

Research Article

Expansion of amino acid homo-sequences in proteins: Insights into the role of amino acid homo-polymers and of the protein context in aggregation

R. P. Menon and A. Pastore*

National Institute for Medical Research, London NW7 1AA (UK), Fax: +44 20 89064477,
e-mail: apastor@nimr.mrc.ac.uk

Received 3 March 2006; received after revision 19 April 2006; accepted 22 May 2006
Online First 22 June 2006

Abstract. Expansion of amino acid homo-sequences, such as polyglutamines or polyalanines, in proteins has been directly implicated in various degenerative diseases through a mechanism of protein misfolding and aggregation. However, it is still unclear how the nature of the expansion and the protein context influence the tendency of a protein to aggregate. Here, we have addressed these questions using spinocerebellar ataxia type-3 (ATX3) protein, the best characterised of the polyglutamine pro-

teins, chosen as a model system. Using a transfected mammalian cell line, we demonstrate that ATX3 aggregation is noticeably reduced by deletion or replacement of regions other than the polyglutamine tract. The nature of the amino acid homo-sequences also has a strong influence on aggregation. From our studies, we draw general conclusions on the effect of the protein architecture and of the amino acid homo-sequence on pathology.

Keywords. Amyloidogenesis, ataxin-3, Josephin, misfolding diseases, polyglutamine expansion, protein aggregation.

Introduction

An increasing number of human genetic diseases have been shown to be associated with an anomalous expansion of low complexity repeats of amino acid sequences [1–3]. A unifying hallmark of these pathologies appears to be the formation of nuclear protein inclusions, which are highly cytotoxic and may cause pathology through a loss of function mechanism, associated with gain of abnormal functions [1, 4]. It is now widely accepted that there is a direct link between expansion and the aggregation properties of the corresponding proteins, although the mechanism by which expansion leads to misfolding and aggregation remains unclear. One of the best characterised among these human diseases is probably the family of diseases that involves expansion of polyglut-

amine (polyQ) repeats: at least eight dominant spinocerebellar ataxias (SCA) occur when the length of a polyQ tract within the coding regions of the respective proteins is above a given threshold [5]. Several other human pathologies have more recently been linked to expansion of polyalanine (polyA) repeats (reviewed in [1, 3]). Other disease-linked homo-sequences may yet be identified since long amino acid tracts of almost all types of homo-sequences have been shown to have tendency to aggregate [6].

Whether the repeat regions are involved in any specific function in the corresponding non-pathological proteins is still unknown. It has been suggested that long stretches of homo-sequences are involved in processes that necessitate the assembly of large multi-protein complexes, such as transcription and signalling [7]. An important point, which is still controversial, is the mechanism that leads from expansion to pathology: whether homo-se-

* Corresponding author.

quences are all that is needed to induce toxic aggregation and whether the rest of the protein sequence has any role in aggregate formation. Various independent reports have shown that some of the polyQ proteins tend to aggregate even in their non-expanded form and that protein context has a strong effect on the solubility and stability of polyQ stretches [8–12]. At least two of the polyQ proteins linked to pathologies have also been shown to contain other regions, distinct from the polyQ tract, that are able to aggregate and lead to amyloid formations similar to those observed in other degenerative diseases also when produced independently from the polyQ region or in their non-expanded form [9–11]. An apparently different school of thought suggests instead that polyQ expansion is sufficient to lead to pathology and that this may occur as a consequence of a proteolytic event, which frees the homo-sequence from the rest of the protein [13–17].

In an attempt to clarify this controversy and to study further the effect of poly-amino acid expansion in cell aggregation, we have used a model system in which the aggregation tendencies of different proteins and protein domains were analysed in terms of their ability to form visible aggregates when overexpressed in mammalian cells. Using this system, we have addressed two specific questions. First, we considered the relative role of the protein context in the aggregation properties of ataxin 3 (ATX3), chosen as a paradigmatic example. This is a small protein responsible for Machado-Joseph disease/spinocerebellar ataxia type 3 (MJD/SCA3) [18]. ATX3, which has been shown to be involved in the ubiquitin/proteasome pathway [19–23], contains an N-terminal Josephin domain, which spans the first 182 residues and is the only constitutively structured region of the protein, and a C-terminal region that contains the polyQ tract [24–26]. The Josephin domain was shown to be able to aggregate and form amyloid fibres also in the absence of a polyQ tract [9], thus strongly suggesting the presence of at least two distinct amyloidogenic sequences. In the present study, using different ATX3 constructs in which we have both varied the length of the polyQ tract and changed the nature of the protein context, we have assessed the relative importance of Josephin over polyQ expansion *in vivo*.

As a corollary of this study, we have used the same mammalian cell system to compare the aggregation tendencies of different amino acid homo-polymers. From our study, we show that, while expansion of the polyQ tract is necessary to trigger aggregation, the protein context has strong effects in modulating the role of polyQ and that different amino acid homo-repeats have different tendencies to aggregate.

