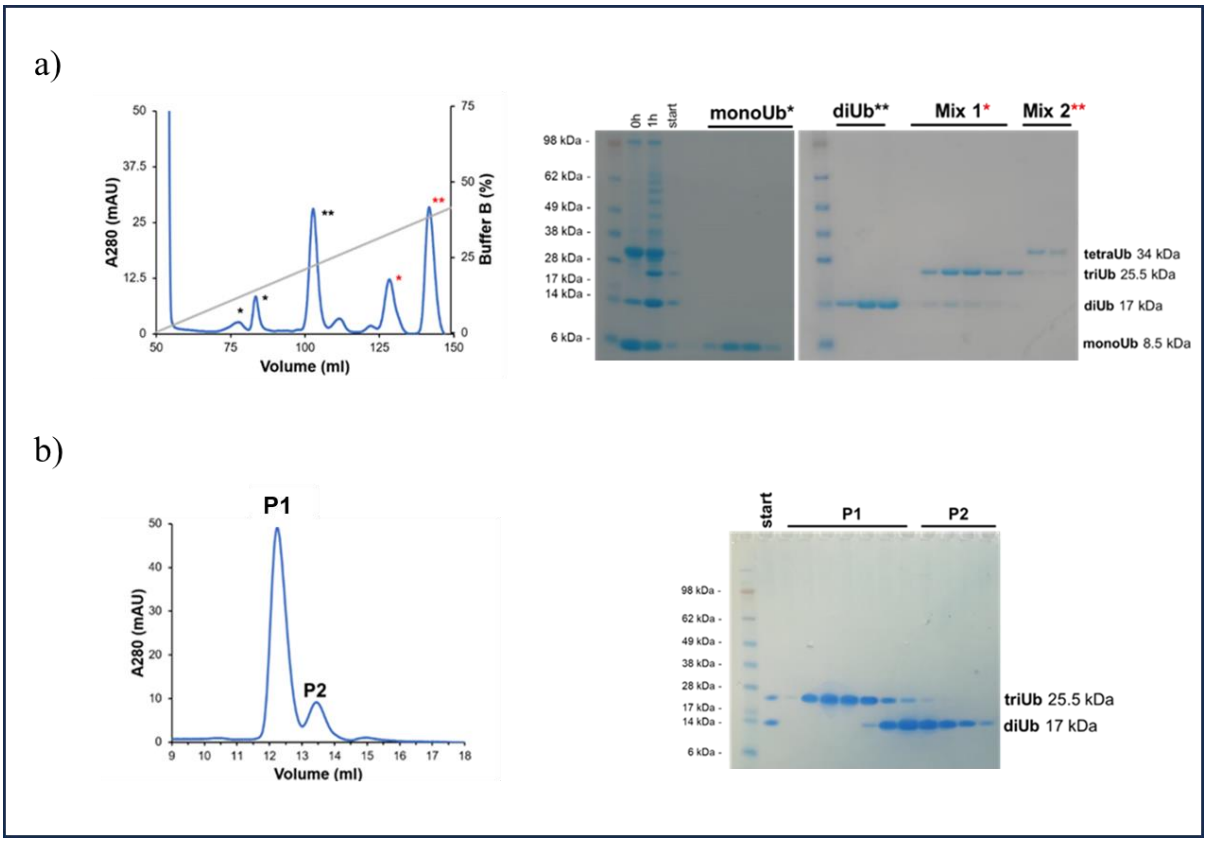
Supplementary Figure S3. Purification steps of recombinant Atx3-Q55. On the left, chromatographic profile. Left axis indicates absorbance at 280 nm, while right axis indicates % elution buffer. On the right, SDS-PAGE analysis related to the fractions a) Affinity chromatography of the first IMAC step performed on His-tagged protein. b) Second IMAC step performed after tag removal. Before TEV and after TEV refers to tagged protein digestion by TEV. c) Size exclusion chromatography by Hi load 16/600 75s. Monomeric Atx3-Q55 elutes in peak comprising 52-58 ml. MBP-Atx3-Q55: 90 kDa; Atx3-Q55: 46 kDa.

Purification steps of recombinant Atx3-Q70



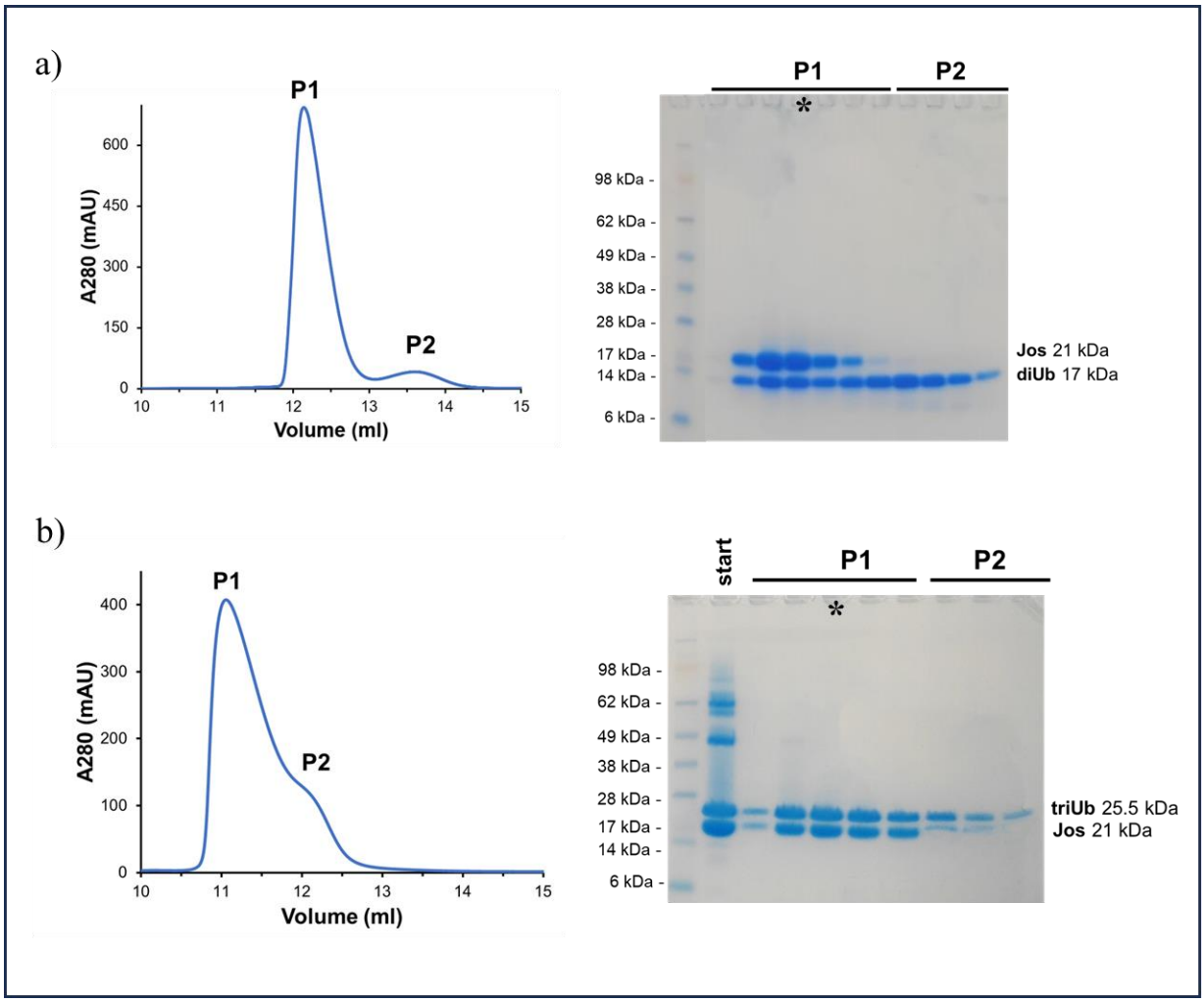
Supplementary Figure S4. Purification steps of recombinant Atx3-Q70. On the left, chromatographic profile. Left axis indicates absorbance at 280 nm, while right axis indicates % elution buffer. On the right, SDS-PAGE analysis related to the fractions. a) Affinity chromatography of the first IMAC step performed on His-tagged protein. b) Second IMAC step performed after His-tag removal. Before TEV and after TEV refers to His-tagged protein digestion by TEV to remove the His tag. c) Ionic exchange purification step by Capto Q. On the left, detail of elution step with progressively higher concentration of salt (buffer B, 1 M NaCl); Asterisk indictates the fractions selected for the last purification step. d) Size exclusion chromatography by Increase 10/300 75s. Monomeric Atx3 Q70 elutes around 10 ml. MBP-Atx3-Q70: 92 kDa; Atx3-Q70: 48 kDa.

PolyUb chains purification steps



Supplementary Figure S5. PolyUb chains purification steps. On the left, chromatographic profile; on the right, related SDS-PAGE. In chromatograms: blue curves indicate A280 values, grey curves correspond to buffer B percentage. a) Cationic exchange purification step. Even using a long gradient different polyUb chains coelute (i.e. mix 1 and mix 2) due to the the similar pI; b) polyUb clean-up by SEC. Even using SEC, part of polyUb chains with different MW still coelute. The combination of ionic exchange chromatography with a long high salt gradient allows the separation of the different polyUb species formed, that tend to co-elute because of their pI similarity. Thus, size exclusion chromatography steps were performed to better separate the different polyUb species according to different MW.

Purification of Jos-diUb and -triUb complexes



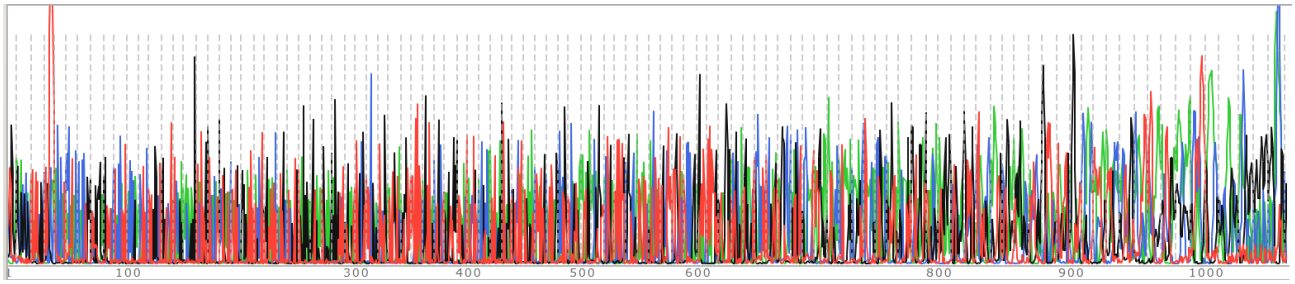
Supplementary Figure S6. Jos-diUb and -triUb complexes purification. On the left, chromatographic profile; on the right related SDS-PAGE. Both complexes were performed by Increase 10/300 75s. a) Jos-diUb complex purification. Retention volume: 12-12.5ml, MW: 38 kDa; b) Jos-triUb complex purification. Retention volume:11-11.5 ml, MW: 46.5 kDa. Asterisks marked the fraction used for aggregation assays.

