Trinucleotide expansions leading to an extended poly-L-alanine segment in the poly (A) binding protein PABPN1 cause fibril formation

TILL SCHEUERMANN,[1,4] BARBE SCHULZ,[2] ALFRED BLUME,[3] ELMAR WAHLE,[2] RAINER RUDOLPH,[1] AND ELISABETH SCHWARZ[1]

[1]Martin-Luther-Universität Halle-Wittenberg, Institut für Biotechnologie and [2]Institut für Biochemie, 06120 Halle, Germany
[3]Institut für Physikalische Chemie, 06108 Halle, Germany

(RECEIVED May 21, 2003; FINAL REVISION August 7, 2003; ACCEPTED August 14, 2003)

Abstract

The nuclear poly(A) binding protein (PABPN1) stimulates poly(A) polymerase and controls the lengths of poly(A) tails during pre-mRNA processing. The wild-type protein possesses 10 consecutive Ala residues immediately after the start methionine. Trinucleotide expansions in the coding sequence result in an extension of the Ala stretch to maximal 17 Ala residues in total. Individuals carrying the trinucleotide expansions suffer from oculopharyngeal muscular dystrophy (OPMD). Intranuclear inclusions consisting predominantly of PABPN1 have been recognized as a pathological hallmark of the genetic disorder. To elucidate the molecular events that lead to disease, recombinant PABPN1, and N-terminal fragments of the protein with varying poly-L-alanine stretches were analyzed. As the full-length protein displayed a strong tendency to aggregate into amorphous deposits, soluble N-terminal fragments were also studied. Expansion of the poly-L-alanine sequence to the maximal length observed in OPMD patients led to an increase of α-helical structure. Upon prolonged incubation the protein was found in fibrils that showed all characteristics of amyloid-like fibers. The lag-phase of fibril formation could be reduced by seeding. Structural analysis of the fibrils indicated antiparallel β-sheets.

Keywords: Amyloid-like fibrils; seeding; trinucleotide expansions; poly-L-alanine; oculopharyngeal muscular dystrophy

Several human disorders are caused by trinucleotide expansions, including the often-cited and well-studied Huntington's disease and spinocerebellar ataxia type I, two disorders caused by genetic polymorphisms resulting in the expansion of Gln stretches (Trottier et al. 1995; Klement et al. 1998; Cummings and Zoghbi 2000). For both Huntington and ataxin, less than 40 consecutive Gln residues represent the wild-type situation. Above this threshold, additional Gln residues result in abnormal phenotypes with a clear correlation between the number of surplus Gln and early onset of the disease. Pathologically, the diseases are characterized by the deposition of the mutant proteins in amyloid fibers. Although fibril formation is generally thought to be caused by conformational changes leading to β-sheet structures, the actual cause finally leading to the disease is still under debate. In addition to poly-L-glutamine extensions, gene polymorphisms causing the expansion of Ala stretches have also been reported. Short poly-L-alanine repeats of about 10 residues are commonly found in nature. Their biochemical function is, however, so far unknown except for spider fiber protein where the poly-L-alanine segments endow the fibers with crystalline properties (Simmons et al. 1996). Poly-L-alanine extensions causing diseases have so far been detected exclusively in transcription factors and the nuclear

Reprint requests to: Elisabeth Schwarz, Martin-Luther-Universität Halle-Wittenberg, Institut für Biotechnologie, Kurt-Mothes-Str. 3, 06120 Halle, Germany; e-mail: elisabeth.schwarz@biochemtech.uni-halle.de; fax: + 49-345-55-27-013.
[4]Present address: Roche Diagnostics GmbH, Nonnenwald 2, 82372 Penzberg, Germany
Article and publication are at http://www.proteinscience.org/cgi/doi/10.1110/ps.03214703.

poly(A) binding protein PABPN1 (Muragaki et al. 1996; Mundlos et al. 1997; Brais et al. 1998; Brown et al. 1998; Goodman et al. 2000; Crisponi et al. 2001; Galant and Carroll 2002; Ronshaugen et al. 2002). In contrast to the CAG polymorphisms with extreme extensions that result in poly-L-glutamine stretches of up to 180 residues, GCG repeats leading to an extension of Ala segments appear moderate and are genetically stable. So far, repeats of only up to 10 additional Ala residues have been reported that extend the normally occurring 10–18 Ala segments (Muragaki et al. 1996; Mundlos et al. 1997; Brais et al. 1998; Brown et al. 1998; Crisponi et al. 2001).

The nuclear poly(A) binding protein (PABPN1) stimulates polyadenylation and controls the length of poly(A) tails (Wahle 1991; Wahle and Rüegsegger 1999). Immediately following the start methionine, PABPN1 has a natural 12 Ala stretch that is interrupted after the tenth Ala residue by a single Gly residue (Fig. 1). Trinucleotide expansions result in an extension of the wild-type 10 Ala sequence before the Gly residue to a maximum of 17 Ala residues. Individuals carrying the extensions develop the disease oculopharyngeal muscular dystrophy (OPMD) that is characterized by swallowing difficulties, eyelid drooping, and limb weakness (Brais et al. 1998). Biopsy material from patients revealed intranuclear inclusions of palisade-like structures in muscle fibers (Tomé et al. 1997). Strikingly, PABPN1 is the main constituent of the fibers together with ubiquitin, proteasomal subunits, and poly(A) RNAs (Calado et al. 2000).

PABPN1 is a 33-kD protein containing a single RNA binding domain (residues 161–258) consisting of a eukaryotic RNA recognition motif that is flanked by an acidic N and a basic C terminus (Fig. 1). Analysis of truncation constructs indicated that the C terminus (residues 259–306) contributes to RNA binding (Kühn et al. 2003). Residues 125–161 are predicted to form an α-helix, and are required for the stimulation of the poly(A) polymerase (Kerwitz et al. 2003). The function of the first 124 amino acids during polyadenylation is yet unknown. A recent publication revealed that an N-terminal fragment of PABPN1 comprising the first 145 residues interacts with the transcription factor SKIP and enhances synthesis of myogenic factors (Kim et al. 2001).

The molecular mechanism by which the Ala extensions cause OPMD is completely unknown. To learn more about the molecular steps leading to the disease, we carried out biophysical comparisons of the wild-type protein with PABPN1 variants containing the most extreme extension of seven additional Ala (PABPN1-[+7-]Ala) and a PAPBN1 variant lacking the complete N-terminal Ala stretch (PABPN1-ΔAla) (Fig. 1). No structural differences due to the additional poly-L-alanines were detected via circular dichroism (CD) analysis. However, when N-terminal fragments were investigated, the variant carrying seven additional Ala residues showed increased α-helical structure and exhibited a pronounced propensity to form fibrils, in contrast to the fragments derived from wild-type and the Ala deletion. Seeding with preformed fibrils reduced the lag phase of fibril formation.

