Functional interactions as a survival strategy against abnormal aggregation

Laura Masino,* Giuseppe Nicastro,* Lesley Calder,* Michele Vendruscolo,† and Annalisa Pastore*,[1]

*Medical Research Council National Institute for Medical Research, London, UK; and †Department of Chemistry, University of Cambridge, Cambridge, UK

ABSTRACT **Protein aggregation is under intense scrutiny because of its role in human disease. Although increasing evidence indicates that protein native states are highly protected against aggregation, the specific protection mechanisms are poorly understood. Insight into such mechanisms can be gained through study of the relatively few proteins that aggregate under native conditions. Ataxin-3, the protein responsible for Spinocerebellar ataxia type 3, a polyglutamine expansion disease, represents one of such examples. Polyglutamine expansion is central for determining solubility and aggregation rates of ataxin-3, but these properties are profoundly modulated by its N-terminal Josephin domain. This work aims at identifying the regions that promote Josephin fibrillogenesis and rationalizing the mechanisms that protect Josephin and nonexpanded ataxin-3 from aberrant aggregation. Using different biophysical techniques, aggregation propensity predictions and rational design of amino acid substitutions, we show that Josephin has an intrinsic tendency to fibrillize under native conditions and that fibrillization is promoted by two solvent-exposed patches, which are also involved in recognition of natural substrates, such as ubiquitin. Indeed, designed mutations at these patches or substrate binding significantly reduce Josephin aggregation kinetics. Our results provide evidence that protein nonpathologic function can play an active role in preventing aberrant fibrillization and suggest the molecular mechanism whereby this occurs in ataxin-3.—Masino, L., Nicastro, G., Calder, L., Vendruscolo, M., Pastore, A. Functional interactions as a survival strategy against abnormal aggregation. *FASEB J.* 25, 45–54 (2011). www.fasebj.org**

Key Words: ataxin-3 · fibrillogenesis · native-like aggregation · neurodegeneration · polyglutamine proteins · structure

An increasing number of human neurodegenerative diseases has been associated with the formation of ordered aggregates mainly formed by misfolded proteins (1, 2). This phenomenon does not seem to be restricted to a small number of examples, since the number of peptides and proteins that, under appropriate conditions, have been shown to form fibers is much greater than the number of sequences known to be implicated in human disease, thus leading to the hypothesis that most proteins can misfold and aggregate into fibers (3). This view, together with the medical (2), biological (4), and biotechnological (5) interest associated with protein aggregation, raises the important question of which characteristics are responsible for preventing or promoting fibrillogenesis and whether it is possible to predict, and ultimately modulate, the fibrillogenic propensity of different amino acid sequences. A set of principles has been identified. First, the information defining the fibrillogenesis propensity appears to be stored within the amino acid sequences (6). Second, sequences must have some tendency to adopt structural motifs compatible with the ordered or semiordered structures observed in the repetitive assembly proper of fibers, such as β-rich tertiary structure (7, 8). Finally, fibrillogenic sequences need to be at least partially if not completely accessible to solvent (9–11). This feature is common in peptides (such as the Alzheimer's Aβ peptide) and in intrinsically unstructured proteins (such as α-synuclein, responsible for Parkinson's disease) but is less frequently observed in globular proteins, in which potential β-prone sequences are usually well buried in the protein core even in cases of proteins associated to misfolding diseases. *In vitro* aggregation of these proteins often requires destabilization of the native fold, for instance using organic solvents (*e.g.*, TFE), heat, or mechanical stress.