Annex I

Sanger sequencing pMAL Atx3 Q70 3UIMs

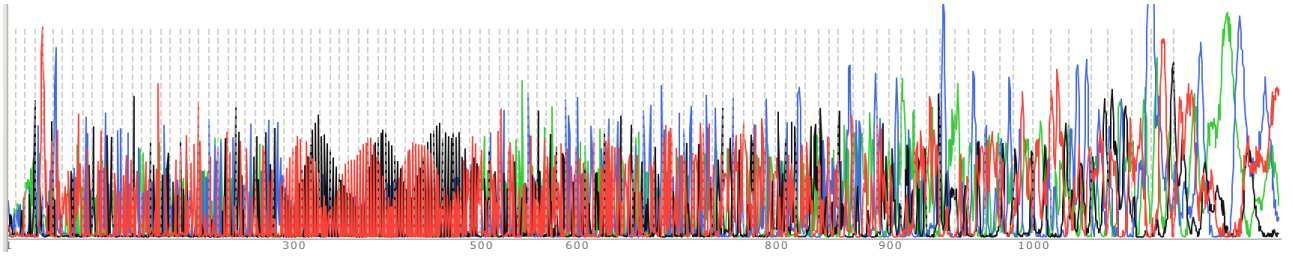
MBP-Factor X c.s. -HisTag – TEV c.s- Atx3 Q70

Sequence_malE (fwd primer)



NNNGNNCAANGCGCAGATNATTCGAGCTCGAACAACAACAACAATACACAATAACAACAACCTCGGGATCGAG
GGAAGGATTTCACATCACCATCATCACCACGAAAACCTGTACTTTCAGGGCATGGAGTCCATCTTCCACGAGA
AACAAGAAGGCTCACTTTGTGCTCAACATTGCCTGAATAACTTATTGCAAGGAGAATATTTTAGCCCTGTGGA
ATTATCCTCAATTGCACATCAGCTGGATGAGGAGGAGAGGATGAGAATGGCAGAAGGAGGAGTTACTAGTGAA
GATTATCGCACGTTTTTACAGCAGCCTTCTGGAAATATGGATGACAGTGGTTTTTTCTCTATTCAGGTTATAA
GCAATGCCTTGAAAGTTTGGGGTTTAGAACTAATCCTGTTCAACAGTCCAGAGTATCAGAGGCTCAGGATCGA
TCCTATAAATGAAAGATCATTTATATGCAATTATAAGGAACACTGGTTTACAGTTAGAAAATTAGGAAAACAG
TGGTTTAACTTGAATTCTCTCTTGACGGGTCCAGAATTAATATCAGATACATATCTTGCACTTTTCTTGGCTC
AATTACAACAGGAAGGTTATTCTATATTTGTCGTTAAGGGTGATCTGCCAGATTGCGAAGCTGACCAACTCCT
ACAGATGATTAGGGTCCAACAGATGCATCGACCAAAACTTATTGGAGAAGAATTAGCACAACTAAAAGAGCAA
AGAGTCCATAAAACAGACCTGGAAACGAGTGTTAGAAGCAAATGATGGCTCAGAATGTTAGACGAAGATGAGG
AGGACTTGCAGAGGGCTCTGGCACTAAGTCGCCAAGAAATTGACATGGGAAGATGAGGAAGCAGATCTCCGCA
GGGCTATTCAGCTAAGTATGCAAGGTAGTTCCAGGAAAACATATCTCAGGATATGACACAGACATCNNGTACA
AAATCTTACTTNAGAAAGAGCTNNGAAGAAGAACGANAANCCTACTTTGAAAAAACAGCAGCAAAGCAGNAAC
AGCAGCAGCAGCAGCAGGCAACTNGCAG

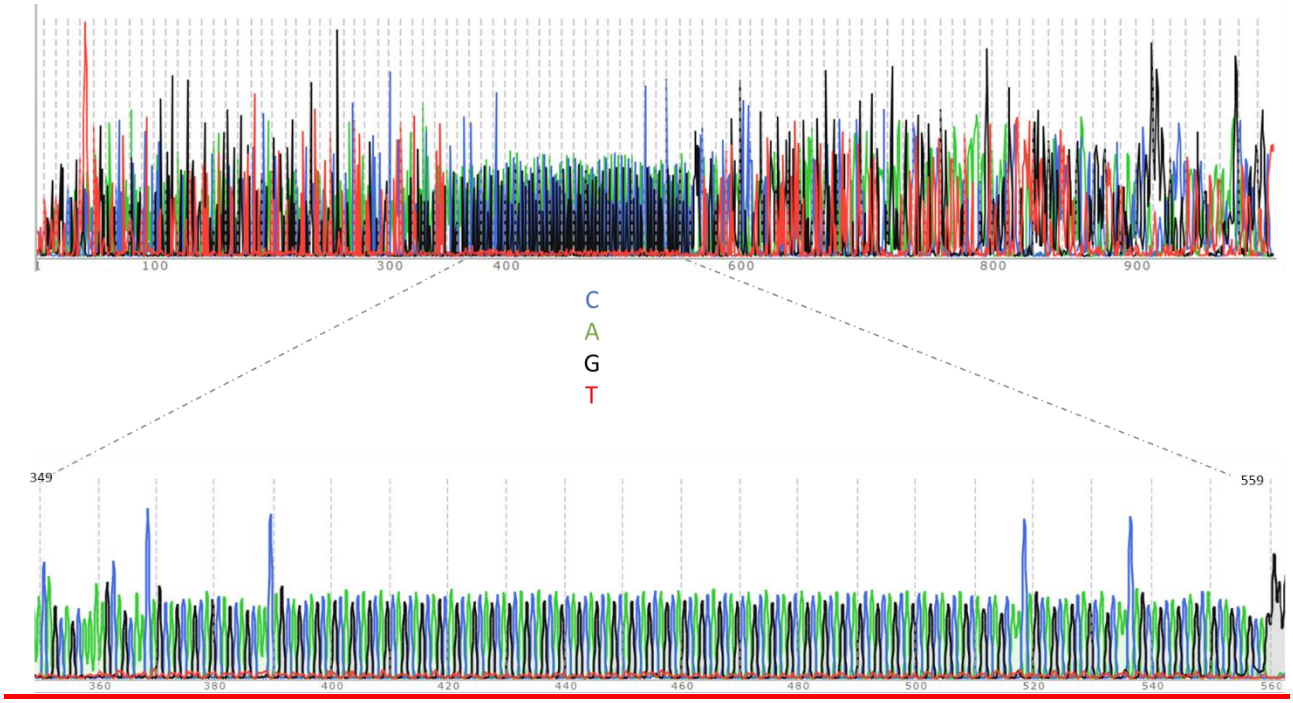
Sequence_pMalrev (rev primer)



GTNNAGCANTGCCTNNAGNNNGNNNNNNNNNTATCNNNNNNCGTCCAANNNNNNNCAGNTCGNNNNNNATNAAG
ATCNTAATGCATATAGAACNNTGGNNNCAGTTANNNTAGGNNCAGNNTTNNNTGANTNTCTCTGACGGTTCAN
NNNATATCAGATNCATATCTGCACTTTCTGGCTNATACCACCNNAAGGGTATCTATATTGTCGNTANGGTGAT
CTGCCAGATTGCGAAGCTGACCAACTCTACAGATGATTAGGGTCCAACAGATGCATCGACCCAAACTTATTGG
AGAAGAATTAGCACAACTAAAAGAGCAAAGAGTCCATAAAACAGACCTGGAACGAGTGTTAGAAGCAAATGAT

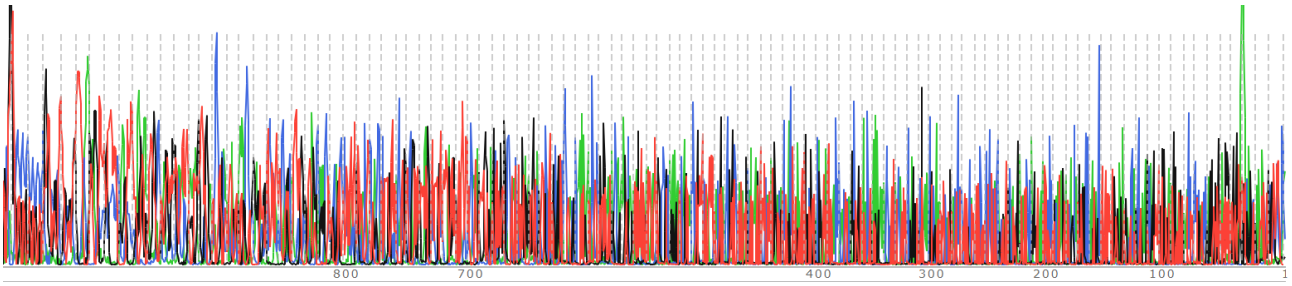
GGCTCAGGAATGTTAGACGAAGATGAGGAGGACTTGCAGAGGGCTCTGGCACTAAGTCGCCAAGAAATTGACA
TGGAAGATGAGGAAGCAGATCTCCGCAGGGCTATTCAGCTAAGTATGCAAGGTAGTTCCAGAAACATATCTCA
AGATATGACACAGACATCAGGTACAAATCTTACTTCAGAAGAGCTTCGGAAGAGACGAGAAGCCTACTTTGAA
AAACAGCAGCAAAAGCAGCAACAGCAGCAGCAGCAGCAGCAACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC
AGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCA
GCAGCAGCAGCAGCAGCAGCAGCAACAGCAGCAGCAGCAGCAACAGCAGCAGCAGCAGCAGCAGCAGGGGGAC
CTATCAGGACAGAGTTCACATCCATGTGAAAGGCCAGCCACCAGTTCAGGAGCACTTGGGAGTGATCTAGGTG
ATGCTATGAGTGAAGAAGACATGCTTCAGGCAGCTGTGACCATGTCTTTAGAAACTGTCAGAAATGATTTGAA
AACAGAAGGGAAGAAGTAGGAATTCCCTGCAGGTAATGAATTCCCTGCAGGTAATTAAATAAGCTTCAAATAA
AACGAAAGGGCTCAGTCGATATANACTGGGCCTTTCGTTTTATNCTGTTGCTTGNNG

Sequence_p170 (fwd primer)



CNNNNTTNAANNAGNATTAGGGTCCNACAGATGCATCNNNTCNAAACTTATTGGAGAAGAATTAGCACAACTA
AAAGAGCAAAGAGTCCATAAAACAGACCTGGAACGAGTGTTAGAAGCAAATGATGGCTCAGGAATGTTAGACG
AAGATGAGGAGGACTTGCAGAGGGCTCTGGCACTAAGTCGCCAAGAAATTGACATGGAAGATGAGGAAGCAGA
TCTCCGCAGGGCTATTCAGCTAAGTATGCAAGGTAGTTCCAGAAACATATCTCAAGATATGACACAGACATCA
GGTACAAATCTTACTTCAGAAGAGCTTCGGAAGAGACGAGAAGCCTACTTTGAAAAACAGCAGCAAAAGCAGC
AACAGCAGCAGCAGCAGCAGCAACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCA
GCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG
CAGCAACAGCAGCAGCAGCAGCAACAGCAGCAGCAGCAGCAGCAGGGGGACCTATCAGGACAGAGTTCAC
ATCCATGTGAAAGGCCAGCCACCAGTTCAGGAGCACTTGGGAGTGATCTAGGTGATGCTATGAGTGAAGAAGA
CATGCTTCAGGCAGCTGTGACCATGTCTTTAGAAACTGTCAGAAATGATTTGAAAACAGAAGGGAAGAAGTAG
GAATTCCCTGCAGGTAATGAATTCCCTGCAGGTAATTAAATAAGCTTCAAATAAAACGAAAGGCTCAGTCGAA
AGACTGGGCCTTTCGTTTTATCTGTTGTTTGTCGGTGAACGCTCTCCTGAGTAGGACAAATCCNNNGGAGCGG
ATTTTGAACGTTGCGAAGCAACGGTCNNNGTTGGCGGGNNNNACGCCCGCATNAACTGCAGGCATCAAATTAA
GCAGAANNNATCCTGACGGATGNNNN

Sequence_q71 (rev primer)



NNGNCNNNNNTNNNTNNCCTGNNTNANNNNNNNCATGGATNGCCANNNNNNNATATCNCCNTNGCCANAGCNN
GCNTNANNTAAAGGCGAACAGCGATNNCANNNNNNNCGTNGGGCANGNNNNNATCGACNNNAGCAAAGTNNNN
NANNNGTAACGGGTACTGCCGANNNNNGGGGTCAACCCANCCANCCGTTTCGTTGGCGTGCTGAGCGCAGGTA
TTAACGCCNCCAGTCCGAACAAAGAGCTGGCAAAAAGAGTTCCTCGAAAACTATCTGCTGACTGATGAAAGGT
CTGGAAAGCGGTTAATAAAGACAAACCCGCTGGGTGCCGTAGCGCTGAAGTCTTACGAGGAAGAGTTGGGTGA
AAGATCCGCGTATTGCCGCCACTATGGAAAACGCCCAGAAAGGTGAAATCATGCCGAACATCCCGCAGATGTC
CGCTTTCTGGTATGCCGTGCGTACTGCGGTGATCAACGCCGCCAGCGGTCGTCAGACTGTCGATGAAGCCCTG
AAAGACGCGCAGACTAATTCGAGCTCGAACAACAACAACAATAACAATAACAACAACCTCGGGATCGAGGGAA
GGATTTCACATCACCATCATCACCACGAAAACCTGTACTTTCAGGGCATGGAGTCCATCTTCCACGAGAAACA
AGAAGGCTCACTTTGTGCTCAACATTGCCTGAATAACTTATTGCAAGGAGAATATTTTAGCCCTGTGGAATTA
TCCTCAATTGCACATCAGCTGGATGAGGAGGAGAGGATGAGAATGGCAGAAGGAGGAGTTACTAGTGAAGATT
ATCGCACGTTTTTACAGCAGCCTTCTGGAAATATGGATGACAGTGGTTTTTTCTCTATTCAGGTTATAAGCAA
TGCCTTGAAAGTTTGGGGTTTAGAACTAATCCTGTTCAACAGTCCAGAGTATCAGAGGCTCAGGATCGATCCT
ATAAATGAAAGATCATTTATATGCAATTATAAGGAACACTGGTTTACAGTTAGAAAATTAGGAAAACAGTGGT
TTAACTTGAATTCTCTCTTGACGGGTCCAGAATTAATATCAGATACATATCTTGCACTTTTCTTGGNNCAATT
ACAACAGNAAGNGTTATTCTATATTTGTCGTAAGGNGTNNNGG