Materials and methods

Expression constructs. ATX3 cDNA with 72 or 18 CAG repeats was PCR amplified and cloned into pBudCE4.1

vector (Invitrogen) using standard protocols to obtain V5 tagged mammalian expression constructs (ATX3_Q18 and ATX3_Q72). These were used as templates to introduce frameshift mutations so that the glutamine encoding CAG codons were changed to AGC and GCA, to code for serines (ATX3_S18 and ATX3_S72) and alanines (ATX3_A18 and ATX3_A72), respectively. The frameshifts were achieved by PCR taking advantage of the non strand-displacing action of *pfu* Turbo DNA polymerase (Stratagene) and employing 5' phosphorylated mutagenic primers incorporating the frameshift that anneals to complementary sequences on opposite strands of the plasmid. The following primers were used: 5'-GAGAAGCCTACTTTGAAAAAGCAGCAA-AAGCAGCAACAGC-3' and 5'-GCTGTTGCTGCTTT-TGCTGCTTTTTTCAAAGTAGGCTTCTC-3' for serine repeats, and 5'-GAGAAGCCTACTTTGAAAAAGCAGCAA-AAGCAGCAACAGC-3' and 5'-GCTGTTGCTGCTTTT-TGCTGCTTTTTTCAAAGTAGGCTTCTC-3' for alanine repeats. After PCR (25 cycles with annealing and extension temperatures of 55 °C and 72 °C each and 6 min extension time), the template DNA was digested with *DpnI* and 5 µL of the PCR product was directly transformed into XL-10 Gold competent cells (Stratagene). Individual clones were selected and verified by DNA sequencing. The frameshifts resulted in ATX3 constructs truncated 10 (alanine repeat) and 23 (serine repeat) residues downstream of the repeats. The 3' sequences were then corrected using a further round of mutagenesis. The following sense primers, along with their complementary antisense primers (not listed) were used: 5'-AGCAGCAGCCGGGACCTATCAGGACAGAGTTCACATCC-3' for serine constructs and 5'-GGATGTGAACTCTGTCCTGATAGGTCCTGCTGCTGCTG-3' for alanine constructs. ATX3 lacking the Josephin domain was amplified by PCR using the 5'-AAGACTGGTACCATGAAAACAGACCTGGAACGAATGTTAGAAGC-3' and 5'-TAGATCCTCGAGTCTGTCAGATAAAGTGTGAAGGTAGCGAAC-3' primers. These were cloned into pBudCE4.1 vector using *KpnI* and *XhoI* sites to generate Josephin-deleted ATX3 expression vectors (ATX3_Q18-ΔJos and ATX3_Q72-ΔJos). The AXH domain DNA sequence from HBP1 was amplified by PCR and cloned in-frame into these constructs to create HBP1 AXH replaced constructs (ATX3_Q18-HBP1 and ATX3_Q72-HBP1). Expression vectors encoding C-terminal fragments of ATX3_Q72, corresponding to putative caspase cleavage products, were generated by cloning ATX3_Q72 lacking the first 217 or first 225 amino acids into pBudCE4.1 vector (ATX3_Q72-Δ1-217 and ATX3_Q72-Δ1-225).

ATX3_Q18-GFP and ATX3_Q72-GFP were made by amplifying sequences corresponding to the C-terminal polyQ region and the nuclear localisation signal (NLS) from the DNA of the ATX3_Q18 and ATX3_Q72 constructs and

cloning them downstream of green fluorescent protein (GFP) using primers 5'-AACGTCAGCTTCCTTCG-GAAGAGACGAGAAGCCTACTTTG-3' and 5'-ATC-GACGGATCCCTATGGATGTGAACTCTGTCCT-GATAG-3'. Using oligonucleotide incorporating codons for the repeats, 24 repeats of glutamines, serines or leucines were cloned downstream of GFP. The complementary oligonucleotides with *Xho*I and *Eco*RI sticky ends were annealed and cloned between the respective sites in the pEGFPC2 vector (Clontech), resulting in the GFP_Q24, GFP_S24 and GFP_L24 constructs. The oligonucleotide sequences were as follows: 5'-TCCAGG(CAG)²⁴TAG-3' and 5'-AATTCTA(CTG)²⁴CC-3' for glutamine repeats, 5'-TCCAGG(AGC)²⁴TAG-3' and 5'-AATTCTA (GCT)²⁴CC-3' for serine repeats, and 5'-TCGAGG(CTG)²⁴TAG-3' and 5'-AATTCTA (CAG)²⁴CC-3' for leucine repeats.

Cells, transfection and confocal microscopy. COS cells were grown in chamber slides. Transfection and confocal microscopy were carried out as described earlier [27]. After fixation and permeabilization, cells transfected with constructs expressing ATX3 were incubated either with anti-ATX3 N-terminal antibodies (a generous gift from Prof. Erich Wanker) at 1:1000 dilution or with anti-V5 antibodies (Invitrogen, 1:500 dilution) and FITC-conjugated secondary antibodies. Cells were incubated with primary and secondary antibodies for 1 h each. Nuclei were counterstained with DAPI (10 min). Cells were mounted as described earlier and labelled cells were visualised with a Leica laser scanning confocal microscope (TCS-SP1). Detection of the fluorescence signals of cells expressing GFP and GFP fusion proteins was achieved by using appropriate filter sets (excitation 488 nm/emission 505–530 nm). Images were acquired as single transcellular 0.2- μ m optical sections in the Z plane and averaged over 16 scans/frame. Fluorescent aggregates are readily identifiable by their size and intensity as well as compartmentalised appearance. Fluorescence was further confirmed as aggregate-associated and distinct from diffused staining by the presence across several optical sections in the Z plane. When necessary, images were merged using the Image J program (NIH, Bethesda). All transfections were carried out in duplicate and repeated at least three times to check for consistency. For each construct, ~250 cells expressing the protein were examined and counted in terms of aggregate formation. Results were shown as a percentage of total cells examined.

Western blot analysis. COS cells were seeded in duplicate in equal density on chamber slides and transfected with the appropriate DNA constructs. For each pair of chambers, cells from one chamber were processed for microscopy as described above 48 h after transfection, while the cells in the duplicate chamber were lysed in sample buffer. Aliquots of the cell lysates (20 μ L) were subjected

to electrophoresis and Western blot analysis using anti-V5 or anti-GFP antibodies. Western blot analysis against a housekeeping gene, glyceraldehyde-3-phosphate dehydrogenase (GAPDH), was carried out in parallel as a loading control, using comparable amounts of cell lysates.