In contrast, a small but increasing number of protein examples seem to form fibrils with no requirement of destabilization of the native state. These are globular, well-folded proteins that undergo misfolding and fibrillogenesis also under native-like conditions (12). In these cases, new principles have to be established to

[1] Correspondence: MRC National Institute for Medical Research, The Ridgeway, Mill Hill, London NW7 1AA, U.K. E-mail: apastor@nimr.mrc.ac.uk

doi: 10.1096/fj.10-161208

This article includes supplemental data. Please visit *http://www.fasebj.org* to obtain this information.

rationalize the elements responsible for this behavior. These observations also pose the essential question of which mechanisms have evolved to prevent aberrant aggregation and to enable correct protein function. Quality control pathways that maintain protein homeostasis in the cell, such as the heat-shock response and the unfolded protein response, are currently under intense investigation (13). Bioinformatics studies have also recently suggested that an additional contribution to protection mechanisms against aggregation may be encoded within protein sequences themselves, through specific residue patterns that promote stabilizing interactions and protect solvent-exposed aggregation-prone regions (14). However, the general mechanisms that protect protein native states from harmful aggregation are not yet known in great detail. Study of the proteins that undergo aggregation under native-like conditions is thus central to a better understanding of these mechanisms.

One such example is ataxin-3, the protein responsible for the neurodegenerative disease spinocerebellar ataxia type 3 or Machado-Joseph disease (SCA3/MJD). SCA3, which is a member of the class of misfolding and aggregation pathologies that also includes Alzheimer's and Parkinson's diseases, is caused by the anomalous expansion of glutamine repeats within ataxin-3 above the pathological threshold of ~55 consecutive glutamine residues (15). The presence of intracellular aggregates of ataxin-3 has been observed in SCA3 patients and has been directly linked to the pathology (16). Because of their strong connection with SCA3 pathogenesis, both the aggregation properties and the biological function of ataxin-3 have been intensively studied (17, 18 and references therein). The only folded region of ataxin-3 is an N-terminal domain known as Josephin, whereas the C terminus of the protein, containing the polyglutamine (polyQ) tract, is mostly unstructured (19). Ataxin-3 is ubiquitously expressed and seems to be involved in the cellular pathways of transcriptional repression (20, 21) and protein degradation through the ubiquitin (Ub)/proteasome complex (22–24). It interacts with polyUb chains of $\geq$4 subunits through the Josephin domain, and Ub interaction motifs (UIMs) in the C-terminal region. By determining the 3-dimensional structure of Josephin and characterizing its enzymatic activity, we confirmed that ataxin-3 is a cysteine protease with Ub hydrolase activity (25, 26).

Aggregation of full-length nonexpanded (that is, containing <40 glutamine repeats) ataxin-3 was initially studied under destabilizing conditions using thermal, pressure-induced, or chemical denaturation (27–30), but it was subsequently reported that the protein also forms fibers under native conditions (31–34). We showed that the Josephin domain plays a pivotal role in modulating aggregation of full-length nonexpanded ataxin-3, by dictating its stability and self-association properties (29). Further studies resulted in consistent conclusions, suggesting that the early stages of ataxin-3 fibrillization are triggered by Josephin self-association and that the nucleating core consists mainly of Josephin (32–34). These reports provided direct evidence that the polyQ region is not involved in the first steps of ataxin-3 aggregation, as expanded ataxin-3 and variants that lack the polyQ tract share the same kinetic mechanism (32). The characterization of the properties of the Josephin domain is thus clearly a key step toward a detailed understanding of ataxin-3 aggregation.

We previously studied the Josephin domain in the absence of the C-terminal portion of ataxin-3 and showed that it aggregates and forms fibers on thermal stress (29). Dimerization of Josephin during its purification under nondenaturing conditions and fibrillization under partially denaturing conditions was also observed in a subsequent report (34). The high tendency of Josephin to aggregate even at room temperature forced us to use new samples after each nuclear magnetic resonance (NMR) data collection set (which typically lasts 3–6 d). These observations suggested that Josephin could be at least partially responsible for the unique properties of the full-length protein also under native conditions.