Merged sequence

TGAAAGATCCGCGTATTGCCGCCACTATGGAAAACGCCCAGAAAGGTGAAATCATGCCGAACATCCCGCAGAT
GTCCGCTTTCTGGTATGCCGTGCGTACTGCGGTGATCAACGCCGCCAGCGGTCGTCAGACTGTCGATGAAGCC
CTGAAAGACGCGCAGACTAATTCGAGCTCGAACAACAACAACAATAACAATAACAACAACCTCGGGATCGAGG
GAAGGATTTCACATCACCATCATCACCACGAAAACCTGTACTTTCAGGGCATGGAGTCCATCTTCCACGAGAA
ACAAGAAGGCTCACTTTGTGCTCAACATTGCCTGAATAACTTATTGCAAGGAGAATATTTTAGCCCTGTGGAA
TTATCCTCAATTGCACATCAGCTGGATGAGGAGGAGAGGATGAGAATGGCAGAAGGAGGAGTTACTAGTGAAG
ATTATCGCACGTTTTTACAGCAGCCTTCTGGAAATATGGATGACAGTGGTTTTTTCTCTATTCAGGTTATAAG
CAATGCCTTGAAAGTTTGGGGTTTAGAACTAATCCTGTTCAACAGTCCAGAGTATCAGAGGCTCAGGATCGAT
CCTATAAATGAAAGATCATTTATATGCAATTATAAGGAACACTGGTTTACAGTTAGAAAATTAGGAAAACAGT
GGTTTAACTTGAATTCTCTCTTGACGGGTCCAGAATTAATATCAGATACATATCTTGCACTTTTCTTGGCTCA
ATTACAACAGGAAGGTTATTCTATATTTGTCGTTAAGGGTGATCTGCCAGATTGCGAAGCTGACCAACTCCTA
CAGATGATTAGGGTCCAACAGATGCATCGACCAAAACTTATTGGAGAAGAATTAGCACAACTAAAAGAGCAAA
GAGTCCATAAAACAGACCTGGAACGAGTGTTAGAAGCAAATGATGGCTCAGGAATGTTAGACGAAGATGAGGA
GGACTTGCAGAGGGCTCTGGCACTAAGTCGCCAAGAAATTGACATGGAAGATGAGGAAGCAGATCTCCGCAGG
GCTATTCAGCTAAGTATGCAAGGTAGTTCCAGAAACATATCTCAAGATATGACACAGACATCAGGTACAAATC
TTACTTCAGAAGAGCTTCGGAAGAGACGAGAAGCCTACTTTGAAAAACAGCAGCAAAAGCAGCAACAGCAGCA
GCAGCAGCAGCAACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG
CAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAACAGC
AGCAGCAGCAGCAACAGCAGCAGCAGCAGCAGCAGGGGGACCTATCAGGACAGAGTTCACATCCATGTGA
AAGGCCAGCCACCAGTTCAGGAGCACTTGGGAGTGATCTAGGTGATGCTATGAGTGAAGAAGACATGCTTCAG
GCAGCTGTGACCATGTCTTTAGAAACTGTCAGAAATGATTTGAAAACAGAAGGGAAGAAGTAG

Translated merged sequence

…aaagatccgcgtattgccgccactatggaaaacgcccagaaaggtgaaatcatgccgaac
…K D P R I A A T M E N A Q K G E I M P N
atcccgcagatgtccgctttctggtatgccgtgcgtactgcggtgatcaacgccgccagc
I P Q M S A F W Y A V R T A V I N A A S
ggtcgtcagactgtcgatgaagccctgaaagacgcgcagactaattcgagctcgaacaac
G R Q T V D E A L K D A Q T N S S S N N
aacaacaataacaataacaacaacctcgggatcgagggaaggatttcacatcaccatcat
N N N N N N N N L G I E G R I S H H H H
caccacgaaaacctgtactttcagggcatggagtccatcttccacgagaaacaagaaggc
H H E N L Y F Q G M E S I F H E K Q E G
tcactttgtgctcaacattgcctgaataacttattgcaaggagaatattttagccctgtg
S L C A Q H C L N N L L Q G E Y F S P V
gaattatcctcaattgcacatcagctggatgaggaggagaggatgagaatggcagaagga
E L S S I A H Q L D E E E R M R M A E G
ggagttactagtgaagattatcgcacgtttttacagcagccttctggaaatatggatgac
G V T S E D Y R T F L Q Q P S G N M D D
agtggtttttttctctattcaggttataagcaatgccttgaaagtttggggtttagaacta
S G F F S I Q V I S N A L K V W G L E L
atcctgttcaacagtccagagtatcagaggctcaggatcgatcctataaatgaaagatca
I L F N S P E Y Q R L R I D P I N E R S
tttatatgcaattataaggaacactggtttacagttagaaaattaggaaaacagtggttt
F I C N Y K E H W F T V R K L G K Q W F
aacttgaattctctcttgacgggtccagaattaatatcagatacatatcttgcacttttc
N L N S L L T G P E L I S D T Y L A L F
ttggctcaattacaacaggaaggttattctatatttgtcgttaagggtgatctgccagat
L A Q L Q Q E G Y S I F V V K G D L P D
tgcgaagctgaccaactcctacagatgattagggtccaacagatgcatcgaccaaaactt
C E A D Q L L Q M I R V Q Q M H R P K L
attggagaagaattagcacaactaaaagagcaaagagtccataaaacagacctggaacga
I G E E L A Q L K E Q R V H K T D L E R
gtgttagaagcaaatgatggctcaggaatgttagacgaagatgaggaggacttgcagagg
V L E A N D G S G M L D E D E E D L Q R
gctctggcactaagtcgccaagaaattgacatggaagatgaggaagcagatctccgcagg
A L A L S R Q E I D M E D E E A D L R R
gctattcagctaagtatgcaaggtagttccagaaacatatctcaagatatgacacagaca
A I Q L S M Q G S S R N I S Q D M T Q T
tcaggtacaaatcttacttcagaagagcttcggaagagacgagaagcctactttgaaaaa
S G T N L T S E E L R K R R E A Y F E K
cagcagcaaaagcagcaacagcagcagcagcagcagcaacagcagcagcagcagcagcag
Q Q Q K Q Q Q Q Q Q Q Q Q Q Q Q Q Q Q Q
cagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag
Q Q Q Q Q Q Q Q Q Q Q Q Q Q Q Q Q Q Q Q
cagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcaacagcagcagcag
Q Q Q Q Q Q Q Q Q Q Q Q Q Q Q Q Q Q Q Q
cagcaacagcagcagcagcagcagcaggggacctatcaggacagagttcacatcca
Q Q Q Q Q Q Q Q Q Q G D L S G Q S S H P
tgtgaaaggccagccaccagttcaggagcacttgggagtgatctaggtgatgctatgagt
C E R P A T S S G A L G S D L G D A M S
gaagaagacatgcttcaggcagctgtgaccatgtctttagaaactgtcagaaatgatttg
E E D M L Q A A V T M S L E T V R N D L
aaaacagaagggaagaagtag
K T E G K K –

ANNEX I. Sanger sequencing results of Atx3-Q70. Atx3 Q70 construct was sequenced using malE and p170 as forward primers and pMAL and q71 as reverse primers. In purple, MBP sequence; in brown, Factor X sequence, in yellow, MBP sequence: in cyan, TEV cleavage site and in green, Atx3 coding sequence. CAG repetitions coding for glutamine (Q) are underlined in red. Merged sequencing results show in frame Atx3-Q70 construct, that is translated as expected.

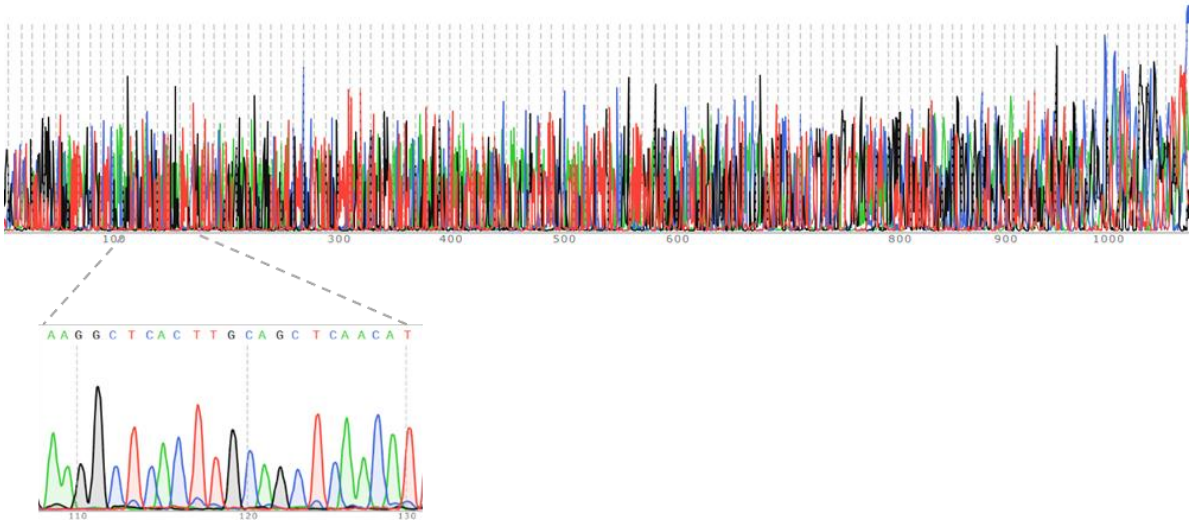
Annex II

Sanger sequencing Jos C14A

6His-GST-TEVcs-Jos1-182 C14A

Sequence_pGEX (fwd primer) in bold, GCA coding for Alanine 14

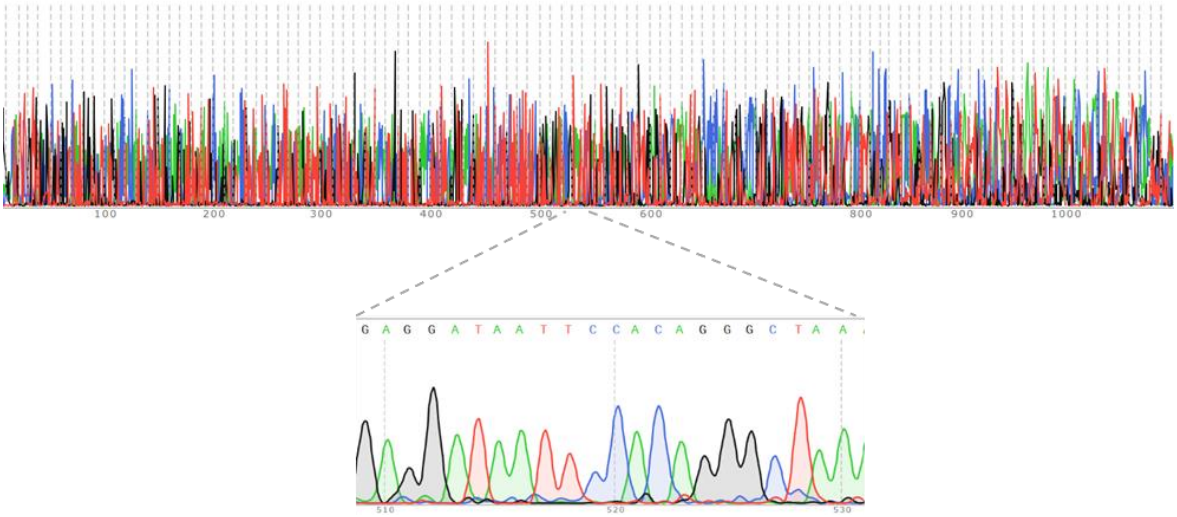
GNNNNNNNGNCNCTCCTCCACTAGTGGATCTGGTGGTGGTGGCGGATCCATGAGCGAGAATCTTTATTTTCAG
GGCGCCATGGAGTCCATCTTCCACGAGAAACAAGAAGGCTCACTT**GCA**GCTCAACATTGCCTGAATAACTTA
TTGCAAGGAGAATATTTTAGCCCTGTGGAATTATCCTCAATTGCACATCAGCTGGATGAGGAGGAGAGGATGA
GAATGGCAGAAGGAGGAGTTACTAGTGAAGATTATCGCACGTTTTTACAGCAGCCTTCTGGAAATATGGATGA
CAGTGGTTTTTTCTCTATTCAGGTTATAAGCAATGCCTTGAAAGTTTGGGGTTTAGAACTAATCCTGTTCAAC
AGTCCAGAGTATCAGAGGCTCAGGATCGATCCTATAAATGAAAGATCATTTATATGCAATTATAAGGAACACT
GGTTTACAGTTAGAAAATTAGGAAAACAGTGGTTTAACTTGAATTCTCTCTTGACGGGTCCAGAATTAATATC
AGATACATATCTTGCACTTTTCTTGGCTCAATTACAACAGGAAGGTTATTCTATATTTGTTGTTAAGGGTGAT
CTGCCAGATTGCGAAGCTGACCAACTCCTGCAGATGATTAGGTAAGCGGCCGCGAGCTCGGATCCGGCTGCTA
ACAAAGCCCGAAAGGAAGCTGAGTTGGCTGCTGCCACCGCTGAGCAATAACTAGCATAACCCCTTGGGGCCTC
TAAACGGGTCTTGAGGGGTTTTTTGCTGAAAGGAGGAACTATATCCGGATATCCACAGGACGGGTGTGGTCGC
CATGATCGCGTAGTCGATAGTGGCTCCAAGTAGCGAAGCGAGCAGGACTGGGCGGCGGGCAAAGCGGTCGGAC
AGTGCTCCGAGAACGGGTGCGCATAGAAATTGCATCAACGCATATAGCGCTAGCAGCACGCCATAGTGACTGG
CGATGCTGTCGGAATGGACGATATCCCGCAAGAGGCCCGGCAGTACCGGCATAACCAAGCTATTGCCTACAGC
ATCCAGGGGGAACGGGGCCGAGGATGGACGATGAACCGANNTGTTAGAATTTATATANN