Results

Secondary structure and stability of PABPN1

Secondary structure predictions of PABPN1 using PredictProtein by Expasy indicated a low secondary structure content of the protein apart from the ribonucleoprotein (RNP–) domain (Fig. 1A). The low content of predicted secondary structural elements at the termini of PABPN1 might be due to the high content of Pro and Glu (20 Pro and 24 Glu within 125 residues) in the N terminus and the highly positively charged C terminus (11 Arg within 40 residues). PABPN1-WT, PABPN1-(+7)Ala and PABPN1-ΔAla showed a strong tendency to aggregate during purification. However, the aggregates did not exhibit any amyloid-like structure as analyzed by electron microsocopy, Congo red staining, or

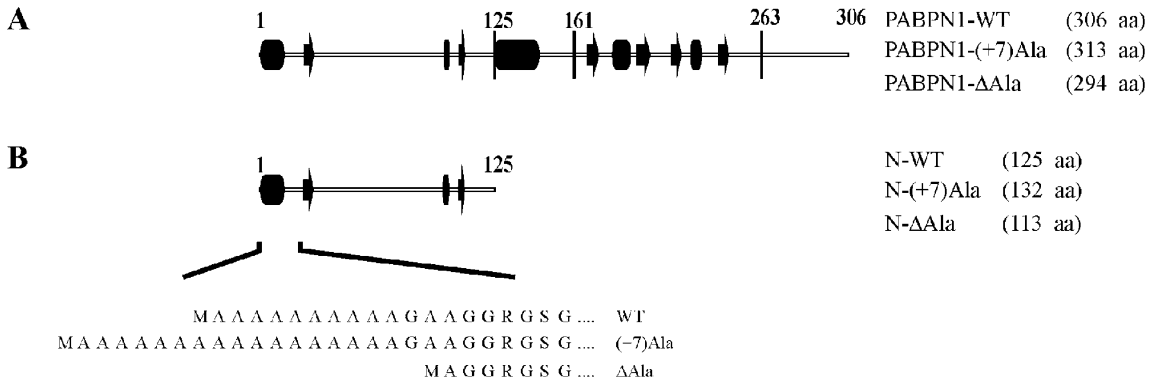

Figure 1. (*A*) Schematic presentation of the PABPN1. Cylinders indicate α-helical segments; arrows, β-sheets. (*B*) Schematic presentation of the N-terminal fragments. Amino acids for the three different variants are indicated.

FTIR-spectroscopy (data not shown). Unspecific aggregation could be retarded by buffers of high ionic strength such as Tris buffer with 1.5 M KCl. The computer-predicted low content of secondary structural elements was confirmed by far-UV CD spectroscopy (data not shown). Comparison of the CD spectra of the three PABPN1 variants, PABPN1-WT, PABPN1-(+7)Ala, and PABPN1-ΔAla, revealed no structural differences that could have been caused by the presence of surplus Ala or the deletion of all Ala residues in the N terminus (data not shown). Chemical denaturation of the variants with guanidinium thiocyanate was irreversible, and unfolding at low guanidinium concentrations was accompanied by aggregation as observed by light scattering (data not shown). Denaturant-dependent unfolding curves did not differ between the three variants. Thus, neither a stabilizing nor destabilizing effect of the presence or absence of poly-L-alanine segments could be detected.

N-terminal constructs of PABPN1

As the full-length proteins displayed the above mentioned high propensity for unspecific aggregate formation, characterization of truncated constructs of PABPN1 was considered to reveal more information about poly-L-alanine-induced fibril formation. To this end, a tryptic digestion of PABPN1-WT was performed. Proteolysis products were analyzed by N-terminal sequencing and mass spectrometry after purification by RP-HPLC. A major fragment comprising residues 126–263 was identified (data not shown). As proteolysis at residue 125 is likely to reflect a natural domain boundary, N-terminal variants with N-terminal His-tags were constructed that comprised residues 1–125 of the wild-type sequence (N-WT), the corresponding seven Ala extended version (N-[+7]Ala) and the deletion variant (N-ΔAla; Fig. 1B). In contrast to the full-length proteins, all truncated constructs lost the tendency to aggregate. Far-UV CD-spectroscopy indicated a low content of secondary structural elements as for the full-length proteins (Fig. 2A). Comparison of the spectra of N-WT and N-ΔAla showed only a slight loss of secondary structure in N-ΔAla. In contrast, the addition of seven Ala resulted in a different CD-spectrum that suggested a gain in secondary structure. The difference spectrum of N-(+7)Ala and N-WT indicated that the increased CD signal is due to α-helical structure (Fig. 2B). Evaluation of the stability of the three variants by chemically or thermally induced unfolding did not show any cooperative transition. Thus, tertiary contacts may be absent in the N-terminal fragments.

Fibril formation caused by poly-L-alanine extension

As tissue from patients with poly-L-alanine extensions shows characteristic fibrillar inclusions, fibril formation was studied with the N-terminal fragments. In initial experiments, protein concentrations of 2 mM (30 mg/mL for N-WT) were routinely chosen and incubation was carried out at 37°C. In the first experiments, samples were analyzed for binding to thioflavin T (ThT), a histological dye that is commonly used for the detection of amyloid fibrils (LeVine III 1993). Binding of ThT to amyloid fibrils results in fluorescence at 490 nm upon excitation at 450 nm, whereas unbound ThT does not fluoresce at this wavelength. N-(+7)Ala showed ThT binding after an incubation time or lag phase of 7 d, indicating formation of amyloid-like fibrils. For N-WT, an incubation time of ca. 27 d was required for detecting ThT fluorescence. N-ΔAla did not show any interaction with ThT within an observation time of 100 d. A possible proteolytic degradation of the samples was excluded by SDS-PAGE analysis. For confirmation of amyloid-like structures, negatively stained electron microscopy (EM) of both N-(+7)Ala and N-WT was performed. Fibrils of both N-(+7)Ala and N-WT appeared as unbranched filaments with a diameter of 12–14 nm (Fig. 3). Experiments with the same protein concentrations at 20°C resulted in lag phases of ca. 12 weeks for N-(+7)Ala. To explore whether fibril formation may be caused by the N-terminal His-tag rather than the presence of poly-L-alanine