Here, we present a systematic analysis of the aggregation behavior of Josephin. The work aims at identifying the regions that promote Josephin fibrillogenesis and at rationalizing the mechanisms that protect Josephin, and consequently nonexpanded ataxin-3, from aberrant aggregation. Using complementary techniques, including NMR spectroscopy, size-exclusion chromatography (SEC), electron microscopy (EM), predictions of aggregation propensity, and rational design of amino acid substitutions capable of altering this propensity, we show that the regions on the surface of Josephin involved in fibrillogenesis comprise two highly conserved aggregation-prone patches that also allow substrate recognition of the enzyme. They do not need to be previously disrupted since we demonstrate that isolated Josephin, as full-length ataxin-3, also aggregates spontaneously without the need of denaturing conditions. Substrate binding or designed mutations within the patches involved in target recognition significantly slow aggregation kinetics. Our work provides a new perspective for understanding how evolution might result, for functional reasons, in regions characterized by a significant fibrillogenic tendency, and suggests a protection mechanism through which aggregation can be modulated or even prevented.

MATERIALS AND METHODS

Protein expression and purification

Josephin domains were expressed in *Escherichia coli* BL21(DE3) and purified as described previously (35). Uniformly ^{15}N-labeled samples were obtained by growing *E. coli* in synthetic medium containing $^{15}NH_4Cl$ as the sole nitrogen source. To prevent cysteine oxidation, protein purification and all experiments were performed in the presence of 2–5 mM β-mercaptoethanol or 5–10 mM Tris(2-carboxyethyl)phosphine hydrochloride (TCEP; Sigma, St. Louis, MO, USA). The clone of human

Ub, a generous gift of Paul Driscol (National Institute for Medical Research, London, UK), was overexpressed in *E. coli* strain BL21 as an untagged construct using a modified pET vector. The protein was purified by ion-exchange chromatography using Q-Sepharose matrix (GE Healthcare, New York, NY, USA) equilibrated with 50 mM Tris-HCl, pH 8.0. The flow-through containing Ub was further purified by SEC using G-100 Sephadex resin (Sigma-Aldrich). Protein concentrations were determined using UV absorption with calculated extinction coefficients at 280 nm of 25,440 M^{-1} cm^{-1} for wild-type Josephin and Jos_I77KQ78K, 19940 M^{-1} cm^{-1} for Jos_W87K, and 1280 M^{-1} cm^{-1} for Ub. The purity of the samples (always >95%) was assessed by SDS-PAGE and mass spectrometry.

NMR spectroscopy

NMR data were recorded using Varian Inova spectrometers (Varian Inc., Walnut Creek, CA, USA) operating at 600 and 800 MHz ^{1}H frequencies equipped with a pulsed field *z*-gradient unit. Aggregation kinetics were monitored by acquiring 2-D ^{1}H-^{15}N HSQC spectra at 38°C at regular time intervals. The NMR signal was obtained by calculating the peak volumes corresponding to backbone amide groups in 2-D 1H-15N HSQC spectra. Wild-type and mutant Josephin samples were in 20 mM sodium phosphate buffer at pH 6.5, at a concentration of 250 μM. Experiments performed in the presence of Ub were recorded at a 3:1 Ub:Josephin ratio.

Aggregation time course analysis

Samples of wild-type and mutant Josephin domains were incubated at 37°C without shaking at a concentration of 14 μM. When Ub was present, Ub concentration was 593 μM, thus providing >90% saturation of both binding sites in wild-type Josephin (26). The buffer used was 20 mM sodium phosphate, pH 6.5, 10 mM TCep, 0.025% NaN_3. At the specified time points, aliquots were removed from the same stock solutions for SEC, SDS-PAGE, and EM measurements. Monodispersion of the samples prior to aggregation was confirmed by NMR, SEC, and EM data. SDS-PAGE showed that the proteins remained intact throughout the monitored incubation period.

SEC

Analytical SEC experiments were performed on a Superdex 75 10/300 GL Tricorn column (GE Healthcare) equilibrated in PBS (2 mM phosphate, pH 7.5, 5 mM KCl, and 140 mM NaCl). Aliquots of Josephin stock solutions incubated at 37°C were injected into the column at different time points. The signal was monitored at 215 and 280 nm. Peak areas were calculated using GE Healthcare-supplied software.

EM measurements

Diluted Josephin samples were applied to carbon-coated, glow-discharged grids and stained with 1% sodium silicotungstate (pH 7). Protein concentration was 2–3 μM. The grids were viewed under minimum-dose, accurate defocus conditions with a Jeol 1200EX instrument operated at 100 kV.