Sequence_T7 (rev primer) in bold, GCA coding for Alanine 14

NTAGCTTGCCAAGCCCAACCATTTGGGTGGTGGCCAAAAGAGCGGCAGAGATTTCAATGTTGAAGGAGCGGTT
TGGATATTAAATGCGGTGTTCGAAAATTGCATATAGTAAAGCTTTGAAACTCTCAAGTTGATTTTCTTAGCCA
GNTACCTGAAATGCTGAAAATGTTCGAAGATCGTTATGTCATAAAACATATTTAAATGGTGATCATGTAACCC
ATCCTGACTTCATGTTGTATGACGCTCTTGATGTTGTTTTATACATGGACCCAAATGTGCCTGGATGCGTTCC
CCAAAATTAGTTTTGTTTTAAAAAACGTATTGAAGCTATCCCACAAATTGATAAGTACTTGAAATCCAGCAAG
TATATAGCATGGCCTTTGCAGGGCTGGCAAGCCACGTTTGGTGGTGGCGACCATCCTCCAACTAGTGGATCTG
GTGGTGGTGGCGGATCCATGAGCGAGAATCTTTATTTTCAGGGCGCCATGGAGTCCATCTTCCACGAGAAACA
AGAAGGCTCACTT**GCA**GCTCAACATTGCCTGAATAACTTATTGCAAGGAGAATATTTTAGCCCTGTGGAATT
ATCCTCAATTGCACATCAGCTGGATGAGGAGGAGAGGATGAGAATGGCAGAAGGAGGAGTTACTAGTGAAGAT
TATCGCACGTTTTTACAGCAGCCTTCTGGAAATATGGATGACAGTGGTTTTTTCTCTATTCAGGTTATAAGCA

ATGCCTTGAAAGTTTGGGGTTTAGAACTAATCCTGTTCAACAGTCCAGAGTATCAGAGGCTCAGGATCGATCC
TATAAATGAAAGATCATTTATATGCAATTATAAGGAACACTGGTTTACAGTTAGAAAATTAGGAAAACAGTGG
TTTAACTTGAATTCTCTCTTGACGGGTCCAGAATTAATATCAGATACATATCTTGCACTTTTCTTGGCTCAAT
TACAACAGGAAGGTTATTCTATATTTGTTGTTAAGGGTGATCTGCCAGATTGCGAAGCTGACCAACTCCTGCA
GATGATTAGGTAAGCGGCCGCGAGCTCGGATCCGGCTGCTAACAAAGCCCGAAAGGAAGCTANGNNGCNCNNN
NCC



Merged sequence

ATGAAACATCACCATCACCATCACAACACTAGTAGCAATTCCATGTCCCCTATACTAGGTTATTGGAAAATTA
AGGGCCTTGTGCAACCCACTCGACTTCTTTTGGAATATCTTGAAGAAAAATATGAAGAGCATTTGTATGAGCG
CGATGAAGGTGATAAATGGCGAAACAAAAAGTTTGAATTGGGTTTGGAGTTTCCCAATCTTCCTTATTATATT
GATGGTGATGTTAAATTAACACAGTCTATGGCCATCATACGTTATATAGCTGACAAGCACAACATGTTGGGTG
GTTGTCCAAAAGAGCGTGCAGAGATTTCAATGCTTGAAGGAGCGGTTTTGGATATTAGATACGGTGTTTCGAG
AATTGCATATAGTAAAGACTTTGAAACTCTCAAAGTTGATTTTCTTAGCAAGCTACCTGAAATGCTGAAAATG
TTCGAAGATCGTTTATGTCATAAAACATATTTAAATGGTGATCATGTAACCCATCCTGACTTCATGTTGTATG
ACGCTCTTGATGTTGTTTTATACATGGACCCAATGTGCCTGGATGCGTTCCCAAAATTAGTTTGTTTTAAAAA
ACGTATTGAAGCTATCCCACAAATTGATAAGTACTTGAAATCCAGCAAGTATATAGCATGGCCTTTGCAGGGC
TGGCAAGCCACGTTTGGTGGTGGCGACCATCCTCCAACTAGTGGATCTGGTGGTGGTGGCGGATCCATGAGCG
AGAATCTTTATTTTCAGGGCGCCATGGAGTCCATCTTCCACGAGAAACAAGAAGGCTCACTT**GCA**GCTCAAC
ATTGCCTGAATAACTTATTGCAAGGAGAATATTTTAGCCCTGTGGAATTATCCTCAATTGCACATCAGCTGGA
TGAGGAGGAGAGGATGAGAATGGCAGAAGGAGGAGTTACTAGTGAAGATTATCGCACGTTTTTACAGCAGCCT
TCTGGAAATATGGATGACAGTGGTTTTTTCTCTATTCAGGTTATAAGCAATGCCTTGAAAGTTTGGGGTTTAG
AACTAATCCTGTTCAACAGTCCAGAGTATCAGAGGCTCAGGATCGATCCTATAAATGAAAGATCATTTATATG
CAATTATAAGGAACACTGGTTTACAGTTAGAAAATTAGGAAAACAGTGGTTTAACTTGAATTCTCTCTTGACG
GGTCCAGAATTAATATCAGATACATATCTTGCACTTTTCTTGGCTCAATTACAACAGGAAGGTTATTCTATAT
TTGTTGTTAAGGGTGATCTGCCAGATTGCGAAGCTGACCAACTCCTGCAGATGATTAGGTAA

Translated sequence

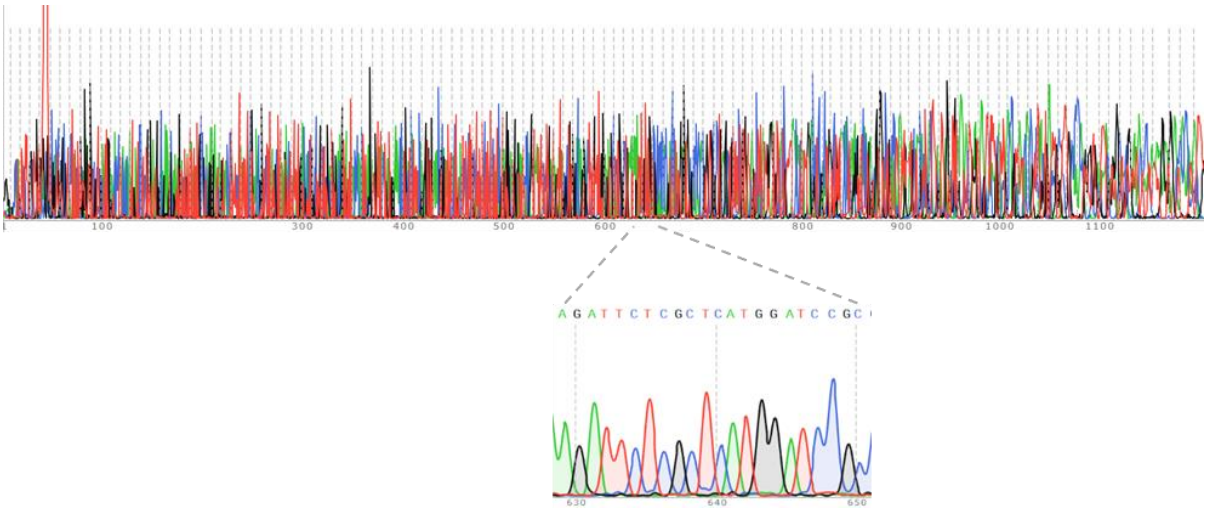
```
atgaaacatcaccatcaccatcacaacactagtagcaattccatgtcccctatactaggt
 M  K  H  H  H  H  H  H  N  T  S  S  N  S  M  S  P  I  L  G
tattggaaaattaagggccttgtgcaacccactcgacttcttttggaatatcttgaagaa
 Y  W  K  I  K  G  L  V  Q  P  T  R  L  L  L  E  Y  L  E  E
aaatatgaagagcatttgtatgagcgcgatgaaggtgataaatggcgaaacaaaaagttt
 K  Y  E  E  H  L  Y  E  R  D  E  G  D  K  W  R  N  K  K  F
gaattgggtttggagtttcccaatcttccttattatattgatggtgatgttaaattaaca
 E  L  G  L  E  F  P  N  L  P  Y  Y  I  D  G  D  V  K  L  T
cagtctatggccatcatacgttatatagctgacaagcacaacatgttgggtggttgtcca
 Q  S  M  A  I  I  R  Y  I  A  D  K  H  N  M  L  G  G  C  P
aaagagcgtgcagagatttcaatgcttgaaggagcggttttggatattagatacggtgtt
 K  E  R  A  E  I  S  M  L  E  G  A  V  L  D  I  R  Y  G  V
tcgagaattgcatatagtaaagactttgaaactctcaaagttgattttcttagcaagcta
 S  R  I  A  Y  S  K  D  F  E  T  L  K  V  D  F  L  S  K  L
cctgaaatgctgaaaatgttcgaagatcgtttatgtcataaaacatatttaaatggtgat
 P  E  M  L  K  M  F  E  D  R  L  C  H  K  T  Y  L  N  G  D
catgtaacccatcctgacttcatgttgtatgacgctcttgatgttgttttatacatggac
 H  V  T  H  P  D  F  M  L  Y  D  A  L  D  V  V  L  Y  M  D
ccaatgtgcctggatgcgttcccaaaattagtttgttttaaaaaacgtattgaagctatc
 P  M  C  L  D  A  F  P  K  L  V  C  F  K  K  R  I  E  A  I
ccacaaattgataagtacttgaaatccagcaagtatatagcatggcctttgcagggctgg
 P  Q  I  D  K  Y  L  K  S  S  K  Y  I  A  W  P  L  Q  G  W
caagccacgtttggtggtggcgaccatcctccaactagtggatctggtggtggtggcgga
 Q  A  T  F  G  G  G  D  H  P  P  T  S  G  S  G  G  G  G  G
tccatgagcgagaatctttattttcagggcgccatggagtccatcttccacgagaaacaa
 S  M  S  E  N  L  Y  F  Q  G  A  M  E  S  I  F  H  E  K  Q
gaaggctcacttgcagctcaacattgcctgaataacttattgcaaggagaatattttagc
 E  G  S  L  A  A  Q  H  C  L  N  N  L  L  Q  G  E  Y  F  S
cctgtggaattatcctcaattgcacatcagctggatgaggaggagaggatgagaatggca
 P  V  E  L  S  S  I  A  H  Q  L  D  E  E  E  R  M  R  M  A
gaaggaggagttactagtgaagattatcgcacgttttacagcagccttctggaaatatg
 E  G  G  V  T  S  E  D  Y  R  T  F  L  Q  Q  P  S  G  N  M
gatgacagtggtttttctctattcaggttataagcaatgccttgaaagtttggggttta
 D  D  S  G  F  F  S  I  Q  V  I  S  N  A  L  K  V  W  G  L
gaactaatcctgttcaacagtccagagtatcagaggctcaggatcgatcctataaatgaa
 E  L  I  L  F  N  S  P  E  Y  Q  R  L  R  I  D  P  I  N  E
agatcatttatatgcaattataaggaacactggtttacagttagaaaattaggaaaacag
 R  S  F  I  C  N  Y  K  E  H  W  F  T  V  R  K  L  G  K  Q
tggtttaacttgaattctctcttgacgggtccagaattaatatcagatacatatcttgca
 W  F  N  L  N  S  L  L  T  G  P  E  L  I  S  D  T  Y  L  A
cttttcttggctcaattacaacaggaaggttattctatatttgttgttaagggtgatctg
 L  F  L  A  Q  L  Q  Q  E  G  Y  S  I  F  V  V  K  G  D  L
ccagattgcgaagctgaccaactcctgcagatgattaggtaa
 P  D  C  E  A  D  Q  L  L  Q  M  I  R  -
```