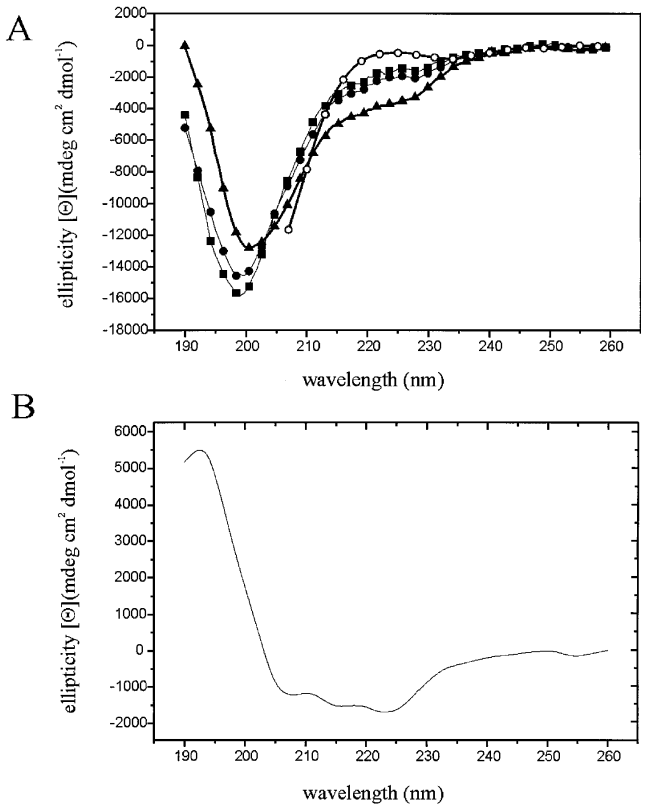

Figure 2. (*A*) Far-UV CD spectra of N-WT (filled circles), N-(Δ)Ala (filled squares), N-(+7)Ala (filled triangles), and N-WT unfolded in 6 M GdmCl (open circles). (*B*) Calculated difference spectrum of N-(+7)Ala and N-WT.

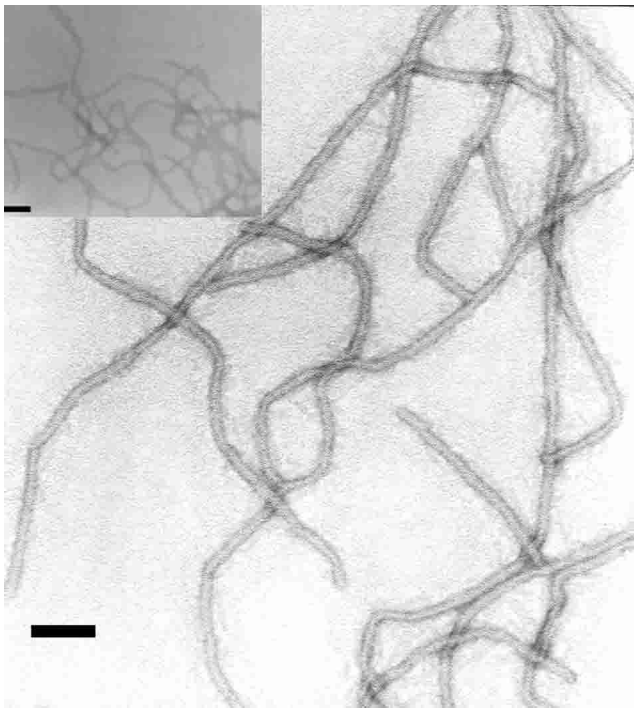

Figure 3. Negatively stained electron micrograph of uranyl acetate stained fibrils from N-(+7)Ala fibrils and N-(+7)Ala of which the His-tag had been removed before fibrillation by CNBr treatment (*inset*). Fibrils formed at a protein concentration of 1.4 mM at 37°C. The bar corresponds to 90 nm.

segments, the His-tag of N-(+7)Ala was removed by CNBr cleavage, controlled by SDS-PAGE and mass spectrometry (data not shown). Fibril formation with the cleaved variant was confirmed by ThT fluorescence (data not shown) and EM (Fig. 3, inset). As the morphology of fibrils formed by N-(+7)Ala after CNBr treatment did not differ from that of the untreated material, for convenience, all further experiments were carried out using the tagged variants. In further assays, the concentration dependence of fibril formation was tested at 37°C. Decreasing the concentration of N-(+7)Ala to 2 mg/mL (140 μM) resulted in a lag phase of ca. 35 d while at a concentration of 1 mg/mL (70 μM) even after 100 d no fibrils were detected. Lag phases of N-WT below concentrations of 30 mg/mL (2 mM) could not yet be determined due to incubation times longer than 50 d.

Next, seeding experiments were carried out. Addition of sonicated preformed fibrils of N-(+7)Ala to not yet fibrillized N-(+7)Ala led to a drastic decrease of the lag phases of N-(+7)Ala (Fig. 4A). Interestingly, upon seeding, fibril formation occurred even at those low protein concentrations where no fibrils were detected without seeds. The onset of fibril formation correlates with the length of the poly-L-alanine sequence (Fig. 4B): Comparison of the kinetics of fibril formation of N-WT and N-(+7)Ala by ANS fluorescence, a dye allowing a more sensitive detection of fibrils (Fig. 4C), shows that the lag phase of fibril formation is significantly prolonged in case of N-WT, even though the protein concentration was only half of that of N-(+7)Ala (Fig. 4B). Once the fibrils were formed they could not be