Identification of aggregation-prone regions

To predict the fibrillogenic segments of Josephin, we considered the bioinformatics methods Aggrescan, Tango, Waltz, Pasta, and Zyggregator, all accessible through Web servers (14, 36–40). Aggrescan data were normalized by running the program on all sequences available in Swiss-Prot and taking the resulting highest a4vAHS value as the maximum aggregation propensity value. For Waltz, we used the "high-specificity" threshold, providing less false positives (36). Aggregation-prone surfaces were calculated using the Zyggregator method, as discussed by Pechmann *et al.* (14).

RESULTS

Identification of aggregation-prone regions in Josephin

To identify fibrillogenic fragments, we first resorted to *in silico* predictions that could be validated experimentally. The amino acid sequence of Josephin was analyzed by different bioinformatics software (Aggrescan, Tango, Waltz, Pasta, and Zyggregator; refs. 14, 36–40). These methods are based on different approaches, which range from threading of a sequence into parallel or antiparallel fibrillar β-structures, to calculating the enthalpic and entropic costs associated with the conformational transition between folded and aggregated structures, to using energy functions, which take into account the physicochemical properties of a given amino acid sequence. All of these methods analyze sequence information without preknowledge of the structure adopted in the nonaggregated/native state and assume the information to be sequential. Although conceptually very different, all methods consistently identified the region comprising residues 73–96 of the sequence as having a high aggregation propensity (**Fig. 1*A***). Two additional stretches predicted by 4 of the methods to have high aggregation propensities are formed by residues 159–167 and 144–154 (the latter with a lower score). It was interesting to note that, although far from each other along the sequence, when mapped onto the 3-D Josephin structure, the first and second stretches are close together on the surface and contribute to form an exposed hydrophobic patch, comprising residues I77, I92, L93, and F163 (25) (Fig. 1*B*). This patch maps to the groove formed between the main body of the domain and an α-helical hairpin that is directly contiguous to the catalytic triad of the enzyme. The surface is essential for recognition of Ub, the preferential Josephin substrate target, and for positioning it for cleavage (25, 26). The first stretch indicated by the predictions also comprises residues V86 and W87. These residues are spatially close to Y27 and F28 and form a distinct aromatic patch that is solvent exposed (Fig. 1*B*). This surface coincides with a distinct binding site, which can accommodate either Ub or the Ub-like domain of HHR23 proteins (26).

The regions predicted to be responsible for the fibrillogenic behavior of Josephin, thus, include two exposed highly aggregation-prone patches that are functionally important. The residues in these regions are highly conserved from humans to plants, confirming their crucial functional role (Fig. 1*C*). This suggests that Josephin fibrillogenesis and function could represent competing processes promoted by the same regions.