Sanger sequencing Jos C18A

6His-GST-TEVcs-Jos1-182 C18A

Sequence_T7 (rev primer) in bold GCG coding for alanine 18

ATGTGCNGAAACAAAAAGGTTGAATGGTTTGAGGTTNCCATCTNNNNNATATTGATGGNNATGNNN
NTACCCAGTCCTATGGCCATCATCGTTATAAAGCTGANAGCCACAACATGTTGGGTGGTTGTCCCA
AAAGAGCGTGCAGAGATTTCAATGCTTGAAGGAGCGGTTTTGGATATTAGATACGGTGTTTCGAGA
ATTGCATATAGTAAAGACTTTGAAACTCTCAAAGTTGATTTTCTTAGCAAGCTACCTGAAATGCTG
AAAATGTTCGAAGATCGTTTATGTCATAAAACATATTTAAATGGTGATCATGTAACCCATCCTGAC
TTCATGTTGTATGACGCTCTTGATGTTGTTTTATACATGGACCCAATGTGCCTGGATGCGTTCCCA
AAATTAGTTTGTTTTAAAAAACGTATTGAAGCTATCCCACAAATTGATAAGTACTTGAAATCCAGC
AAGTATATAGCATGGCCTTTGCAGGGCTGGCAAGCCACGTTTGGTGGTGGCGACCATCCTCCAACT
AGTGGATCTGGTGGTGGTGGCGGATCCATGAGCGAGAATCTTTATTTTCAGGGCGCCATGGAGTCC
ATCTTCCACGAGAAACAAGAAGGCTCACTTTGTGCTCAACATGCGCTGAATAACTTATTGCAAGGA
GAATATTTTAGCCCTGTGGAATTATCCTCAATTGCACATCAGCTGGATGAGGAGGAGAGGATGAGA
ATGGCAGAAGGAGGAGTTACTAGTGAAGATTATCGCACGTTTTTACAGCAGCCTTCTGGAAATATG
GATGACAGTGGTTTTTTCTCTATTCAGGTTATAAGCAATGCCTTGAAAGTTTGGGGTTTAGAACTA
ATCCTGTTCAACAGTCCAGAGTATCAGAGGCTCAGGATCGATCCTATAAATGAAAGATCATTTATA
TGCAATTATAAGGAACACTGGTTTACAGTTAGAAAATTAGGAAAACAGTGGTTTAACTTGAATTCT
CTCTTGACGGGTCCAGAATTAATATCAGATACATATCTTGCACTTTTCTTGGCTCAATTACAACAG
GAAGGTTATTCTATATTTGTTGTTAAGGGTGATCTGCCAGATTGCGAAGCTGACCAACTCCTGCAG
ATGATTAGGTAAGCGGCCGCGAGCTCGGATCCNGCNGCTAACAAAGCCCGAAAGGAAGGCTANGTN
TTGCNNCNT



Merged sequence

:::::::ATGCTTGAAGGAGCGGTTTTGGATATTAGATACGGTGTTTCGAGAATTGCATATAGTAA
AGACTTTGAAACTCTCAAAGTTGATTTTCTTAGCAAGCTACCTGAAATGCTGAAAATGTTCGAAGA
TCGTTTATGTCATAAAACATATTTAAATGGTGATCATGTAACCCATCCTGACTTCATGTTGTATGA
CGCTCTTGATGTTGTTTTATACATGGACCCAATGTGCCTGGATGCGTTCCCAAAATTAGTTTGTTT
TAAAAAACGTATTGAAGCTATCCCACAAATTGATAAGTACTTGAAATCCAGCAAGTATATAGCATG
GCCTTTGCAGGGCTGGCAAGCCACGTTTGGTGGTGGCGACCATCCTCCAACTAGTGGATCTGGTGG

TGGTGGCGGATCCATGAGCGAGAATCTTTATTTTCAGGGCGCCATGGAGTCCATCTTCCACGAGAA
ACAAGAAGGCTCACTTTGTGCTCAACATGCGCTGAATAACTTATTGCAAGGAGAATATTTTAGCCC
TGTGGAATTATCCTCAATTGCACATCAGCTGGATGAGGAGGAGAGGATGAGAATGGCAGAAGGAGG
AGTTACTAGTGAAGATTATCGCACGTTTTTACAGCAGCCTTCTGGAAATATGGATGACAGTGGTTT
TTTCTCTATTCAGGTTATAAGCAATGCCTTGAAAGTTTGGGGTTTAGAACTAATCCTGTTCAACAG
TCCAGAGTATCAGAGGCTCAGGATCGATCCTATAAATGAAAGATCATTTATATGCAATTATAAGGA
ACACTGGTTTACAGTTAGAAAATTAGGAAAACAGTGGTTTAACTTGAATTCTCTCTTGACGGGTCC
AGAATTAATATCAGATACATATCTTGCACTTTTCTTGGCTCAATTACAACAGGAAGGTTATTCTAT
ATTTGTTGTTAAGGGTGATCTGCCAGATTGCGAAGCTGACCAACTCCTGCAGATGATTAGGTAA

Translated sequence

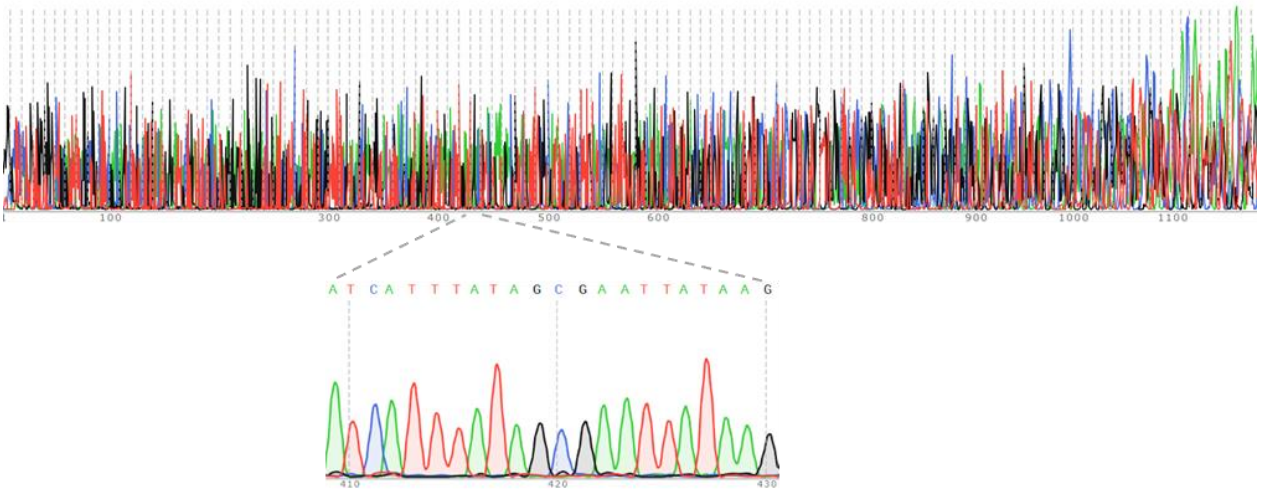
```
atgcttgaaggagcggttttggatattagatacggtgtttcgagaattgcatatagtaaa
 M  L  E  G  A  V  L  D  I  R  Y  G  V  S  R  I  A  Y  S  K
gactttgaaactctcaaagttgattttcttagcaagctacctgaaatgctgaaaatgttc
 D  F  E  T  L  K  V  D  F  L  S  K  L  P  E  M  L  K  M  F
gaagatcgtttatgtcataaaacatatttaaatggtgatcatgtaacccatcctgacttc
 E  D  R  L  C  H  K  T  Y  L  N  G  D  H  V  T  H  P  D  F
atgttgtatgacgctcttgatgttgttttatacatggacccaatgtgcctggatgcgttc
 M  L  Y  D  A  L  D  V  V  L  Y  M  D  P  M  C  L  D  A  F
ccaaaattagtttgttttaaaaaacgtattgaagctatcccacaaattgataagtacttg
 P  K  L  V  C  F  K  K  R  I  E  A  I  P  Q  I  D  K  Y  L
aaatccagcaagtatatagcatggcctttgcagggctggcaagccacgtttggtggtggc
 K  S  S  K  Y  I  A  W  P  L  Q  G  W  Q  A  T  F  G  G  G
gaccatcctccaactagtggatctggtggtggtggcggatccatgagcgagaatctttat
 D  H  P  P  T  S  G  S  G  G  G  G  G  S  M  S  E  N  L  Y
tttcagggcgccatggagtccatcttccacgagaaacaagaaggctcactttgtgctcaa
 F  Q  G  A  M  E  S  I  F  H  E  K  Q  E  G  S  L  C  A  Q
catgcgctgaataacttattgcaaggagaatattttagccctgtggaattatcctcaatt
 H  A  L  N  N  L  L  Q  G  E  Y  F  S  P  V  E  L  S  S  I
gcacatcagctggatgaggaggagaggatgagaatggcagaaggaggagttactagtgaa
 A  H  Q  L  D  E  E  E  R  M  R  M  A  E  G  G  V  T  S  E
gattatcgcacgttttacagcagccttctggaaatatggatgacagtggtttttctct
 D  Y  R  T  F  L  Q  Q  P  S  G  N  M  D  D  S  G  F  F  S
attcaggttataagcaatgccttgaaagtttggggtttagaactaatcctgttcaacagt
 I  Q  V  I  S  N  A  L  K  V  W  G  L  E  L  I  L  F  N  S
ccagagtatcagaggctcaggatcgatcctataaatgaaagatcatttatatgcaattat
 P  E  Y  Q  R  L  R  I  D  P  I  N  E  R  S  F  I  C  N  Y
aaggaacactggtttacagttagaaaattaggaaaacagtggtttaacttgaattctctc
 K  E  H  W  F  T  V  R  K  L  G  K  Q  W  F  N  L  N  S  L
ttgacgggtccagaattaatatcagatacatatcttgcacttttcttggctcaattacaa
 L  T  G  P  E  L  I  S  D  T  Y  L  A  L  F  L  A  Q  L  Q
caggaaggttattctatatttgttgttaagggtgatctgccagattgcgaagctgaccaa
 Q  E  G  Y  S  I  F  V  V  K  G  D  L  P  D  C  E  A  D  Q
ctcctgcagatgattaggtaa
 L  L  Q  M  I  R  -
```

Sanger sequencing Jos C114A

6His-GST-TEVcs-Jos1-182 C114A

Sequence 3 (fwd primer) in bold GCG coding for alanine 114

GGNNNCNGACNCTCCTCCACTAGTGGATCTGGTGGTGGTGGCGGATCNCNTGAGCGAGAATCTTTATTTTCAG
GGCGCCATGGAGTCCATCTTCCACGAGAAACAAGAAGGCTCACTTTGTGCTCAACATTGCCTGAATAACTTAT
TGCAAGGAGAATATTTTAGCCCTGTGGAATTATCCTCAATTGCACATCAGCTGGATGAGGAGGAGAGGATGAG
AATGGCAGAAGGAGGAGTTACTAGTGAAGATTATCGCACGTTTTTACAGCAGCCTTCTGGAAATATGGATGAC
AGTGGTTTTTTCTCTATTCAGGTTATAAGCAATGCCTTGAAAGTTTGGGGTTTAGAACTAATCCTGTTCAACA
GTCCAGAGTATCAGAGGCTCAGGATCGATCCTATAAATGAAAGATCATTTATA**GCG**AATTATAAGGAACACT
GGTTTACAGTTAGAAAATTAGGAAAACAGTGGTTTAACTTGAATTCTCTCTTGACGGGTCCAGAATTAATATC
AGATACATATCTTGCACTTTTCTTGGCTCAATTACAACAGGAAGGTTATTCTATATTTGTTGTTAAGGGTGAT
CTGCCAGATTGCGAAGCTGACCAACTCCTGCAGATGATTAGGTAAGCGGCCGCGAGCTCGGATCCGGCTGCTA
ACAAAGCCCGAAAGGAAGCTGAGTTGGCTGCTGCCACCGCTGAGCAATAACTAGCATAACCCCTTGGGCCTCT
AAACGGGTCTTGAGGGGTTTTTTGCTGAAAGGAGGAACTATATCCGGATATCCACAGGACGGGTGTGGTCGCC
ATGATCGCGTAGTCGATAGTGGCTCCAAGTAGCGAAGCGAGCAGGACTGGGCGGCGGCCAAAGCGGTCGGACA
GTGCTCCGAGAACGGGTGCGCATAGAAATTGCATCAACGCATATAGCGCTAGCAGCACGCCATAGTGACTGGC
GATGCTGTCGGAATGGACGATATCCCGCAAGAGGCCCGGCAGTACCGGCATAACCAAGCCTATGCCTACAGCA
TCCAGGGTGACGGTGCCGAGGATGACGATGAGCCGCATTGTAGATTTCATACACGGTGCCTGACTGCGTAGCA
ATTAACTGGGATAAACTACCGNATTAAAGCTTATCGATGATAAGCTGTCAAACATGGAAATTCTAGAAAACTT
ATCGGCATNNAAGTGAACTGCAT