Results


















General strategy. A number of different protein constructs were designed (for a summary, see Table 1) and their effect on the tendency to form cellular aggregates tested in transiently transfected mammalian cell lines (COS cells), using fluorescence microscopy experiments. To confirm that the observed differences in visible aggregate formation were not merely due to differences in protein expression levels in the transfected cells expressing different DNA constructs, Western blot analysis of all the cell lysates was carried out. Parallel transfections were carried out in twin chamber slides: cells from one chamber were processed for microscopy, while cells from the duplicate chamber were lysed and subjected to Western blot. Expression analysis of the housekeeping gene product GAPDH, which was detected using specific antibodies, confirmed equal loading of the cell lysates; the results shown in Fig. 1 indicate that the reported differences in aggregate formation were not due to variations in the expression of particular proteins.

The polyQ length determines different tendencies to aggregate. First, we compared the expression pattern of ATX3 in COS cells when they were transiently transfected with ATX3 constructs carrying polyQ tracts with 18 (ATX3_Q18) and 72 (ATX3_Q72) glutamines. We observed that when expressing ATX3 with a non-pathological polyQ tract (ATX3_Q18), ~85% of the cells consistently exhibited diffused nucleocytoplasmic formations, which stained for ATX3 with occasional appearance of small foci (Fig. 2a). Only a minority of the ATX3_Q18 cells (15%) exhibited larger aggregates (Fig. 2b). In contrast, ~60% of the cells that expressed ATX3_Q72 had very large aggregates exceeding several micrometres in size (Fig. 2c). Very few cells exhibited a completely diffused staining pattern and most of the remaining 40% cells had several small aggregates (Fig. 2d). Large aggregates were found in the cytoplasm and were mostly perinuclear in both ATX3_Q18 and ATX3_Q72 transfected cells, especially those of ATX3_Q72.

We conclude that, in our model system, polyQ expansion consistently increases the cell tendency to form large intracellular aggregates.

The effect of the protein context on the aggregation of ATX3. We then probed the effect of deleting or replacing the only globular region of ATX3, the Josephin domain,

Table 1. Summary of the constructs used in this study and pattern of expression observed in%. Cartoon representations of the constructs are given in the second column in which Josephin, the AXH of HBP1 and GFP are shown with a rectangle, circle and hexagon respectively. The homopolymeric sequences are shown in grey where as the rest of the tail, when present, is shown in black.

| Construct | Cartoon | None or little aggregation | | Aggregation | |
|-----------------|---|----------------------------|---------------------------|------------------|----------------------|
| | | Diffused expression | Diffused/small aggregates | Small aggregates | Large aggregates |
| ATX3_Q18 |  | | 85 | | 15 |
| ATX3_Q18-ΔJos |  | 98 | | 2 | |
| ATX3_Q18-HBP1 |  | 98 | | 2 | |
| ATX3_Q18-GFP |  | 100 | | | |
| ATX3_Q72 |  | | | 40 | 60 |
| ATX3_Q72-ΔJos |  | 64 | | 30 | 6(nuc) ^a |
| ATX3_Q72-Δ1-217 |  | 66 | | 28 | 6(nuc) ^a |
| ATX3_Q72-Δ1-225 |  | 64 | | 31 | 5(nuc) ^a |
| ATX3_Q72-HBP1 |  | 78 | | 22 | |
| ATX3_Q72-GFP |  | 79 | | | 21(nuc) ^a |
| ATX3_A72 |  | | 65 | | 35 |
| ATX3_S72 |  | | | | 100 |
| GFP |  | 100 | | | |
| GFP_Q24 |  | 100 | | | |
| GFP_A24 |  | 100 | | | |
| GFP_S24 |  | 100 | | | |
| GFP_L24 |  | | | | 100 |

^a Exclusively nuclear aggregates.

on cellular aggregate formation. We deleted the Josephin domain from expanded and non-expanded ATX3 (ATX3_Q18-ΔJos and ATX3_Q72-ΔJos). Additionally, we made two constructs in which the Josephin domain was replaced with the AXH domain of HBP1 (ATX3_Q18-HBP1 and ATX3_Q72-HBP1), a domain with no known tendency to aggregate [11]. These constructs resulted in proteins comparable in size to full-length ATX3. Finally, the ATX3 C terminus was fused to the GFP, which is well characterised, highly soluble and easy to detect (ATX3_Q18-GFP and ATX3_Q72-GFP).

Transfection of cells with non-expanded ATX3_Q18-ΔJos and ATX3_Q18-HBP1 constructs showed that deletion or replacement of Josephin resulted in diffused expression and absence of aggregates in 98% of the cells examined. Only 2% or less of these cells showed some aggregates (Fig. 3a, b). These were smaller in size than those observed in a minority (15%) of the ATX3_18Q-expressing cells (Fig. 2b), and bore more resemblance to the small-foci-forming majority of cells expressing ATX3_Q18. Likewise, replacement of Josephin with GFP had a similar effect: ATX3_Q18-GFP had an appearance similar to that of GFP alone which, in the same cells, showed a diffused expression pattern both in the nucleus and in the cytoplasm (Fig. 4a). This behaviour should be compared with the pattern observed in the