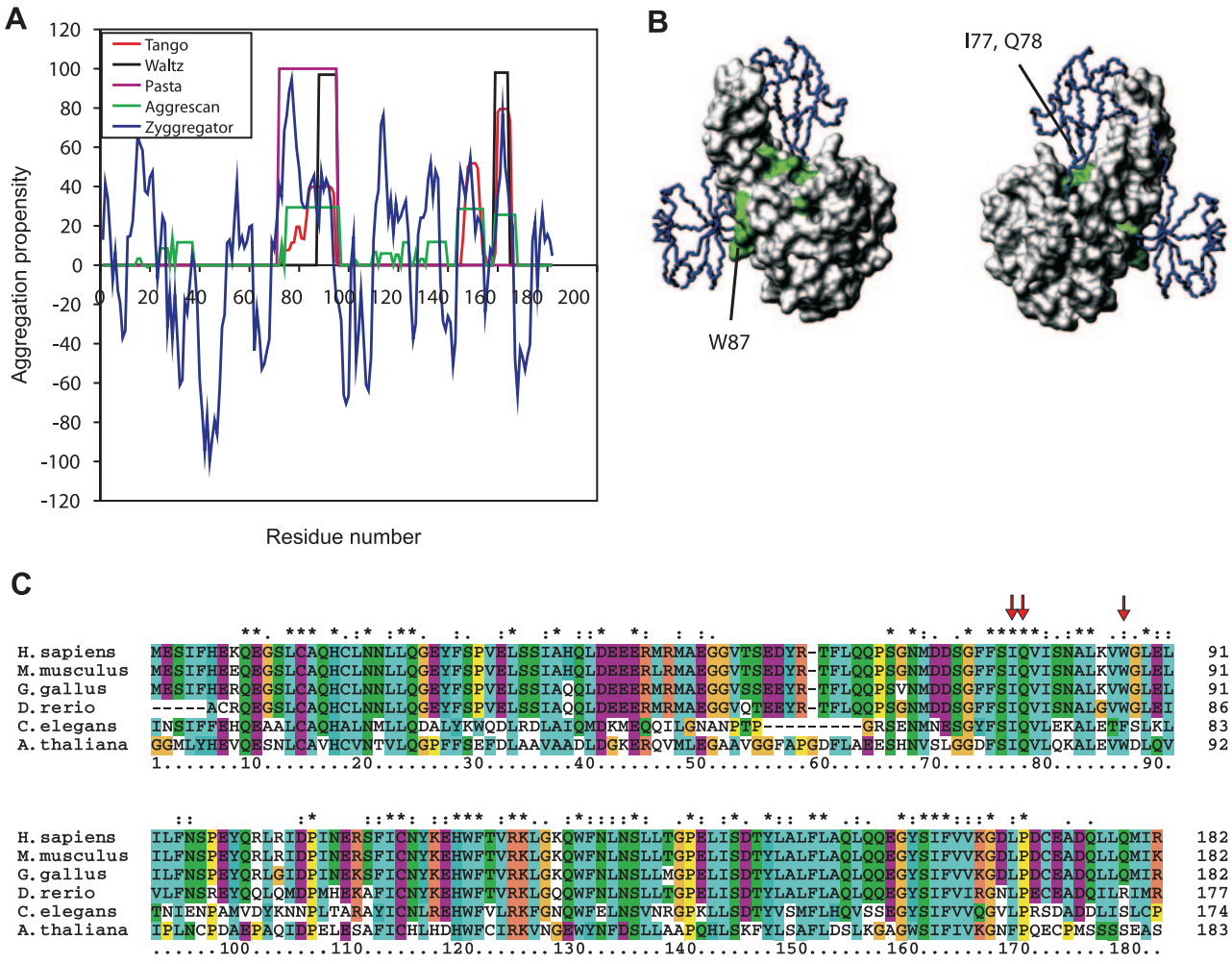

Figure 1. Prediction of the fibrillogenic regions of Josephin. *A*) Aggregation propensity profiles of Josephin as predicted by Aggrescan, Tango, Waltz, Pasta, and Zyggregator (14, 37–40). Since the relative scales are not comparable, data were arbitrarily normalized to 100. As compared to the other program, Zyggregator (39) gives a score, defined as in ref. 14, both for aggregation-prone (positive) and aggregation-resistant (negative) residues. *B*) Surface representation of Josephin (PDB entry: 2jri); the 2 exposed hydrophobic patches that form the interfaces with Ub are shown in green. Residues mutated within these regions are indicated. Ub molecules are shown in blue. *C*) Sequence alignment of Josephin from *Homo sapiens*, *Mus musculus*, *Gallus gallus*, *Danio rerio*, *Caenorhabditis elegans*, and *Arabidopsis thaliana*. Stars indicate sequence identity. Colons and dots indicate decreasing degrees of conservation. Numbering refers to the sequence of human Josephin. Red arrows indicate mutated residues I77, Q78, and W87.