Merged sequence

ATGAAACATCACCATCACCATCACAACACTAGTAGCAATTCCATGTCCCCTATACTAGGTTATTGGAAAATTA
AGGGCCTTGTGCAACCCACTCGACTTCTTTTGGAATATCTTGAAGAAAAATATGAAGAGCATTTGTATGAGCG
CGATGAAGGTGATAAATGGCGAAACAAAAAGTTTGAATTGGGTTTGGAGTTTCCCAATCTTCCTTATTATATT
GATGGTGATGTTAAATTAACACAGTCTATGGCCATCATACGTTATATAGCTGACAAGCACAACATGTTGGGTG
GTTGTCCAAAAGAGCGTGCAGAGATTTCAATGCTTGAAGGAGCGGTTTTGGATATTAGATACGGTGTTTCGAG
AATTGCATATAGTAAAGACTTTGAAACTCTCAAAGTTGATTTTCTTAGCAAGCTACCTGAAATGCTGAAAATG
TTCGAAGATCGTTTATGTCATAAAACATATTTAAATGGTGATCATGTAACCCATCCTGACTTCATGTTGTATG
ACGCTCTTGATGTTGTTTTATACATGGACCCAATGTGCCTGGATGCGTTCCCAAAATTAGTTTGTTTTAAAAA
ACGTATTGAAGCTATCCCACAAATTGATAAGTACTTGAAATCCAGCAAGTATATAGCATGGCCTTTGCAGGGC
TGGCAAGCCACGTTTGGTGGTGGCGACCATCCTCCAACTAGTGGATCTGGTGGTGGTGGCGGATCCATGAGCG

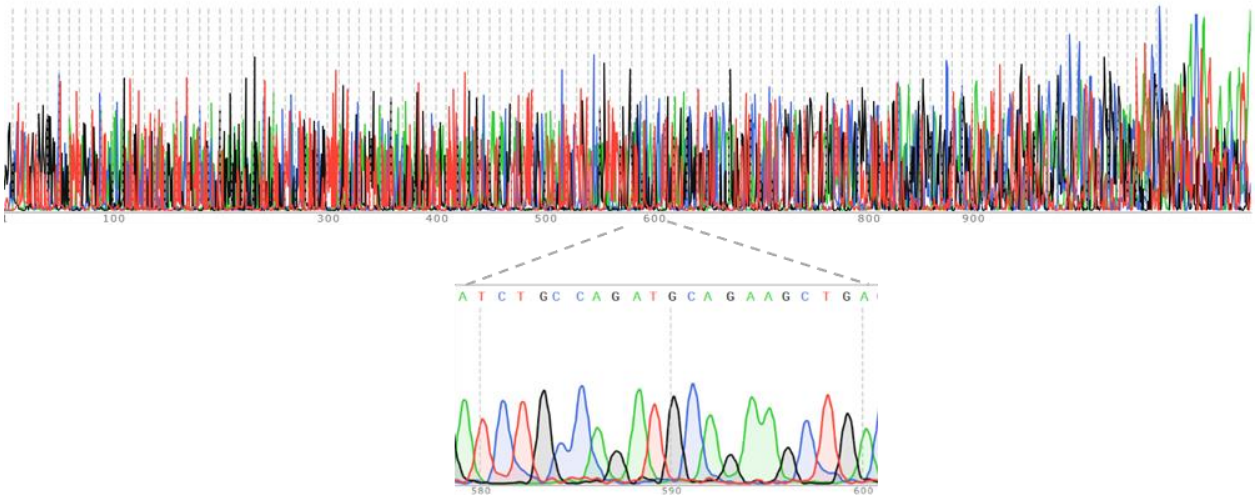
AGAATCTTTATTTTCAGGGCGCCATGGAGTCCATCTTCCACGAGAAACAAGAAGGCTCACTTTGTGCTCAACA
TTGCCTGAATAACTTATTGCAAGGAGAATATTTTAGCCCTGTGGAATTATCCTCAATTGCACATCAGCTGGAT
GAGGAGGAGAGGATGAGAATGGCAGAAGGAGGAGTTACTAGTGAAGATTATCGCACGTTTTTACAGCAGCCTT
CTGGAAATATGGATGACAGTGGTTTTTTCTCTATTCAGGTTATAAGCAATGCCTTGAAAGTTTGGGGTTTAGA
ACTAATCCTGTTCAACAGTCCAGAGTATCAGAGGCTCAGGATCGATCCTATAAATGAAAGATCATTTATA**GC**
AAATTATAAGGAACACTGGTTTACAGTTAGAAAATTAGGAAAACAGTGGTTTAACTTGAATTCTCTCTTGAC
GGGTCCAGAATTAATATCAGATACATATCTTGCACTTTTCTTGGCTCAATTACAACAGGAAGGTTATTCTATA
TTTGTTGTTAAGGGTGATCTGCCAGATTGCGAAGCTGACCAACTCCTGCAGATGATTAGGTAA

Sanger sequencing Jos C172A

6His-GST-TEVcs-Jos1-182 C172A

Sequence pGEX (fwd primer) in bold GCA coding for alanine 172

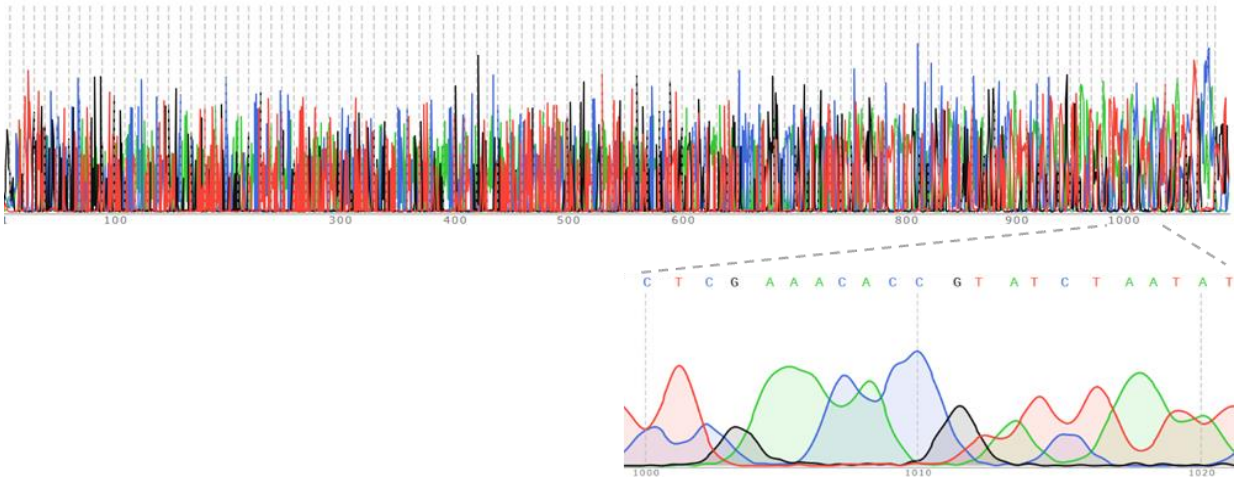
GGNNNNGCACTCCTCCACTAGTGGATCTGGTGGTGGTGGCGGATCCATGAGCGAGAATCTTTATTTTCAGGGC
GCCATGGAGTCCATCTTCCACGAGAAACAAGAAGGCTCACTTTGTGCTCAACATTGCCTGAATAACTTATTGC
AAGGAGAATATTTTAGCCCTGTGGAATTATCCTCAATTGCACATCAGCTGGATGAGGAGGAGAGGATGAGAAT
GGCAGAAGGAGGAGTTACTAGTGAAGATTATCGCACGTTTTTACAGCAGCCTTCTGGAAATATGGATGACAGT
GGTTTTTTCTCTATTCAGGTTATAAGCAATGCCTTGAAAGTTTGGGGTTTAGAACTAATCCTGTTCAACAGTC
CAGAGTATCAGAGGCTCAGGATCGATCCTATAAATGAAAGATCATTTATATGCAATTATAAGGAACACTGGTT
TACAGTTAGAAAATTAGGAAAACAGTGGTTTAACTTGAATTCTCTCTTGACGGGTCCAGAATTAATATCAGAT
ACATATCTTGCACTTTTCTTGGCTCAATTACAACAGGAAGGTTATTCTATATTTGTTGTTAAGGGTGATCTGC
CAGAT**GCA**GAAGCTGACCAACTCCTGCAGATGATTAGGTAAGCGGCCGCGAGCTCGGATCCGGCTGCTAACA
AAGCCCGAAAGGAAGCTGAGTTGGCTGCTGCCACCGCTGAGCAATAACTAGCATAACCCCTTGGGGCCTCTAA
ACGGGTCTTGAGGGGTTTTTTGCTGAAAGGAGGAACTATATCCGGATATCCACAGGACGGGTGTGGTCGCCAT
GATCGCGTAGTCGATAGTGGCTCCAAGTAGCGAAGCGAGCAGGACTGGGCGGCGGCCAAAGCGGTCGGACAGT
GCTCCGAGAACGGGTGCGCATAGAAATTGCATCACGCATATAGCGCTAGCAGCACGCCATAGTGACTGGCGAT
GCTGTCGGAATGGACGATATCCCGCAGAGGCCCGGCAGTACCGGCATAACCAAACCTATGCCTACAGCATCCA
GGGTGACGGTGCCGAGGATGGACGATGAGCCGCATTGTTAGATTTCATACACCGGCGCCTGACTGCGTTAGCG



Sequence T7 (rev primer) in bold GCA coding for alanine 172

AACAAGCANNAAACATTGTTGGGNTGGTGGTCCAAAGNGGNGGGCAGAAGATTTCAATGCTTGAAGGAGCGGT
TTTGGATATTAGATACGGTGTTTCGAGAATTGCATATAGTAAAGACTTTGAAACTCTCAAAGTTGATTTTCTT
AGCAAGCTACCTGAAATGCTGAAAATGTTCGAAGATCGTTTATGTCATAAAACATATTTAAATGGTGATCATG
TAACCCATCCTGACTTCATGTTGTATGACGCTCTTGATGTTGTTTTATACATGGACCCAATGTGCCTGGATGC
GTTCCCAAAATTAGTTTGTTTTAAAAAACGTATTGAAGCTATCCCACAAATTGATAAGTACTTGAAATCCAGC
AAGTATATAGCATGGCCTTTGCAGGGCTGGCAAGCCACGTTTGGTGGTGGCGACCATCCTCCAACTAGTGGAT
CTGGTGGTGGTGGCGGATCCATGAGCGAGAATCTTTATTTTCAGGGCGCCATGGAGTCCATCTTCCACGAGAA
ACAAGAAGGCTCACTTTGTGCTCAACATTGCCTGAATAACTTATTGCAAGGAGAATATTTTAGCCCTGTGGAA
TTATCCTCAATTGCACATCAGCTGGATGAGGAGGAGAGGATGAGAATGGCAGAAGGAGGAGTTACTAGTGAAG
ATTATCGCACGTTTTTACAGCAGCCTTCTGGAAATATGGATGACAGTGGTTTTTTCTCTATTCAGGTTATAAG

CAATGCCTTGAAAGTTTGGGGTTTAGAACTAATCCTGTTCAACAGTCCAGAGTATCAGAGGCTCAGGATCGAT
CCTATAAATGAAAGATCATTTATATGCAATTATAAGGAACACTGGTTTACAGTTAGAAAATTAGGAAAACAGT
GGTTTAACTTGAATTCTCTCTTGACGGGTCCAGAATTAATATCAGATACATATCTTGCACTTTTCTTGGCTCA
ATTACAACAGGAAGGTTATTCTATATTTGTTGTTAAGGGTGATCTGCCAGAT**GCA**GAAGCTGACCAACTCCT
GCAGATGATTAGGTAAGCGGCCGCGAGCTCGGATCCGGCTGCTAACAAAGCCCGAAAGGAAGCGANTGNNCNC
TNNNT