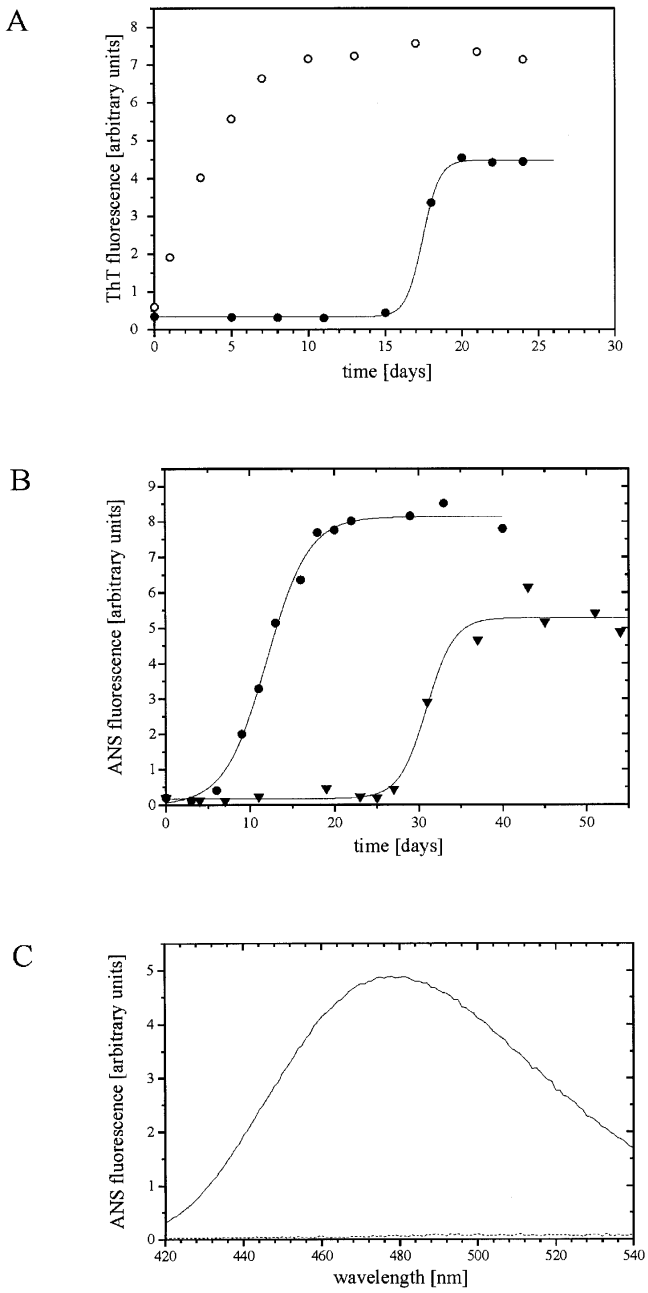

Figure 4. (*A*) Fibril formation of N-(+7)Ala evaluated by ThT fluorescence with seeds (open circles) and without seeds (filled circles). Fibrillation without seeds was fitted to a sigmoidal curve (solid line). Fibrils formed at 37°C and at a protein concentration of 560 μM (8 mg/mL). (*B*) Kinetics of fibril formation followed by ANS fluorescence. Fibrils of N-(+7)Ala (circles) and N-WT (triangles) formed at 37°C at protein concentrations of 1.1 mM and 2.2 mM, respectively. (*C*) Fibril-dependent ANS fluorescence. N-(+7)Ala was fibrillized at a concentration of 2 mM. ANS binding was tested before (dotted line) and after 6 weeks (solid line) incubation at 37°C.

solubilized by either a seven day incubation in 6 M guanidinium chloride, pH 2.0, or in 1% SDS at 60°C. Even treatment with proteinase K at a 1 : 1 ratio (w/w) over the period of 3 d did not result in degradation of the fibrils.

Structural characterization of the fibrils

Depending on the flanking amino acids, poly-L-alanine peptides can adopt either α-helical or β-sheet structures (Marqusee et al. 1989; Blondelle et al. 1997; Miller et al. 2001; Perutz et al. 2002). FTIR spectroscopy is suited to determine the secondary structure of polypeptides and proteins. FTIR spectra of fibrils formed and dissolved in D_2O were recorded and the shape of the amide I band in the wave number region from 1600–1700 cm^{-1} was analyzed by Fourier self-deconvolution for band frequencies characteristic for different secondary structural elements. The FTIR spectrum of fibrillar N-(+7)Ala significantly differed from that of soluble protein (Fig. 5). The spectrum of the fibrils revealed a high-frequency, low-intensity band at 1689 cm^{-1} and a high-intensity band with low frequency at 1621 cm^{-1}. The positions of these two bands are characteristic of antiparallel β-structures, and have also been observed in other amyloid fibrils (Byler and Susi 1986; Come et al. 1993; Nguyen et al. 1995; Conway et al. 2000; Kim et al. 2002). The band at 1673 cm^{-1} cannot be unambiguously assigned.

Discussion

Poly-L-alanine stretches occur in many proteins in nature, the most prominent example being spider silk fibroin (Gatesy et al. 2001). Recently, poly-L-alanine stretches have also been recognized in nucleic acid binding proteins involved in transcriptional regulation. With the exception of dragline spider fibroin, where poly-L-alanine stretches confer the tensile properties via β-sheet structures (Simmons et al. 1996; van Beek et al. 2000), to our knowledge nothing is known about the biochemical role of poly-L-alanines in other proteins. Recently, a number of publications described the linkage of poly-L-alanine expansions and genetic diseases (Muragaki et al. 1996; Mundlos et al. 1997; Brais et al. 1998; Brown et al. 1998; Goodman et al. 2000; Crisponi et al. 2001). Aside from PABPN1, all the other proteins affected by poly-L-alanine expansions are transcription factors. Truncation studies employing the transcription factors Engrailed (En) and Even-skipped (Eve) from *Drosophila* indicated that the Ala segments may mediate gene repression (Han and Manley 1993a,b). Still, the molecular mechanism awaits detailed analysis.

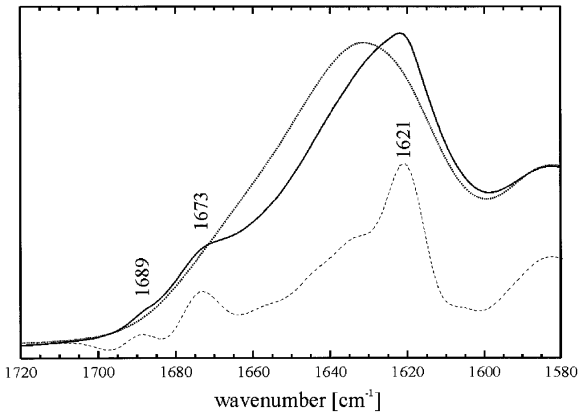

Figure 5. FTIR spectrum of N-(+7)Ala in D_2O at a concentration of 1.3 μM. (Solid line) Undeconvoluted spectrum of fibrillized N-(+7)Ala, (dashed line) deconvoluted spectrum of fibrillized N-(+7)Ala, (dotted line) spectrum of not yet-fibrillized material.

The only RNA processing protein that is known to cause a poly-L-alanine linked disease is PABPN1. To analyze the biophysical changes caused by extended poly-L-alanine stretches in PABPN1, the wild-type form was compared to both the form possessing the most extreme expansion occurring in human and an artificial form lacking the entire N-terminal Ala stretch. Biophysical analysis of the full-length proteins was hampered by their pronounced tendency to aggregate. No fibrillar structures could be detected by Congo red staining, FTIR, and EM. As aggregation could be retarded by buffers of high ionic strength, it seems likely that association of molecules occurs by ionic interaction of the oppositely charged N- and C-termini of PABPN1. In vivo, ionic interactions may be suppressed by binding to the poly(A) tail of mRNA, the natural target of PABPN1.