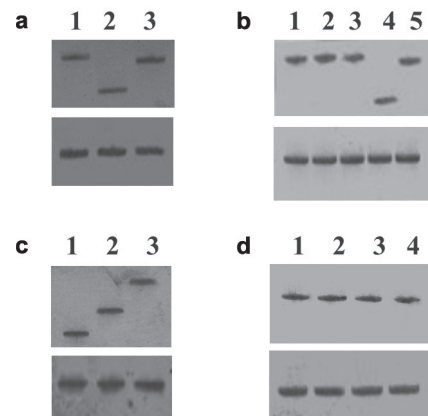


Figure 1. Protein expression analysis of cells transfected with ATX3 and GFP constructs. Lysates of COS cells that were seeded at similar density, and transfected and grown under the same conditions as those processed for microscopy, were subjected to Western blot analysis using appropriate antibodies recognising the overexpressed proteins (upper sections in *a-d*) or GAPDH, which was used as a loading control (lower sections of the panels). The panels represent the following transfections: *a*: lane 1, ATX3_Q18; lane 2, ATX3_Q18-ΔJos; lane 3, ATX3_Q18-HBP1. *b*: lane 1, ATX3_Q72; lane 2, ATX3_A72; lane 3, ATX3_S72; lane 4, ATX3_Q72-ΔJos; lane 5, ATX3_Q72-HBP1. *c*: lane 1, GFP; lane 2, ATX3_Q18-GFP; lane 3, ATX3_Q72-GFP. *d*: lane 1, GFP_24Q; lane 2, GFP_24L; lane 3, GFP_24S; lane 4, GFP.

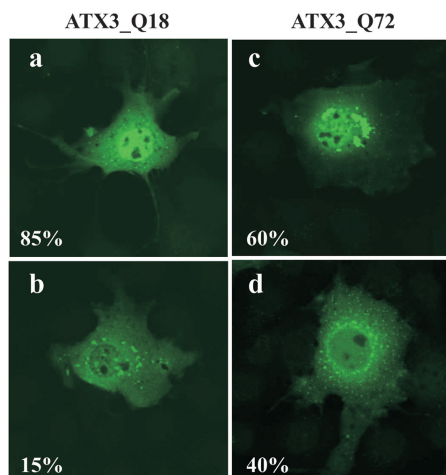


Figure 2. Expression pattern of ATX3 in transfected COS cells. COS cells were transiently transfected with expanded (ATX3_Q72) or non-expanded (ATX3_Q18) ATX3. After staining, the cells were subjected to confocal immunofluorescence microscopy. Majority of ATX3_Q18 (a) cells exhibit a mostly diffused staining for the protein, while about 15% appear to have large aggregates (b). In contrast, 60% of cells transfected with ATX3_Q72 have large, often perinuclear, aggregates (c) whereas the remainder of cells have smaller aggregates (d).

ATX3_Q18 cells in which a minor, yet significant portion of the cells displayed large cytoplasmic aggregates (*cf.* Figs 2, 4b).

When the effect of expanded constructs was analysed, the staining of most cells expressing ATX3_Q72- Δ Jos

or ATX3_Q72-HBP1 appeared diffused with no visible aggregates (data not shown). This is at strong variance with what observed for ATX3_Q72 (Fig. 2c, d). Where the cells appeared to have aggregates (*i.e.* 30% of cells for ATX3_Q72- Δ Jos and 22% of cells for ATX3_Q72-HBP1), these were much smaller than those observed in the majority of ATX3_Q72-expressing cells (*cf.* Fig. 3e, d and Fig. 2c). In ~6% of the cells expressing ATX3_Q72- Δ Jos, the aggregates were exclusively nuclear (data not shown) at variance with what we had observed for the full-length expanded or non-expanded ATX3. This is interesting in light of the recent proposal that proteolytic cleavage of ATX3 could precipitate nuclear aggregation and cytotoxicity of the protein [16, 17, 28]. To mimic the effect of this proteolysis, which is thought to occur at caspase sites of ATX3, we made two more constructs with N-terminal deletions (ATX3_Q72- Δ 1-217 and ATX3_Q72- Δ 1-225). They correspond to the two C-terminal caspase cleavage products of ATX3. Once again, in most of the cells the staining was diffused and without visible aggregates. In both cases, ca. 30% of the cells appeared to have small cytoplasmic aggregates (Fig. 3e, g), whereas 5–6% of the cells had larger exclusively nuclear formations (Fig. 3f, h). Cells expressing ATX3_Q72-GFP were comparatively less prone to aggregation than ATX3_Q72 cells (only 21%, of the cells contained aggregates) (Fig. 4c). All the aggregates observed in ATX3_Q72-GFP experiments were exclusively nuclear.