Merged sequence

ATGAAACATCACCATCACCATCACAACACTAGTAGCAATTCCATGTCCCCTATACTAGGTTATTGGAAAATTA
AGGGCCTTGTGCAACCCACTCGACTTCTTTTGGAATATCTTGAAGAAAAATATGAAGAGCATTTGTATGAGCG
CGATGAAGGTGATAAATGGCGAAACAAAAAGTTTGAATTGGGTTTGGAGTTTCCCAATCTTCCTTATTATATT
GATGGTGATGTTAAATTAACACAGTCTATGGCCATCATACGTTATATAGCTGACAAGCACAACATGTTGGGTG
GTTGTCCAAAAGAGCGTGCAGAGATTTCAATGCTTGAAGGAGCGGTTTTGGATATTAGATACGGTGTTTCGAG
AATTGCATATAGTAAAGACTTTGAAACTCTCAAAGTTGATTTTCTTAGCAAGCTACCTGAAATGCTGAAAATG
TTCGAAGATCGTTTATGTCATAAAACATATTTAAATGGTGATCATGTAACCCATCCTGACTTCATGTTGTATG
ACGCTCTTGATGTTGTTTTATACATGGACCCAATGTGCCTGGATGCGTTCCCAAAATTAGTTTGTTTTAAAAA
ACGTATTGAAGCTATCCCACAAATTGATAAGTACTTGAAATCCAGCAAGTATATAGCATGGCCTTTGCAGGGC
TGGCAAGCCACGTTTGGTGGTGGCGACCATCCTCCAACTAGTGGATCTGGTGGTGGTGGCGGATCCATGAGCG
AGAATCTTTATTTTCAGGGCGCCATGGAGTCCATCTTCCACGAGAAACAAGAAGGCTCACTTTGTGCTCAACA
TTGCCTGAATAACTTATTGCAAGGAGAATATTTTAGCCCTGTGGAATTATCCTCAATTGCACATCAGCTGGAT
GAGGAGGAGAGGATGAGAATGGCAGAAGGAGGAGTTACTAGTGAAGATTATCGCACGTTTTTACAGCAGCCTT
CTGGAAATATGGATGACAGTGGTTTTTTCTCTATTCAGGTTATAAGCAATGCCTTGAAAGTTTGGGGTTTAGA
ACTAATCCTGTTCAACAGTCCAGAGTATCAGAGGCTCAGGATCGATCCTATAAATGAAAGATCATTTATATGC
AATTATAAGGAACACTGGTTTACAGTTAGAAAATTAGGAAAACAGTGGTTTAACTTGAATTCTCTCTTGACGG
GTCCAGAATTAATATCAGATACATATCTTGCACTTTTCTTGGCTCAATTACAACAGGAAGGTTATTCTATATT
TGTTGTTAAGGGTGATCTGCCAGAT**GCA**GAAGCTGACCAACTCCTGCAGATGATTAGGTAA

Translated sequence

atgaaacatcaccatcaccatcacaacactagtagcaattccatgtcccctatactaggt
M K H H H H H H N T S S N S M S P I L G
tattggaaaattaagggccttgtgcaacccactcgacttcttttggaatatcttgaagaa
Y W K I K G L V Q P T R L L L E Y L E E
aaatatgaagagcatttgtatgagcgcgatgaaggtgataaatggcgaaacaaaaagttt
K Y E E H L Y E R D E G D K W R N K K F
gaattgggtttggagtttcccaatcttccttattatattgatggtgatgttaaattaaca
E L G L E F P N L P Y Y I D G D V K L T
cagtctatggccatcatacgttatatagctgacaagcacaacatgttgggtggttgtcca
Q S M A I I R Y I A D K H N M L G G C P
aaagagcgtgcagagatttcaatgcttgaaggagcggttttggatattagatacggtgtt
K E R A E I S M L E G A V L D I R Y G V
tcgagaattgcatatagtaaagactttgaaactctcaaagttgattttcttagcaagcta
S R I A Y S K D F E T L K V D F L S K L
cctgaaatgctgaaaatgttcgaagatcgtttatgtcataaaacatatttaaatggtgat
P E M L K M F E D R L C H K T Y L N G D
catgtaacccatcctgacttcatgttgtatgacgctcttgatgttgttttatacatggac
H V T H P D F M L Y D A L D V V L Y M D
ccaatgtgcctggatgcgttcccaaaattagtttgttttaaaaaacgtattgaagctatc
P M C L D A F P K L V C F K K R I E A I
ccacaaattgataagtacttgaaatccagcaagtatatagcatggcctttgcagggctgg
P Q I D K Y L K S S K Y I A W P L Q G W
caagccacgtttggtggtggcgaccatcctccaactagtggatctggtggtggtggcgga
Q A T F G G G D H P P T S G S G G G G G
tccatgagcgagaatctttattttcagggcgccatggagtccatcttccacgagaaacaa
S M S E N L Y F Q G A M E S I F H E K Q
gaaggctcactttgtgctcaacattgcctgaataacttattgcaaggagaatattttagc
E G S L C A Q H C L N N L L Q G E Y F S
cctgtggaattatcctcaattgcacatcagctggatgaggaggagaggatgagaatggca
P V E L S S I A H Q L D E E E R M R M A
gaaggaggagttactagtgaagattatcgcacgtttttacagcagccttctggaaatatg
E G G V T S E D Y R T F L Q Q P S G N M
gatgacagtggtttttttctctattcaggttataagcaatgccttgaaagtttggggttta
D D S G F F S I Q V I S N A L K V W G L
gaactaatcctgttcaacagtccagagtatcagaggctcaggatcgatcctataaatgaa
E L I L F N S P E Y Q R L R I D P I N E
agatcatttatatgcaattataaggaacactggtttacagttagaaaattaggaaaacag
R S F I C N Y K E H W F T V R K L G K Q
tggtttaacttgaattctctcttgacgggtccagaattaatatcagatacatatcttgca
W F N L N S L L T G P E L I S D T Y L A
cttttcttggctcaattacaacaggaaggttattctatatttgttgttaagggtgatctg
L F L A Q L Q Q E G Y S I F V V K G D L
ccagatgcagaagctgaccaactcctgcagatgattaggtaa
P D A E A D Q L L Q M I R –

ANNEX II. Sanger sequencing results of four point mutants of Jos. In purple, GST sequence; in yellow, His tag sequence; in cyan, TEV cleavage site and in green, Jos coding sequence. Alanine coding triplets (CGA or CGG) are reported in bold.

Annex III

human Josephin 1-182 in modified pET9d

Aminoacidic sequence HisTag-GST-TEV-human Jos 1-182

```
      10         20         30         40         50         60
HHHHHHNTSS   NSMSPILGYW   KIKGLVQPTR   LLLEYLEEKY   EEHLYERDEG   DKWRNKKFEL

         70         80         90        100        110        120
GLEFPNLPYY   IDGDVKLTQS   MAIIRYIADK   HNMLGGCPKE   RAEISMLEGA   VLDIRYGVSR

        130        140        150        160        170        180
IAYSKDFETL   KVDFLSKLPE   MLKMFEDRLC   HKTYLNGDHV   THPDFMLYDA   LDVVLYMDPM

        190        200        210        220        230        240
CLDAFPKLVC   FKKRIEAIPQ   IDKYLKSSKY   IAWPLQGWQA   TFGGGDHPPT   SGSGGGGGSM

        250        260        270        280        290        300
SENLYFQGAM   ESIFHEKQEG   SLCAQHCLNN   LLQGEYFSPV   ELSSIAHQLD   EEERMRMAEG

        310        320        330        340        350        360
GVTSEDYRTF   LQQPSGNMDD   SGFFSIQVIS   NALKVWGLEL   ILFNSPEYQR   LRIDPINERS

        370        380        390        400        410        420
FICNYKEHWF   TVRKLGKQWF   NLNSLLTGPE   LISDTYLALF   LAQLQQEGYS   IFVVKGDLPD

        430
CEADQLLQMI R
```

	HisTag-GST-TEV-human Josephin 1-182	human Josephin 1-182 after TEV cleavage
Number of amino acids	431	184
Molecular weight	49647.63	21163.96
Theoretical pI	5.41	4.67
Amino acid composition	Ala (A) 18 4.2% Arg (R) 17 3.9% Asn (N) 16 3.7% Asp (D) 26 6.0% Cys (C) 8 1.9% Gln (Q) 20 4.6% Glu (E) 34 7.9% Gly (G) 32 7.4% His (H) 16 3.7% Ile (I) 24 5.6% Leu (L) 52 12.1% Lys (K) 26 6.0% Met (M) 15 3.5% Phe (F) 21 4.9% Pro (P) 19 4.4% Ser (S) 30 7.0% Thr (T) 13 3.0% Trp (W) 7 1.6% Tyr (Y) 21 4.9% Val (V) 16 3.7% Pyl (O) 0 0.0% Sec (U) 0 0.0%	Ala (A) 8 4.3% Arg (R) 8 4.3% Asn (N) 9 4.9% Asp (D) 9 4.9% Cys (C) 4 2.2% Gln (Q) 14 7.6% Glu (E) 17 9.2% Gly (G) 12 6.5% His (H) 4 2.2% Ile (I) 11 6.0% Leu (L) 24 13.0% Lys (K) 6 3.3% Met (M) 5 2.7% Phe (F) 11 6.0% Pro (P) 6 3.3% Ser (S) 15 8.2% Thr (T) 5 2.7% Trp (W) 3 1.6% Tyr (Y) 6 3.3% Val (V) 7 3.8% Pyl (O) 0 0.0% Sec (U) 0 0.0%
Total number of negatively charged(Asp + Glu)	60	26
Total number of positively charged (Arg + Lys)	43	14
Ext. coefficient at 280 nm (assuming all pairs of Cys residues form cysteines)	70290	25690
Abs 0.1% (=1 g/l) (assuming all pairs of Cys residues form cysteines)	1.416	1.214
Ext. coefficient at 280 nm (assuming all Cys residues are reduced)	69790	25440
Abs 0.1% (=1 g/l) (assuming all Cys residues are reduced)	1.406	1.202

ANNEX III. Josephin construct sequence and Protparam

Annex IV

human Atx3 Q14 3UIMs in pMAL

Aminoacidic sequence MBP-FactorXa-HisTag -TEV-human Atx3 full-length Q14 3UIMs.

```
        10         20         30         40         50         60
MKIEEGKLVI WINGDKGYNG LAEVGKKFEK DTGIKVTVEH PDKLEEKFPQ VAATGDGPDI

        70         80         90        100        110        120
IFWAHDRFGG YAQSGLLAEI TPDKAFQDKL YPFTWDAVRY NGKLIAYPIA VEALSLIYNK

       130        140        150        160        170        180
DLLPNPPKTW EEIPALDKEL KAKGKSALMF NLQEPYFTWP LIAADGGYAF KYENGKYDIK

       190        200        210        220        230        240
DVGVDNAGAK AGLTFLVDLI KNKHMNADTD YSIAEAAFNK GETAMTINGP WAWSNIDTSK

       250        260        270        280        290        300
VNYGVTVLPT FKGQPSKPFV GVLSAGINAA SPNKELAKEF LENYLLTDEG LEAVNKDKPL

       310        320        330        340        350        360
GAVALKSYEE ELAKDPRIAA TMENAQKGEI MPNIPQMSAF WYAVRTAVIN AASGRQTVDE

       370        380        390        400        410        420
ALKDAQTNSS SNNNNNNNNN NLGIEGRISH HHHHHENLYF QGMESIFHEK QEGSLCAQHC

       430        440        450        460        470        480
LNNLLQGEYF SPVELSSIAH QLDEEERMRM AEGGVTSEDY RTFLQQPSGN MDDSGFFSIQ

       490        500        510        520        530        540
VISNALKVWG LELILFNSPE YQRLRIDPIN ERSFICNYKE HWFTVRKLGK QWFNLNSLLT

       550        560        570        580        590        600
GPELISDTYL ALFLAQLQQE GYSIFVVKGD LPDCEADQLL QMIRVQQMHR PKLIGEELAQ

       610        620        630        640        650        660
LKEQRVHKTD LERVLEANDG SGMLDEDEED LQRALALSRQ EIDMEDEEAD LRRAIQLSMQ

       670        680        690        700        710        720
GSSRNISQDM TQTSGTNLTS EELRKRREAY FEKQQQKQQQ QQQQQQGD LSGQSSHPCERP

       730        740        750        760
ATSSGALGSD LGDAMSEEDM LQAAVTMSLE TVRNDLKTEGKK
```