Josephin forms fibrillar aggregates also under nondestabilizing conditions

The observation that the fibrillogenic sequences of Josephin cluster together to form surface-exposed hydrophobic patches could be directly relevant to misfolding if aggregation were observed also under nondestabilizing conditions. We investigated this hypothesis by studying Josephin aggregation by NMR spectroscopy. The process was monitored by exploiting the progressive decrease of peak intensities of backbone amide groups in 2-D ^{1}H-^{15}N HSQC spectra, due to the increase in the tumbling time of the forming aggregate, as described previously (29). Josephin is initially folded and monodispersed, as shown by the well-dispersed and sharp resonances. These experiments were performed in 20 mM sodium phosphate buffer, conditions chosen to have slower and thus more easily measurable kinetics. Although a higher ionic strength would mimic the cell environment more accurately, we consistently observed that higher salt concentrations (up to 150 mM) increase Josephin aggregation propensity (data not shown). Even at lower salt concentrations, Josephin still remains highly prone to aggregation under native conditions. The NMR data show that, at 250 μM and 38°C, the peak intensities measured in HSQC spectra for wild-type Josephin become undetectable after ~30 h. Under our conditions, the aggregation kinetics do not exhibit an appreciable lag phase and were fit by single exponential functions (**Fig. 2*A***).

These results are in excellent agreement with SEC data. Aggregation kinetics were monitored by incubating Josephin at 37°C and recording elution profiles at different time points. Our data show that, as the intensity of the