CD analysis of the full-length proteins under high salt conditions did not reveal any structural changes upon either extension of the natural 10 Ala segment to 17 Ala residues or complete deletion of the Ala stretch. It is possible that slight changes in the secondary structure content are not detectable in the context of the full-length proteins as CD analysis of the N-terminal fragments clearly showed an increase in α-helicity in N-(+7)Ala when compared to N-WT or N-ΔAla (Fig. 2B). These data suggest that the 10 Ala residues of N-WT may not be sufficient for the formation of an α-helical structure. Apparently, the capacity of poly-L-alanines to form α-helical structures depends on the protein context, as it has been recently shown that a short peptide consisting of $D_2A_{10}K_2$ adopted an α-helical structure independent of the pH (Perutz et al. 2002).

Investigation of the N-terminal segments for their potential to form amyloid-like fibrillar aggregates revealed that the addition of seven surplus Ala residues resulted in a pronounced propensity of fibril formation. With N-ΔAla no fibril formation was observed. Lag phases of fibril formation with ca. 7 days for N-(+7)Ala and ca. 27 days for N-WT at the same molar concentrations (2 mM) clearly depended on the length of the poly-L-alanine segment. Addition of seeds consisting of preformed fragmented fibrils by N-(+7)Ala drastically reduced the lag phases of both

N-(+7)Ala and N-WT. The reduced lag phases upon seeding very likely reflect elongation of fibrils added as seeds by those polypeptides that had undergone conformational changes from α-helical to β-sheet structures. The concentration of N-(+7)Ala—2 mg/mL (140 μM)—required for fibril formation appears unphysiologically high. Taking into consideration, however, that PABPN1 molecules have been observed on poly(A) tails like beads on a string, local concentrations of PABPN1 might be considerably in excess of those expected in solution (Keller et al. 2000). It has also been reported that self-association of the protein mediated by C-terminal sequences, which are not present in the N-terminal protein fragments studied here, favors the formation of insoluble aggregates, presumably fibrils, in vivo (Fan et al. 2001). Important for the discussion of the molecular scenario underlying the disease onset is the fact that once seeds have formed, elongation of fibrils can proceed at concentrations much lower than 2 mg/mL. As OPMD is an adult-onset disease that usually starts in the sixth decade, the late onset and progressive disease development may reflect long in vivo lag phases. Once the in vivo lag phase is overcome, a rapid deterioration could be caused by preformed fibrils that are likely to mould native structures into the alternative amyloid conformation.

Investigations of N-(+7)Ala fibrils by EM and upon thioflavin T or ANS binding revealed properties of amyloid fibrils. Biophysical investigation via FTIR indicated an antiparallel ß-structure. How can a polypeptide that has been shown to possess α-helical structures adopt a completely different structure concomitant with the loss of its solubility? It is now generally accepted that the long holding dogma "one sequence, one fold" has to be abandoned. A vast amount of literature on amyloid proteins documents that certain, maybe many, proteins can exist in alternate conformations. A striking example is myoglobin with most of its sequence arranged in α-helices. Under conditions that destabilize the native state, myoglobin adopts fibrillar structures consisting of β-sheets (Fändrich et al. 2001). Similarly, conformational changes can be induced in phosphosglycerate kinase (Damaschun et al. 1999). A popular hypothesis suggests that mutations that destabilize ΔG of the native state ease conformational conversions (Perutz et al. 2002). In contrast to many other proteins that possess a globular fold in the native state, that is, tertiary contacts, unfolding of the N-terminal segments of PABPN1 was uncooperative. It is therefore likely that this portion lacks tertiary contacts in the native state. For conversion into fibrils probably only one α-helix has to unfold, and no tertiary contacts have to be broken as proposed for the prion protein, human cystatin C, or human stefin (Janowski et al. 2001; Knaus et al. 2001; Staniforth et al. 2001). Conformational transitions of α-helical into β-sheet structures have also been reported for prion protein peptides (Nguyen et al. 1995).

Trinucleotide-based illnesses have received a lot of attention in the recent past. So far, the major focus has been on poly-L-glutamine repeats as those present in Huntington's disease and spinocerebellar ataxia type 1. A recent publication discusses the possibility that CAG repeats leading to the extension of poly-L–glutamine are accompanied with rare transcriptional or translational + 2 frame shifts that would result in poly-L–alanine sequences (Gaspar et al. 2000). A number of papers have been published that report poly-L–alanine based severe disorders in humans. The molecular causes underlying these diseases are just about to be elucidated. By an OPMD cell culture test system, evidence was obtained that protein aggregation and cell death elicited by poly-L–alanine sequences are linked (Bao et al. 2002). Here, we could demonstrate in vitro fibril formation caused by poly-L–alanine expansions with the N-terminal fragment of PABPN1. Clearly, additional Ala residues ease the conversion into the alternate conformation, although the wild-type fragment could also adopt the amyloid-like state. Fibril formation by the wild-type protein is not entirely unexpected, as the extension of the polyalanine sequence even by a single residue results in a recessive OPMD phenotype, that is, presumably fibril formation, in humans (Brais et al. 1998). Description of the molecular steps that are likely to underly the diseases may eventually lead to the discovery of compounds that impair fibril formation and could serve as therapeutics.

Materials and methods

Cloning of PABPN1 and N-terminal domains of PABPN1

A partially synthetic gene for bovine PABPN1 (GenBank acc. No. X89969, Kühn et al. 2003) was recloned into pET11a (Novagen) via insertion into *Nde*I and *Bam*HI sites. For construction of a variant of PABPN1 carrying seven surplus Ala residues (PABPN1-(+7)Ala), additional codons for Ala were introduced via the QuickChange mutagenesis kit (Stratagene). For mutagenesis, primers 5′ CAC CAT ATG GCG GCA GCC GCG GCG GCA GCG GCA GCA GCA 3′ and the corresponding complementary oligonucleotide were used. Here, the template DNA was in vector pET8c with an N-terminal His-tag. For expression, the mutant DNA was inserted via *Nde*I and *Bam*HI sites into vector pET11a. For constructing a variant of PABPN1 lacking the 10 N-terminal Ala residues following the start methionine (PABPN1-[Δ]Ala), DNA encoding the wild-type sequence was mutagenized via PCR with the mutagenesis primer 5′ CAC ATA TGG CAG GAG GAA GAG GAT CAG G 3′ and the back primer 5′ GCT AGT TAT TGC TCA GCG GTG G 3′. For obtaining N-terminal fragments of PABPN1, DNA was PCR amplified with the primers 5′ CGA AAT TAA TAC GAC TCA C 3′ and 5′ CCA GGA TCC TTT ATC GAG CTT TTA TTG CTT C 3′. PCR products were digested with *Nde*I and *Bam*HI and ligated into pET15b because efficient expression of the fragments was achieved only in the context of the N-terminal His-tag provided by the vector. All constructs were verified by DNA sequencing.