	MBP-FactorXa-HisTag-TEV-human Atx3 Q14 3UIMs	human Atx3 Q14 3UIMs after TEVcleavage
Number of amino acids	763	362
Molecular weight	85589.01	41307.03
Theoretical pI	4.95	4.69
Amino acid composition	Ala (A) 63 8.3% Arg (R) 27 3.5% Asn (N) 46 6.0% Asp (D) 47 6.2% Cys (C) 5 0.7% Gln (Q) 49 6.4% Glu (E) 67 8.8% Gly (G) 54 7.1% His (H) 16 2.1% Ile (I) 39 5.1% Leu (L) 75 9.8% Lys (K) 51 6.7% Met (M) 20 2.6% Phe (F) 28 3.7% Pro (P) 30 3.9% Ser (S) 47 6.2% Thr (T) 33 4.3% Trp (W) 11 1.4% Tyr (Y) 23 3.0% Val (V) 32 4.2% Pyl (O) 0 0.0% Sec (U) 0 0.0%	Ala (A) 19 5.2% Arg (R) 21 5.8% Asn (N) 13 3.6% Asp (D) 23 6.4% Cys (C) 5 1.4% Gln (Q) 39 10.8% Glu (E) 38 10.5% Gly (G) 23 6.4% His (H) 7 1.9% Ile (I) 15 4.1% Leu (L) 43 11.9% Lys (K) 15 4.1% Met (M) 13 3.6% Phe (F) 12 3.3% Pro (P) 9 2.5% Ser (S) 31 8.6% Thr (T) 14 3.9% Trp (W) 3 0.8% Tyr (Y) 7 1.9% Val (V) 12 3.3% Pyl (O) 0 0.0% Sec (U) 0 0.0%
Total number of negatively charged(Asp + Glu)	114	61
Total number of positively charged(Arg + Lys)	78	36
Ext. coefficient at 280 nm (assuming all pairs of Cys residues form cysteines)	95020	27180
Abs 0.1% (=1 g/l) (assuming all pairs of Cys residues form cysteines)	1.110	0.658
Ext. coefficient at 280 nm (assuming all Cys residues are reduced)	94770	26930

ANNEX IV. Ataxin-3 Q14 construct sequence and Protparam

Annex V

human Atx3 full length Q55 3UIMs in pMAL

Aminoacidic Sequence MBP-HisTag -TEV-human Atx3 full-length Q55 3UIMs

```
         10          20          30          40          50          60
MKIEEGKLVI WINGDKGYNG LAEVGKKFEK DTGIKVTVEH PDKLEEKFPQ VAATGDGPDI

         70          80          90         100         110         120
IFWAHDRFGG YAQSGLLAEI TPDKAFQDKL YPFTWDAVRY NGKLIAYPIA VEALSLIYNK

        130         140         150         160         170         180
DLLPNPPKTW EEIPALDKEL KAKGKSALMF NLQEPYFTWP LIAADGGYAF KYENGKYDIK

        190         200         210         220         230         240
DVGVDNAGAK AGLTFLVDLI KNKHMNADTD YSIAEAAFNK GETAMTINGP WAWSNIDTSK

        250         260         270         280         290         300
VNYGVTVLPT FKGQPSKPFV GVLSAGINAA SPNKELAKEF LENYLLTDEG LEAVNKDKPL

        310         320         330         340         350         360
GAVALKSYEE ELAKDPRIAA TMENAQKGEI MPNIPQMSAF WYAVRTAVIN AASGRQTVDE

        370         380         390         400         410         420
ALKDAQTNSS SNNNNNNNNN NLGIEGRISH HHHHHENLYF QGMESIFHEK QEGSLCAQHC

        430         440         450         460         470         480
LNNLLQGEYF SPVELSSIAH QLDEEERMRM AEGGVTSEDY RTFLQQPSGN MDDSGFFSIQ

        490         500         510         520         530         540
VISNALKVWG LELILFNSPE YQRLRIDPIN ERSFICNYKE HWFTVRKLGK QWFNLNSLLT

        550         560         570         580         590         600
GPELISDTYL ALFLAQLQQE GYSIFVVKGD LPDCEADQLL QMIRVQQMHR PKLIGEELAQ

        610         620         630         640         650         660
LKEQRVHKTD LERVLEANDG SGMLDEDEED LQRALALSRQ EIDMEDEEAD LRRAIQLSMQ

        670         680         690         700         710         720
GSSRNISQDM TQTSGTNLTS EELRKRREAY FEKQQQKQQQ QQQQQQQQQQ QQQQQQQQQQ

        730         740         750         760         770         780
QQQQQQQQQQ QQQQQQQQQQ QQQQQQQQGD LSGQSSHPCE RPATSSGALG SDLGDAMSEE

        790         800
DMLQAAVTMS LETVRNDLKT EGKK
```

	MBP-HisTag -TEV-human Atx3 full length Q55	human Atx3 Q55 After TEV cleavage
Number of amino acids	804	403
Molecular weight	90842.37	46560.45
Theoretical pI	4.95	4.69
Amino acid composition	Ala (A) 63 7.8% Arg (R) 27 3.4% Asn (N) 46 5.7% Asp (D) 47 5.8% Cys (C) 5 0.6% Gln (Q) 90 11.2% Glu (E) 67 8.3% Gly (G) 54 6.7% His (H) 16 2.0% Ile (I) 39 4.9% Leu (L) 75 9.3% Lys (K) 51 6.3% Met (M) 20 2.5% Phe (F) 28 3.5% Pro (P) 30 3.7% Ser (S) 47 5.8% Thr (T) 33 4.1% Trp (W) 11 1.4% Tyr (Y) 23 2.9% Val (V) 32 4.0% Pyl (O) 0 0.0% Sec (U) 0 0.0%	Ala (A) 19 4.7% Arg (R) 21 5.2% Asn (N) 13 3.2% Asp (D) 23 5.7% Cys (C) 5 1.2% Gln (Q) 80 19.9% Glu (E) 38 9.4% Gly (G) 23 5.7% His (H) 7 1.7% Ile (I) 15 3.7% Leu (L) 43 10.7% Lys (K) 15 3.7% Met (M) 13 3.2% Phe (F) 12 3.0% Pro (P) 9 2.2% Ser (S) 31 7.7% Thr (T) 14 3.5% Trp (W) 3 0.7% Tyr (Y) 7 1.7% Val (V) 12 3.0% Pyl (O) 0 0.0% Sec (U) 0 0.0%
Total number of negatively charged(Asp + Glu)	114	61
Total number of positively charged(Arg + Lys)	78	36
Ext. coefficient at 280 nm (assuming all pairs of Cys residues form cysteines)	95020	27180
Abs 0.1% (=1 g/l) (assuming all pairs of Cys residues form cysteines)	1.046	0.584
Ext. coefficient at 280 nm (assuming all Cys residues are reduced)	94770	26930
Abs 0.1% (=1 g/l) (assuming all Cys residues are reduced)	1.043	0.578

ANNEX V. Ataxin-3 Q55 construct sequence and Protparam

Annex VI

human Atx3 full length Q70 3UIMs in pMAL

Aminoacidic sequence MBP HisTag TEV-Atx3 Q70 full length

```
        10         20         30         40         50         60
MKIEEGKLVI WINGDKGYNG LAEVGKKFEK DTGIKVTVEH PDKLEEKFPQ VAATGDGPDI

        70         80         90        100        110        120
IFWAHDRFGG YAQSGLLAEI TPDKAFQDKL YPFTWDAVRY NGKLIAYPIA VEALSLIYNK

       130        140        150        160        170        180
DLLPNPPKTW EEIPALDKEL KAKGKSALMF NLQEPYFTWP LIAADGGYAF KYENGKYDIK

       190        200        210        220        230        240
DVGVDNAGAK AGLTFLVDLI KNKHMNADTD YSIAEAAFNK GETAMTINGP WAWSNIDTSK

       250        260        270        280        290        300
VNYGVTVLPT FKGQPSKPFV GVLSAGINAA SPNKELAKEF LENYLLTDEG LEAVNKDKPL

       310        320        330        340        350        360
GAVALKSYEE ELVKDPRIAA TMENAQKGEI MPNIPQMSAF WYAVRTAVIN AASGRQTVDE

       370        380        390        400        410        420
ALKDAQTNSS SNNNNNNNNN NLGIEGRISH HHHHHENLYF QGMESIFHEK QEGSLCAQHC

       430        440        450        460        470        480
LNNLLQGEYF SPVELSSIAH QLDEEERMRM AEGGVTSEDY RTFLQQPSGN MDDSGFFSIQ

       490        500        510        520        530        540
VISNALKVWG LELILFNSPE YQRLRIDPIN ERSFICNYKE HWFTVRKLGK QWFNLNSLLT

       550        560        570        580        590        600
GPELISDTYL ALFLAQLQQE GYSIFVVKGD LPDCEADQLL QMIRVQQMHR PKLIGEELAQ

       610        620        630        640        650        660
LKEQRVHKTD LERVLEANDG SGMLDEDEED LQRALALSRQ EIDMEDEEAD LRRAIQLSMQ

       670        680        690        700        710        720
GSSRNISQDM TQTSGTNLTS EELRKRREAY FEKQQQKQQQ QQQQQQQQQQ QQQQQQQQQQ

       730        740        750        760        770        780
QQQQQQQQQQ QQQQQQQQQQ QQQQQQQQQQ QQQQQQQQQQ QQQGDLSGQS SHPCERPATS

       790        800        810
SGALGSDLGD AMSEEDMLQA AVTMSLETVR NDLKTEGKK
```

	MBP-Factor X c.s. – HisTag – TEV c.s- Atx3 Q70	Atx3 Q70 After TEV cleavage
Number of amino acids	819	418
Molecular weight	92792.38	48482.41
Theoretical pI	4.95	4.69
Amino acid composition	Ala (A) 63 7.8% Arg (R) 27 3.4% Asn (N) 46 5.7% Asp (D) 47 5.8% Cys (C) 5 0.6% Gln (Q) 105 12.8% Glu (E) 67 8.3% Gly (G) 54 6.7% His (H) 16 2.0% Ile (I) 39 4.9% Leu (L) 75 9.3% Lys (K) 51 6.3% Met (M) 20 2.5% Phe (F) 28 3.5% Pro (P) 30 3.7% Ser (S) 47 5.8% Thr (T) 33 4.1% Trp (W) 11 1.4% Tyr (Y) 23 2.9% Val (V) 32 4.0% Pyl (O) 0 0.0% Sec (U) 0 0.0%	Ala (A) 19 4.7% Arg (R) 21 5.2% Asn (N) 13 3.2% Asp (D) 23 5.7% Cys (C) 5 1.2% Gln (Q) 95 22.7% Glu (E) 38 9.4% Gly (G) 23 5.7% His (H) 7 1.7% Ile (I) 15 3.7% Leu (L) 43 10.7% Lys (K) 15 3.7% Met (M) 13 3.2% Phe (F) 12 3.0% Pro (P) 9 2.2% Ser (S) 31 7.7% Thr (T) 14 3.5% Trp (W) 3 0.7% Tyr (Y) 7 1.7% Val (V) 12 3.0% Pyl (O) 0 0.0% Sec (U) 0 0.0%
Total number of negatively charged(Asp + Glu)	114	61
Total number of positively charged(Arg + Lys)	78	36
Ext. coefficient at 280 nm (assuming all pairs of Cys residues form cysteines)	95020	27180
Abs 0.1% (=1 g/l) (assuming all pairs of Cys residues form cysteines)	1.024	0.561
Ext. coefficient at 280 nm (assuming all Cys residues are reduced)	94770	26930
Abs 0.1% (=1 g/l) (assuming all Cys residues are reduced)	1.021	0.555

ANNEX VI. Ataxin-3 Q70 construct sequence and Protparam

Acknowledgements

I would like to express my gratitude to my supervisor Dr Caterina Alfano, whose careful supervision and expertise have guided me through these three years of not only professional but also personal growth. I also want to thank my co-tutor, Prof Valeria Vetri, for her great availability and support. I extend my thanks to my supervisor at the ESRF in Grenoble, Prof Annalisa Pastore, who has been a source of inspiration and valuable advice.

I would like to also thank the members of the Structural Biology group at the European Synchrotron Radiation Facility in Grenoble for their warm hospitality and in particular Dr Mark Tully, Dr Erika Foldesne Dudas and Dr Antonino Caliò for their careful supervision and support in SAXS measurements. A special thanks to Dr Giorgio Schirò, for his availability and his valuable suggestion on modelling.

I would also thank Martyna Podlasiak for the productive exchange of ideas and the i3S Histology and Electron Microscopy platform team for their technical support in microscopy. My thanks also go to Prof Buscarino, Dr Gulli and the ATeN Center AFM laboratory that supported me in the acquisition of AFM measurements shown in this work.

My deepest and most sincere gratitude goes to my colleagues in the Biophysics and Structural Biology group at Fondazione Ri.MED: to Elisa and Raffaele for their daily support and technical expertise, to Maria for her competent tutoring on NMR spectroscopy and to all of them for the time spent together and for the maintenance of high spirits even during the most challenging moments.

A special thanks also to all the members of the Biophysics group at the University of Palermo, always available and supportive, and particularly to Giuseppe, with whom I shared a wonderful period in the mountains of Grenoble. Thanks to all the colleagues of the XXXVI doctoral cycle, with whom I had the pleasure of sharing this journey.

Last but not least, a heartfelt *thank you* to my family, to Emmanuele, and to my lifelong friends, who have accompanied me on this journey with patience and optimism, sharing with me the joys and the challenging moments.