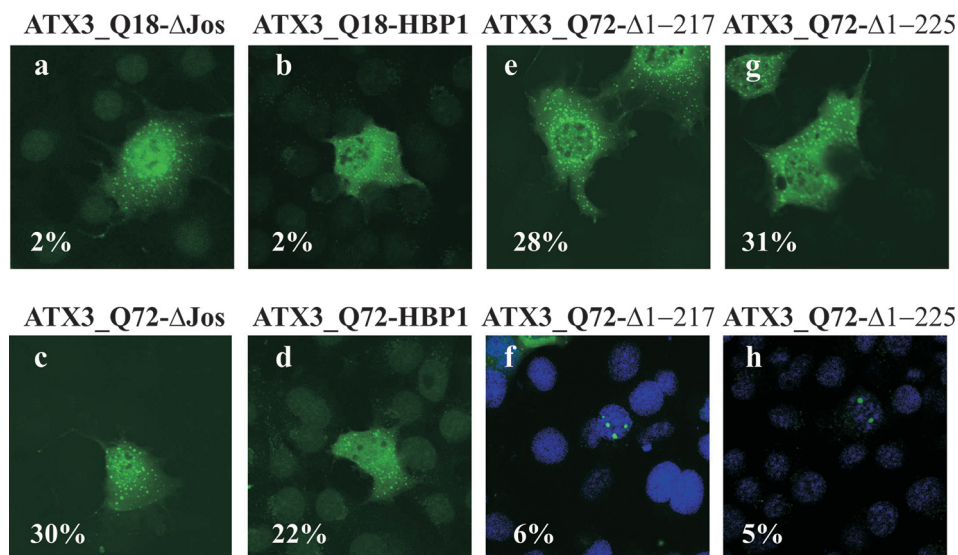


Figure 3. The effect of N-terminal deletion of ATX3 or of swapping the Josephin domain with a non-aggregation-prone domain on aggregate-forming tendency in transfected cells. The panels correspond to cells overexpressing the following constructs: ATX3_Q18- Δ Jos (a), ATX3_Q18-HBP1 (b), ATX3_Q72- Δ Jos (c), ATX3_Q72-HBP1 (d), ATX3_Q72- Δ 1-217 (e, f) and ATX3_Q72- Δ 1-225 (g, h). In all cases, deletion or replacement of Josephin with a domain that was not prone to aggregation appears to prevent aggregate formation in most of the cells (*cf.* Fig. 2). The aggregates appear to be consistently smaller in size in the case of truncated ATX3_Q72-expressing cells compared with that observed for the cells expressing the equivalent full-length constructs shown in Figure 2. Exclusively nuclear aggregate formation was observed in a small number of cells expressing N-terminal deletions of ATX3 (ATX3_Q72- Δ 1-217 and ATX3_Q72- Δ 1-225). In (f) and (h) cells were counterstained with DAPI and images were overlaid to appreciate nuclear (blue) as well as ATX3-associated (green) fluorescence.

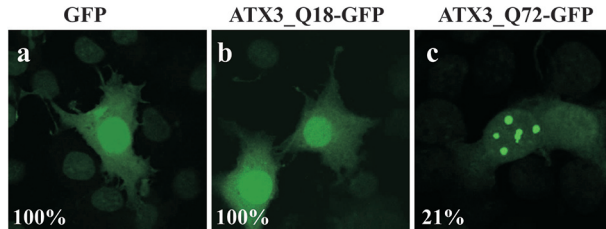


Figure 4. Effect of fusing the expanded and non-expanded ATX3 C terminus to GFP. The expression pattern of cells transfected with constructs in which GFP was fused upstream to the C terminus of ATX3_Q18 or ATX3_Q72 was compared with that of GFP alone. GFP (a) and ATX3_Q18-GFP (b) both appear diffused in both the nucleus and the cytoplasm in all cells examined. While most cells expressing ATX3_Q72-GFP show a similar appearance, ~21% of these cells have large, exclusively nuclear aggregates (c).

These results clearly indicate that protein regions outside the expanded amino acid tracts contribute and modulate the behaviour of polyQ expansion. Their roles appear to become increasingly more important upon expanding the length of the polyQ tract.

Testing the effect of different homo-polymer expansions in ATX3. We then probed the effects of expansion of different poly-amino acid homo-sequences on aggregation. We introduced frameshift mutations in expanded ATX3 to change the polyQ repeats into polyS and polyA. The frameshift was corrected immediately after the poly-amino acid tract, going back to the ATX3 frame.

COS cells were transiently transfected with these constructs. Compared to cells transfected with ATX3_Q72, there was a marked difference in the aggregate forming tendency displayed by ATX3_A72. The vast majority of these cells (~65%) displayed a diffused staining pattern with occasional small foci (not shown), while the rest of the cells had aggregates which, although exceeding several micrometres in size, were generally smaller compared with the large aggregates of ATX3_Q72 cells (Fig. 5a, b). Examination of the cells expressing ATX3_S72 revealed a very different phenotype: all the cells overexpressing ATX3_S72 had very large mostly perinuclear aggregates, indicating that the tendency of ATX3 to form aggregates varies with the nature of the poly-amino acid (Fig. 5c).

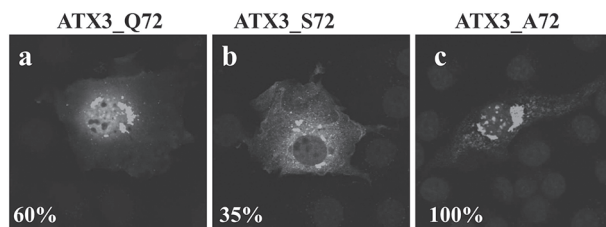


Figure 5. Effect of replacing the polyQ repeats in ATX3_Q72 with polyA and polyS repeats. While ~60% cells expressing ATX3_Q72 have large aggregates (a), fewer ATX3_A72 cells (b) have large aggregates, and their sizes are appreciably reduced. In contrast, all cells expressing ATX3_S72 (c) have large aggregates.

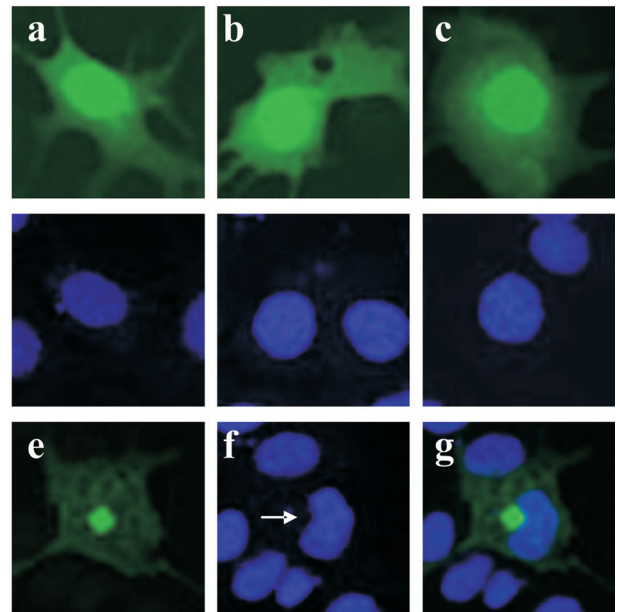


Figure 6. Effect of fusing different amino acid homo-polymers to GFP. Fluorescence analysis of cells transfected with GFP (a), GFP_24Q (b), GFP_24S (c) or GFP_24L (e-g) was carried out. Neither GFP_24Q nor GFP_24S show any significant difference to GFP-expressing cells and the green fluorescence appears as diffused and nucleo-cytoplasmic in all cells examined. Perinuclear aggregate formation and distortion of the nucleus (arrow in f) are observed instead in GFP_24L-expressing cells.