Recombinant protein expression and purification

PABPN1-WT, PABPN1-(+7)Ala, and PABPN1-ΔAla were expressed in *Escherichia coli* (strain BL21[DE3]) using the T7 system from Novagen (pET11a). Cells were induced with 1 mM IPTG at OD_{600} = 0.5–0.8 and harvested 3 h after induction. Cells were resuspended in 50 mM MOPS, pH 6.5, 5 mM EDTA. After cell disruption by high pressure dispersion the supernatant was diluted 1 : 1 with 3 M KCl, 50 mM MOPS, pH 6.5, 50 mM $Na_4P_2O_7$, 1 mM DTT, 5 mM EDTA, resulting in a final KCl concentration of 1.5 M. $(NH_4)_2SO_4$ was added to a final concentration of 1.75 M. The solution was incubated for 1 h at 4°C and centrifuged. The supernatant was subjected to hydrophobic interaction chromatography (HIC). A Fractogel EMD Phenyl (S) column (Novagen) was equilibrated with 1.5 M KCl, 1.75 M $(NH_4)_2SO_4$, 50 mM MOPS, pH 6.5, 50 mM NaP_2O_7, 1 mM DTT, 5 mM EDTA. After loading, washing was performed with the equilibration buffer lacking $(NH_4)_2SO_4$. To prevent ionic aggregation during elution by the decreasing ionic strength, elution was performed by a continuous L-arginine gradient from 0 M to 1 M L-arginine in 50 mM MOPS, pH 6.5, 50 mM NaP_2O_7, 1 mM DTT. Fractions containing PABPN1 were dialyzed in the presence of a strong cation exchange gel material (Fractogel EMD SO^{3-} (S) (Novagen) against 50 mM MOPS, pH 6.5, 1 mM DTT, 5 mM EDTA. The slurry with bound PABPN1 was transferred into a void column and eluted by a KCl gradient (0 M–1.5 M) with 1.5 M KCl, 50 mM MOPS, pH 6.5, 50 mM NaP_2O_7, 1 mM DTT, 5 mM EDTA. PABPN1 containing fractions were loaded on a gel filtration column (HiLoad Superdex 75 prep grade, Amersham Pharmacia Biotech) equilibrated with 1.5 M KCl, 50 mM MOPS, pH 6.5, 50 mM NaP_2O_7, 1 mM DTT, 5 mM EDTA. Protein purity was controlled by SDS-PAGE, and protein concentrations were determined by UV-absorption.

N-terminal fragments of wild-type (N-WT), the variant containing seven additional Ala residues (N-(+7)Ala) and that with the deletion of all 12 N-terminal Ala residues including Gly^{11} (N-ΔAla) (Fig. 1B) could only be expressed as fusions with N-terminal His-tags using the vector pET15b. Harvested cells were resuspended in 50 mM Tris, pH 8.0, 5 mM EDTA. After cell disruption by high pressure dispersion, debris was sedimented by centrifugation. The supernatant was incubated for 5 min at 80°C and centrifuged. The supernatant was diluted 1 : 1 with iso-propanol, centrifuged, and the supernatant was again diluted 1 : 1 with 50 mM Tris, pH 8.0, 5 mM EDTA and loaded on a Q Sepharose Fast Flow (Amersham Pharmacia Biotech) equilibrated with 50 mM Tris, pH 8.0, 5 mM EDTA. Washing was performed with 100 mM NaCl, 50 mM Tris, pH 8.0, 5 mM EDTA. Elution was achieved by a salt step with 200 mM NaCl, 50 mM Tris, pH 8.0, 5 mM EDTA. Fractions with N-terminal fragments were pooled and further purified by gel filtration on a HiLoad Superdex 75 prep grade (Amersham Pharmacia Biotech) equilibrated with 50 mM Tris, pH 8.0, 5 mM EDTA. CNBr cleavage of N-(+7)Ala was performed to test whether the His-tag may cause fibril formation. CNBr cleavage was possible because the start methionine of PABPN1 separating the tag from the PABPN1 sequence is the only internal methionine of this construct. Before cleavage, N-(+7)Ala was dialyzed against water and then lyophilized. Lyophilized material was redissolved in 70% formic acid purged with N_2 containing at least 50-fold molar excess of CNBr over methionine in the protein preparation. The mixture was incubated in the dark for 24 h at room temperature. Subsequently, the material was dialyzed against water, and then against 50 mM Tris, 20 mM imidazol, 300 mM NaCl, pH 8.0. The His-tag was subsequently removed by IMAC.

Circular dichroism (CD)

CD spectra were recorded at 20°C on an Aviv model 62A DS spectropolarimeter. PABPN1-WT (110 μg/mL), PABPN1-(+7)Ala (210 μg/mL), PABPN1-ΔAla (160 μg/mL) were measured in 1-mm quartz cuvettes over the range of 210–260 nm in 1.5 M KCl, 50 mM MOPS, pH 6.5, 50 mM NaP_2O_7. N-WT (1.0 mg/mL), N-(+7)Ala (1.0 mg/mL) and N-ΔAla (0.75 mg/mL) were analyzed in 0.1-mm quartz cuvettes over the range of 190–260 nm in 5 mM MOPS, pH 6.5. The spectrum of denatured N-WT was recorded in 6 M GdmCl, 5 mM MOPS, pH 6.5, at a protein concentration of 1 mg/mL in a 1-mm quartz cuvette. The CD spectra shown represent the average of 10 scans run in 1-nm intervals.

Fibril analysis by ThT (thioflavine T) and ANS (1-anilino-8-naphtalensulfonate)

For fluorescence measurements with the dye thioflavine T (ThT), samples were diluted to final concentrations of 50–200 μg/mL (3.5–13.6 μM) with 5 μM ThT in 150 mM NaCl, 5 mM KH_2PO_4 pH 7.4. Fluorescence spectra were recorded upon excitation at 450 nm from 460 to 600 nm in a Jobin Yvon Spex Fluoromax2 at 20°C. Experiments were performed in 1-cm quartz cuvettes with excitation and emission slit widths of 5 nm. Fibril-dependent ANS fluorescence was recorded at an emission wavelength of 480 nm upon excitation at 370 nm. For fluorescence measurements, samples were diluted to a final concentration of 5 μM with 250 μM ANS in 150 mM NaCl, 5 mM KH_2PO_4, pH 7.4.

Seed preparation

For seeding experiments, preformed fibrils were centrifuged at 70,000 × *g*, washed with 5 mM KH_2PO_4, pH 7.4, 150 mM NaCl, and sonicated. Seeding was performed by adding 0.4% (w/v) seeds to monomeric protein solutions.

Electron microscopy (EM)

Samples were diluted with distilled water to final concentrations of 500 μg/mL and spotted on a carbonized copper grid. The grids were incubated for 3 min, air dried, washed with water, and again air dried. The fibrils were negatively stained with 2% (w/v) uranyl acetate and visualized in a Zeiss EM 900 electron microscope operating at 80 kV.

Fourier transform infrared spectroscopy (FTIR)

Infrared spectra were recorded on a Bruker IFS 66 FTIR spectrometer equipped with a MCT detector under continuous purge with dry nitrogen. For FTIR spectroscopy, fibrils of N-(+7Ala) (20 mg/mL) were allowed to form in D_2O. Fibrillized material was then lyophilized and dissolved in D_2O before spectroscopy. For comparison, soluble protein was lyophilized and subsequently dissolved in D_2O to replace protons by deuterium. Samples were prepared between two CaF_2 windows separated by a Teflon spacer of 50 μm. D_2O spectra were recorded under the same conditions and subtracted from the spectra obtained from fibrils and soluble N-(+7)Ala.

Acknowledgments

We thank Gerd Hause for performing electron microscopy, Angelika Schierhorn for mass spectrometry, and Peter Rücknagel for determination of peptide sequences. We are grateful to Milton Stubbs for polishing the English. The project was supported by the Deutsche Forschungsgemeinschaft (DFG) (SFB 610 to E.S. and E.W.). We also thank the Fonds der Chemischen Industrie and the European Union (E.W.) for support.

The publication costs of this article were defrayed in part by payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 USC section 1734 solely to indicate this fact.

References

Bao, Y.P., Cook, L.J., O'Donovan, D., Uyama, E., and Rubinsztein, D.C. 2002. Mammalian, yeast, bacterial, and chemical chaperones reduce aggregate formation and death in a cell model of oculopharyngeal muscular dystrophy. *J. Biol. Chem.* **277:** 12263–12269.

Blondelle, S.E., Forood, B., Houghten, R.A., and Pérez-Payá, E. 1997. Polyalanine-based peptides as models for self-associated β-pleated-sheet complexes. *Biochemistry* **36:** 8393– 8400.

Brais, B., Bouchard, J.P., Xie, Y.-G., Rochefort, D.L., Chretién, N., Tomé, F.M.S., Lafrenière, R.G., Rommens, J.M., Uyama, E., Nohira, O., et al. 1998. Short GCG expansions in the PABP2 gene cause oculopharyngeal muscular dystrophy. *Nat. Genet.* **18:** 164–167.

Brown, S.A., Warburton, D., Brown, L.Y., Yu, C., Roeder, E.R., Stengel-Rutkowski, S., Hennekam, R.C.M., and Muenke, M. 1998. Holoprosencephaly due to mutations in ZIC2, a homologue of *Drosophila odd-paired*. *Nat. Genet.* **20:** 180–183.

Byler, D.M. and Susi, H. 1986. Examination of the secondary structure of proteins by deconvolved FTIR spectra. *Biopolymers* **25:** 469–487.

Calado, A., Tomé, F.M.S., Brais, B., Rouleau, G.A., Kühn, U., Wahle, E., and Carmo-Fonseca, M. 2000. Nuclear inclusions in oculopharyngeal muscular dystrophy consist of poly(A) binding protein 2 aggregates which sequester poly(A) RNA. *Hum. Mol. Genet.* **9:** 2321–2328.

Come, J.H., Fraser, P.E., and Lansbury Jr., P.T. 1993. A kinetic model for amyloid formation in the prion diseases: Importance of seeding. *Proc. Natl. Acad. Sci.* **90:** 5959–5963.

Conway, K.A., Harper, J.D., and Lansbury Jr., P.T. 2000. Fibrils formed in vitro from α-synuclein and two mutant forms linked to Parkinson's disease are typical amyloid. *Biochemistry* **29:** 2552–2563.

Crisponi, L., Deiana, M., Loi, A., Chiappe, F., Uda, M., Amati, P., Bisceglia, L., Zelante, L., Nagaraja, R., Porcu, S., et al. 2001. The putative forkhead transcription factor FOXL2 is mutated in blepharophimosis/ptosis/epicanthus inversus syndrome. *Nat. Genet.* **27:** 159–166.

Cummings, C.J. and Zoghbi, H.Y. 2000. Fourteen and counting: Unraveling trinucleotide repeat diseases. *Hum. Mol. Genet.* **9:** 909–916.

Damaschun, G., Damaschun, H., Gast, K., and Zirwer, D. 1999. Proteins can adopt totally different folded conformations. *J. Mol. Biol.* **291:** 715–725.

Fan, X., Dion, P., Laganiere, J., Brais, B., and Rouleau, G.A. 2001. Oligomerization of polyalanine expanded PABPN1 facilitates nuclear protein aggregation that is associated with cell death. *Hum. Mol. Genet.* **10:** 2341–2351.

Fändrich, M., Fletcher, M.A., and Dobson, C.M. 2001. Amyloid fibrils from muscle myoglobin. *Nature* **410:** 165–166.

Galant, R. and Carroll, S.B. 2002. Evolution of a transcriptional repression domain in an insect Hox protein. *Nature* **415:** 910–913.

Gaspar, C., Jannatipour, M., Dion, P., Laganière, J., Sequeiros, J., Brais, B., and Rouleau, G.A. 2000. CAG tract of MJD-1 may be prone to frameshifts causing polyalanine accumulation. *Hum. Mol. Genet.* **9:** 1957–1966.

Gatesy, J., Hayashi, C., Motriuk, D., Woods, J., and Lewis, R. 2001. Extreme diversity, conservation, and convergence of spider silk fibroin sequences. *Science* **291:** 2603–2605.

Goodman, F.R., Bacchelli, C., Brady, A.F., Brueton, L.A., Fryns, J.-P., Mortlock, D.P., Innis, J.W., Holmes, L.B., Donnenfeld, A.E., Feingold, M., et al. 2000. Novel HOXA13 mutations and the phenotypic spectrum of Hand-Foot-Genital Syndrome. *Am. J. Hum. Genet.* **67:** 197–202.