As a further step to compare the effect of different amino acid homo-polymers, we made GFP fusion constructs having 24 poly-leucine (polyL), poly-alanine (polyA), polyglutamine (polyQ) or poly-serine (polyS) sequences. GFP_Q24- and GFP_S24-expressing cells appeared no different from GFP-expressing cells, exhibiting diffused nucleo-cytoplasmic fluorescence (Fig. 6a-c). A fluorescence pattern similar to that observed in GFP-expressing cells was also observed for GFP_A24 cells (data not shown). In contrast, the morphology of GFP_L24 cells appeared completely different. In these cells, green fluorescence appeared as a single large perinuclear aggregate agglomerated against the nucleus and distorting its appearance (Fig. 6e-g). Collectively, these data indicate that poly-amino acids of the same length have different aggregation properties and that polyL repeats have a stronger effect than the other poly-amino acids tested, not only regarding aggregate formation tendency but also with respect to their ability to cause microscopically visible cellular architecture defects.

Discussion

Repeats of amino acid homo-polymers have been implicated in an increasing number of human pathologies caused by protein misfolding and aggregation. To clarify

their role in neurodegeneration, we have addressed here two questions: first, we have evaluated the importance of the protein context in polyQ aggregation using ATX3 as a model system; second, we have compared the aggregation tendencies of different poly-amino acid sequences using both ATX3 and GFP as carriers. To address the first question, we assessed the effect of fusing polyQ sequences below and above the pathological range (which for SCA3 is around 50–60 repeats) to different protein carriers by comparing the percentage of cells that contained aggregates. Interestingly, we did not observe any case in which expansion was associated exclusively with nuclear aggregates, such as those reported in SCA3-affected neurons [29]. In our system, expanded ATX3 tended to form predominantly cytoplasmic aggregates. One possibility that would explain this discrepancy is that, in addition to the NLS [30], ATX3 contains an as-yet-unidentified cytoplasmic retention signal, the effect of which may be more pronounced in an overexpression system. Another possible explanation is that the aggregation kinetics of expanded ATX3 are so fast that, in overexpressing cells, aggregation takes place before the protein has had any chance of getting transported across the nuclear membrane. This hypothesis may also explain the mostly perinuclear appearance of the aggregates, which are similar to those observed in ATX3 staining of affected brain tissues where, in addition to ATX3 nuclear aggregates, there is some ‘perinuclear granular staining’ [31]. Although likely, it is not known whether the perinuclear aggregates can affect nuclear transport. In a *Drosophila* model of Huntington’s disease, cytoplasmic aggregates were shown to trap proteins and block axonal transport, suggesting that non-nuclear events brought about by cytoplasmic aggregation play a role in neurodegeneration [32]. Moreover, recent studies on Huntington, the polyQ expanded protein responsible for the Huntington’s pathology, have attributed a critical role to both nuclear and extranuclear polyQ tracts into disease progression [33]. Considering that ATX3 also is a predominantly cytoplasmic protein, it is possible that similar events occur in SCA3.

An important finding of our study is that, while polyQ expansion is essential to observe appreciable aggregation in our model system, the presence and the nature of the protein context strongly modulates the amount, the size and the morphology of the aggregates. We have observed a qualitatively similar effect by deleting the Josephin domain, by replacing it, or by adding to the construct another highly soluble domain. These findings are fully in agreement with results reported for ataxin 1 [11], the protein responsible for SCA1 ataxia, and set a clear hierarchical relationship between polyQ and other fibrillogenic regions of polyQ proteins. Accepting that aggregation is primarily governed by poly-amino acid expansion but modulated by the protein context is crucially important for understanding the mechanism of polyQ aggregation.

We also believe that our findings are not in conflict with the suggestion that proteolytic cleavage could be the basis of a common mechanism of aggregation and toxicity [16, 17, 28]. Our results clearly indicate that C-terminal products of ATX3 with expanded polyQ tracts are capable of forming exclusively nuclear aggregates. It is therefore possible that truncated C-terminal products of ATX3 having expanded polyQ may have a role in initiating the formation of nuclear aggregation. Once this process is initiated, other components may get involved and be sequestered together, including the more aggregation-prone full-length expanded ATX3.

Support, albeit indirect, to this hypothesis is given by studies in yeast models, which have shown that efficient cytoplasmic aggregation and toxicity of polyQ proteins can be detected only in yeast strains containing an endogenous glutamine/asparagine (Q/N)-rich protein, Rnq1p, in its prion form [34]. In this system, the aggregation initiated by Rnq1p seems to lead to heterologous ‘seeding’, which promotes aggregation of polyQ proteins. Rnq1p is one of several known yeast proteins, including New1p, Sup35p and Ure2p, which contain Q/N-rich prion domains [35–37]. Studies on Sup35p and New1p have revealed that, while the Q/N tract mediates sequence-specific aggregation, the adjacent oligo peptide motif is required for the replication and stable inheritance of the aggregates [38, 39]. Interestingly, the Q/N rich domains of New1p and Rnq1p in their prion forms have been shown to facilitate aggregation of a polyQ construct from the mutant ATX3 [40].