Han, K. and Manley, J.L. 1993a. Functional domains of the *Drosophila* Engrailed protein. *EMBO J.* **12:** 2723–2733.

———. 1993b. Transcriptional repression by the *Drosophila* Even-skipped protein: Definition of a minimal repression domain. *Genes & Dev.* **7:** 491–503.

Janowski, R., Kozak, M., Jankowska, E., Grzonka, Z., Grubb, A., Abrahamson, M., and Jaskolski, M. 2001. Human cystatin C, an amyloidogenic protein, dimerizes through three-dimensional domain swapping. *Nat. Struct. Biol.* **8:** 316–320.

Keller, R.W., Kühn, U., Aragón, M., Bornikowa, L., Wahle, E., and Bear, D.G. 2000. The nuclear poly(A) binding protein, PABP2, forms an oligomeric particle covering the length of the poly(A) tail. *J. Mol. Biol.* **297:** 569–583.

Kerwitz, Y., Kühn, U., Lilie, H., Knoth, A., Scheuermann, T., Friedrich, H., Schwarz, E., and Wahle, E. 2003. Stimulation of poly(A) polymerase through a direct interaction with the nuclear poly(A) binding protein allosterically regulated by RNA. *EMBO J.* **22:** 3705–3714.

Kim, Y.-J., Noguchi, S., Hayashi, Y.K., Tsukahara, T., Shimizu, T., and Arahata, K. 2001. The product of an oculopharyngeal muscular dystrophy gene, poly(A)-binding protein 2, interacts with SKIP and stimulates muscle-specific gene expression. *Hum. Mol. Genet.* **10:** 1129–1139.

Kim, Y.-S., Randolph, T.W., Stevens, F.J., and Carpenter, J.F. 2002. Kinetics and energetics of assembly, nucleation, and growth of aggregates and fibrils for an amyloidogenic protein. *J. Biol. Chem.* **277:** 27240–27246.

Klement, I.A., Skinner, P.J., Kaytor, M.D., Yi, H., Hersch, S.M., Clark, H.B., Zoghbi, H.Y., and Orr, H.T. 1998. Ataxin-1 nuclear localization and aggregation: Role in polyglutamine-induced disease in *SCA1* transgenic mice. *Cell* **95:** 41–53.

Knaus, K.J., Morillas, M., Swietnicki, W., Malone, M., Surewicz, W.K., and Yee, V.C. 2001. Crystal structure of the human prion protein reveals a mechanism for oligomerization. *Nat. Struct. Biol.* **8:** 770–774.

Kühn, U., Nemeth, A., Meyer, S., and Wahle, E. 2003. The RNA binding domains of the nuclear poly(A) binding protein. *J. Biol. Chem.* **278:** 16916–16925.

LeVine III, H. 1993. Thioflavine T interaction with synthetic Alzheimer's disease β-amyloid peptides: Detection of amyloid aggregation in solution. *Protein Sci.* **2:** 404–410.

Marqusee, S., Robbins, V.H., and Baldwin, R.L. 1989. Unusually stable helix formation in short alanine-based peptides. *Proc. Natl. Acad. Sci.* **86:** 5286–5290.

Miller, J.S., Kennedy, R.J., and Kemp, D.S. 2001. Short, solubilized polyalanines are conformational chameleons: Exceptionally helical if N- and C-capped with helix stabilizers, weakly to moderately helical if capped with rigid spacers. *Biochemistry* **40:** 305–309.

Mundlos, S., Otto, F., Mundlos, C., Mulliken, J.B., Aylsworth, A.S., Albright, S., Lindhout, D., Cole, W.G., Henn, W., Knoll, J.H.M., et al. 1997. Mutations involving the transcription factor CBFA1 cause cleidocranial dysplasia. *Cell* **89:** 773–779.

Muragaki, Y., Mundlos, S., Upton, J., and Olsen, B.R. 1996. Altered growth and branching patterns synpolydactyly caused by mutations in HOXD13. *Science* **272:** 548–551.

Nguyen, J., Baldwin, M.A., Cohen, F.E., and Prusiner, S.B. 1995. Prion protein peptides induce α-helix to β-sheet conformational transitions. *Biochemistry* **34:** 4186–4192.

Perutz, M.F., Pope, B.J., Owen, D., Wanker, E.E., and Scherzinger, E. 2002. Aggregation of proteins with expanded alanine repeats of the glutamine-rich and asparagine-rich domains of Sup35 and of the amyloid β-peptide of amyloid plaques. *Proc. Natl. Acad. Sci.* **99:** 5596–5600.

Ronshaugen, M., McGinnis, N., and McGinnis, W. 2002. Hox protein mutation and macroevolution of the insect body plan. *Nature* **415:** 914–917.

Simmons, A.H., Michal, C.A., and Jelinski, L.W. 1996. Molecular orientation and two-component nature of the crystalline fraction of spider dragline silk. *Science* **271:** 84–87.

Staniforth, R.A., Giannini, S., Higgins, L.D., Conroy, M.J., Hounslow, A.M., Jerala, R., Craven, C.J., and Waltho, J.P. 2001. Three-dimensional domain swapping in the folded and molten-globule states of cystatins, an amyloid-forming structural superfamily. *EMBO J.* **20:** 4774–4781.

Tomé, F.M.S., Chateau, D., Helbling-Leclerc, A., and Fardeau, M. 1997. Morphological changes in muscle fibers in oculopharyngeal muscular dystrophy. *Neuromuscul. Disord.* **7:** 63–69.

Trottier, Y., Lutz, Y., Stevanin, G., Imbert, G., Devys, D., Cancel, G., Saudou, F., Weber, C., David, G., Tora, L., et al. 1995. Polyglutamine expansion as a pathological epitope in Huntington's disease and four dominant cerebellar ataxias. *Nature* **387:** 403–406.

Van Beek, J.D., Beaulieu, L., Schäfer, H., Demura, M., Asakura, T., and Meier, B.H. 2000. Solid-state NMR determination of the secondary structure of *Samia cynthia ricini* silk. *Nature* **405:** 1077–1079.

Wahle, E. 1991. A novel poly(A)-binding protein acts as a specificity factor in the second phase of messenger RNA polyadenylation. *Cell* **66:** 759–768.

Wahle, E. and Rüegsegger, U. 1999. 3′-end processing of pre-mRNA in eukaryotes. *FEMS Microbiol. Rev.* **23:** 277–295.