To compare the tendency of aggregation of different poly-amino acid sequences, we have also studied their behaviour when linked both to ATX3 and to the non-fibrillogenic GFP. We observed clear differences amongst the different homo-polymers both when they were in the context of the ATX3 protein and when they were directly linked to GFP. Expanded polyA sequences seem overall to be less toxic than polyQ tracts. PolyS and polyL have comparatively stronger phenotypes. Since there are not yet reports about the involvement of polyS or polyL tract expansions in human diseases, it is tempting to suggest that serine and leucine repeats may be too toxic and that any such expansion is not tolerated beyond the embryonic stage, unlike some of the polyQ and polyA expansions. Consistently, a search through the database of human proteins shows that proteins with glutamine or alanine repeats are far more abundant than those with serine repeats or repeats of hydrophobic amino acids such as leucines [6].

In conclusion, our results not only shed light on the different aggregation tendencies of different amino acid homo-polymers but also provide new evidence in favour of a role of the Josephin domain in the aggregation of ATX3 *in vivo*.

Acknowledgements. We thank Erich Wanker for ATX3 antibodies. The work was funded by the EUROSCA project.

- 1 Brown L. Y. and Brown S. A. (2004) Alanine tracts: the expanding story of human illness and trinucleotide repeats. *Trends Genet.* 20, 51–58.
- 2 Manto M. U. (2005) The wide spectrum of spinocerebellar ataxias (SCAs). *Cerebellum* 4, 2–6.
- 3 Albrecht A. and Mundlos S. (2005) The other trinucleotide repeat: polyalanine expansion disorders. *Curr. Opin. Genet. Dev.* 15, 285–293.
- 4 Everett C. M. and Wood N. W. (2004) Trinucleotide repeats and neurodegenerative disease. *Brain* 127, 2385–2405.
- 5 Costa Lima M. A. and Pimentel M. M. (2004) Dynamic mutation and human disorders: the spinocerebellar ataxias. *Int. J. Mol. Med.* 13, 299–302.
- 6 Oma Y., Kino Y., Sasagawa N. and Ishiura, S. (2004) Intracellular localization of homopolymeric amino acid-containing proteins expressed in mammalian cells. *J. Biol. Chem.* 279, 21217–21222.
- 7 Faux N. G., Bottomley S. P., Lesk A. M., Irving J. A., Morrison J. R., de la Banda M. G. and Whisstock, J. C. (2005) Functional insights from the distribution and role of homeopeptide repeat-containing proteins. *Genome Res.* 15, 537–551.
- 8 Chow M. K., Paulson H. L. and Bottomley S. P. (2004) Destabilization of a non-pathological variant of ataxin-3 results in fibrillogenesis via a partially folded intermediate: a model for misfolding in polyglutamine disease. *J. Mol. Biol.* 335, 333–341.
- 9 Masino L., Nicastro G., Menon R. P., Dal Piaz F., Calder L. and Pastore A. (2004) 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.
- 10 Gales L., Cortes L., Almeida C., Melo C. V., do Carmo C. M., Maciel P., Clarke D. T., Damas A. M. and Macedo-Ribeiro S. (2005) 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.
- 11 de Chiara C., Menon R. P., Dal Piaz F., Calder L. and Pastore A. (2005) Polyglutamine is not all: the functional role of the AXH domain in the ataxin-1 protein. *J. Mol. Biol.* 354, 883–893.
- 12 Bhattacharyya A., Thakur A. K., Chellgren V. M., Thiagarajan G., Williams A. D., Chellgren B. W., Creamer T. P. and Wetzel R. (2006) Oligoproline effects on polyglutamine conformation and aggregation. *J. Mol. Biol.* 355, 524–535.
- 13 Kim Y. J., Yi Y., Sapp E., Wang Y., Cuiffo B., Kegel K. B., Qin Z. H., Aronin N. and DiFiglia M. (2001) 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. USA* 98, 12784–12789.
- 14 Sun B., Fan W., Balciunas A., Cooper J. K., Bitan G., Stevenson S., Denis P. E., Young Y., Adler B., Daugherty L., Manoukian R., Elliott G., Shen W., Talvenheimo J., Teplow D. B., Haniu M., Haldankar R., Wypych J., Ross C. A., Citron M. and Richards W. G. (2002) Polyglutamine repeat length-dependent proteolysis of huntingtin. *Neurobiol. Dis.* 11, 111–122.
- 15 Tarlac V. and Storey E. (2003) Role of proteolysis in polyglutamine disorders. *J. Neurosci. Res.* 74, 406–416.
- 16 Berke S. J., Schmied F. A., Brunt E. R., Ellerby L. M. and Paulson H. L. (2004) Caspase-mediated proteolysis of the polyglutamine disease protein ataxin-3. *J. Neurochem.* 89, 908–918.
- 17 Haacke A., Broadley S. A., Boteva R., Tzvetkov N., Hartl F. U. and Breuer P. (2006) Proteolytic cleavage of polyglutamine-expanded ataxin-3 is critical for aggregation and sequestration of non-expanded ataxin-3. *Hum. Mol. Genet.* 15, 555–568.
- 18 Paulson H. L., Das S. S., Crino P. B., Perez M. K., Patel S. C., Gotsdiner D., Fischbeck K. H. and Pittman R. N. (1997) Machado-Joseph disease gene product is a cytoplasmic protein widely expressed in brain. *Ann. Neurol.* 41, 453–462.
- 19 Chai Y., Koppenhafer S. L., Shoemith S. J., Perez M. K. and Paulson, H. L. (1999) Evidence for proteasome involvement in polyglutamine disease: localization to nuclear inclusions in SCA3/MJD and suppression of polyglutamine aggregation *in vitro*. *Hum. Mol. Genet.* 8, 673–682.
- 20 Donaldson K. M., Li W., Ching K. A., Batalov S., Tsai C. C. and Joazeiro C. A. (2003) Ubiquitin-mediated sequestration of normal cellular proteins into polyglutamine aggregates. *Proc. Natl. Acad. Sci. USA* 100, 8892–8897.
- 21 Burnett B., Li F. and Pittman R. N. (2003) The polyglutamine neurodegenerative protein ataxin-3 binds polyubiquitylated proteins and has ubiquitin protease activity. *Hum. Mol. Genet.* 12, 3195–3205.
- 22 Chai Y., Berke S. S., Cohen R. E. and Paulson H. L. (2004) Poly-ubiquitin binding by the polyglutamine disease protein ataxin-3 links its normal function to protein surveillance pathways. *J. Biol. Chem.* 279, 3605–3611.
- 23 Warrick J. M., Morabito L. M., Bilen J., Gordesky-Gold B., Faust L. Z., Paulson H. L. and Bonini N. M. (2005) Ataxin-3 suppresses polyglutamine neurodegeneration in *Drosophila* by a ubiquitin-associated mechanism. *Mol. Cell* 18, 37–48.
- 24 Masino L., Musi V., Menon R. P., Fusi P., Kelly G., Frenkiel T. A., Trottier Y. and Pastore A. (2003) Domain architecture of the polyglutamine protein ataxin-3: a globular domain followed by a flexible tail. *FEBS Lett.* 549, 21–25.
- 25 Nicastro G., Menon R. P., Masino L., Knowles P. P., McDonald N. Q. and Pastore A. (2005) The solution structure of the Josephin domain of ataxin-3: structural determinants for molecular recognition. *Proc. Natl. Acad. Sci. USA* 102, 10493–10498.
- 26 Mao Y., Senic-Matuglia F., Di Fiore P. P., Polo S., Hodsdon M. E. and De Camilli P. (2005) Deubiquitinating function of ataxin-3: insights from the solution structure of the Josephin domain. *Proc. Natl. Acad. Sci. USA* 102, 12700–12705.
- 27 Menon R. P., Gibson T. J. and Pastore A. (2004) The C terminus of fragile X mental retardation protein interacts with the multi-domain Ran-binding protein in the microtubule-organising centre. *J. Mol. Biol.* 343, 43–53.
- 28 Goti D., Katzen S. M., Mez J., Kurtis N., Kiluk J., Ben Haiem L., Jenkins N. A., Copeland N. G., Kakizuka A., Sharp A. H., Ross C. A., Mouton P. R. and Colomer V. (2004) 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.* 24, 10266–10279.
- 29 Paulson H. L., Perez M. K., Trottier Y., Trojanowski J. Q., Subramony S. H., Das S. S., Vig P., Mandel J. L., Fischbeck K. H. and Pittman R. N. (1997) Intranuclear inclusions of expanded polyglutamine protein in spinocerebellar ataxia type 3. *Neuron* 19: 333–344.
- 30 Tait D., Riccio M., Sittler A., Scherzinger E., Santi S., Ognibene A., Maraldi N. M., Lehrach H. and Wanker E. E. (1998) Ataxin-3 is transported into the nucleus and associates with the nuclear matrix. *Hum. Mol. Genet.* 7, 991–997.
- 31 Fujigasaki H., Uchihara T., Koyano S., Iwabuchi K., Yagishita S., Makifuchi T., Nakamura A., Ishida K., Toru S., Hirai S., Ishikawa K., Tanabe T. and Mizusawa H. (2000) 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.
- 32 Lee W. C., Yoshihara M. and Littleton J. T. (2004) Cytoplasmic aggregates trap polyglutamine-containing proteins and block axonal transport in a *Drosophila* model of Huntington's disease. *Proc. Natl. Acad. Sci. USA* 101, 3224–3229.
- 33 Benn C. L., Landles C., Li H., Strand A. D., Woodman B., Sathasivam K., Li S. H., Ghazi-Noori S., Hockley E., Faruque S. M., Cha J. H., Sharpe P. T., Olson J. M., Li X. J. and Bates G. P. (2005) Contribution of nuclear and extranuclear polyQ to neurological phenotypes in mouse models of Huntington's disease. *Hum. Mol. Genet.* 14, 3065–3078.
- 34 Meriin A. B., Zhang X., He X., Newnam G. P., Chernoff Y. O. and Sherman M. Y. (2002) Huntington toxicity in yeast model

- depends on polyglutamine aggregation mediated by a prion-like protein Rnq1. *J. Cell. Biol.* 157, 997–1004.
- 35 DePace A. H., Santoso A., Hillner P. and Weissman J. S. (1998) A critical role for amino-terminal glutamine/asparagine repeats in the formation and propagation of a yeast prion. *Cell* 93, 1241–1252.
- 36 Maddelein M. L. and Wickner R. B. (1999) Two prion-inducing regions of Ure2p are nonoverlapping. *Mol. Cell. Biol.* 19, 4516–4524.
- 37 Chernoff Y. O. (2004) Replication vehicles of protein-based inheritance. *Trends Biotechnol.* 22, 549–552.
- 38 Osherovich L. Z., Cox B. S., Tuite M. F. and Weissman J. S. (2004) Dissection and design of yeast prions. *PLoS Biol.* 2, E86.
- 39 Shkundina I. S., Kushnirov V. V., Tuite M. F. and Ter-Avanesyan M. D. (2006) The role of the N-terminal oligopeptide repeats of the yeast Sup35 prion protein in propagation and transmission of prion variants. *Genetics* 172, 827–835.
- 40 Osherovich L. Z. and Weissman J. S. (2001) Multiple Gln/Asn-rich prion domains confer susceptibility to induction of the yeast [PSI(+)] prion. *Cell* 106, 183–194.



To access this journal online:
<http://www.birkhauser.ch>