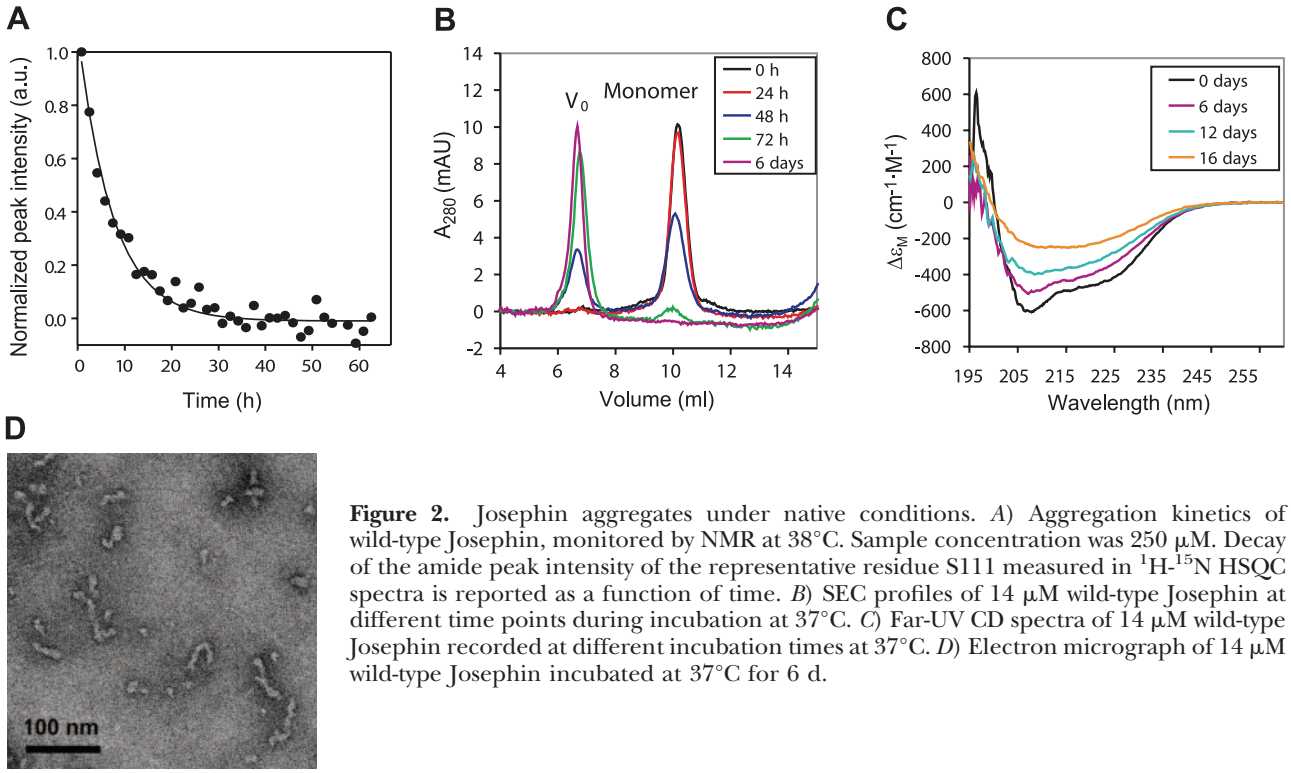

Figure 2. Josephin aggregates under native conditions. *A*) Aggregation kinetics of wild-type Josephin, monitored by NMR at 38°C. Sample concentration was 250 μM. Decay of the amide peak intensity of the representative residue S111 measured in ^{1}H-^{15}N HSQC spectra is reported as a function of time. *B*) SEC profiles of 14 μM wild-type Josephin at different time points during incubation at 37°C. *C*) Far-UV CD spectra of 14 μM wild-type Josephin recorded at different incubation times at 37°C. *D*) Electron micrograph of 14 μM wild-type Josephin incubated at 37°C for 6 d.

peak corresponding to the monomer (elution volume 10.1 ml) progressively decreases at longer incubation times, a peak corresponding to the void volume of the column (elution volume 6.7 ml) appears, corresponding to the aggregate (Fig. 2*B*). The molecular weight of the aggregated species is estimated to be ≥80 kDa from the separation range of the Superdex column, and thus the aggregates comprise ≥4 Josephin units.

We estimated variations of secondary structure at early stages of aggregation and misfolding. Similarly to full-length ataxin-3, Josephin is known to undergo a structural transition between a α-to-β-rich conformation at high temperatures (29). The same behavior was reproduced at 37°C using the same sample used for SEC analysis. These data, therefore, confirm that Josephin aggregation under native conditions is coupled to an increase of β-sheet structure (Fig. 2*C*).

To ascertain whether wild-type Josephin aggregation is concurrent with fibrillization as observed at high temperature (29), the samples were further investigated by EM. Fresh Josephin samples were analyzed prior to incubation as a control. An even spread of stain in the micrographs indicates the presence of protein, but no particles were observed (Supplemental Fig. S1). After 6 d of incubation at 37°C, EM analysis indicates the presence of a heterogeneous population of flexible fibrils ~7–12 nm wide and up to 170 nm long (Fig. 2*D*). Smaller globular oligomers with a diameter of ~7–12 nm are also detected. The fibers, curvilinear and sometimes branched, have a beaded morphology and resemble those previously obtained by slowly heating Josephin to high temperature (29). Interestingly, the fibrils and shorter oligomers obtained with Josephin also closely resemble those reported for full-length both nonexpanded and expanded ataxin-3, under nondenaturing conditions (33, 34, 41).

Taken together, these results confirm that, like ataxin-3, Josephin is also able to misfold and form fibrils under native conditions, provide further evidence for a leading role of this domain in the fibrillogenesis of the full-length protein and suggests the importance of the identified hydrophobic patches for aggregation.

Aggregation is initiated at functionally important regions

To test this hypothesis experimentally, we performed a systematic *in silico* mutagenesis screening using Zyggregator (14, 39). Among different possibilities, we selected 2 mutants, Jos_I77KQ78K and Jos_W87K, in which the native residues are replaced by lysine residues. In our screening, these substitutions are expected to have the strongest effects in reducing aggregation propensity by introducing the minimum number of perturbations in each of the Ub binding sites (**Fig. 3*A***). As previously demonstrated, each set of mutations abolishes one Ub binding site, so that Jos_I77KQ78K loses the ability of binding Ub to the enzyme active site, whereas Jos_W87K retains this binding site only (26). NMR and far-UV CD experiments confirmed that the mutations do not affect Josephin secondary and tertiary structure or significantly alter the fold stability (data not shown).

The aggregation kinetics of the mutants were measured by NMR spectroscopy and compared to those of wild-type Josephin recorded under the same conditions. The mutants are also initially monodispersed and aggregate under physiological conditions without an appreciable lag phase, but the kinetics of aggregation are remarkably different (Fig. 3*B*).

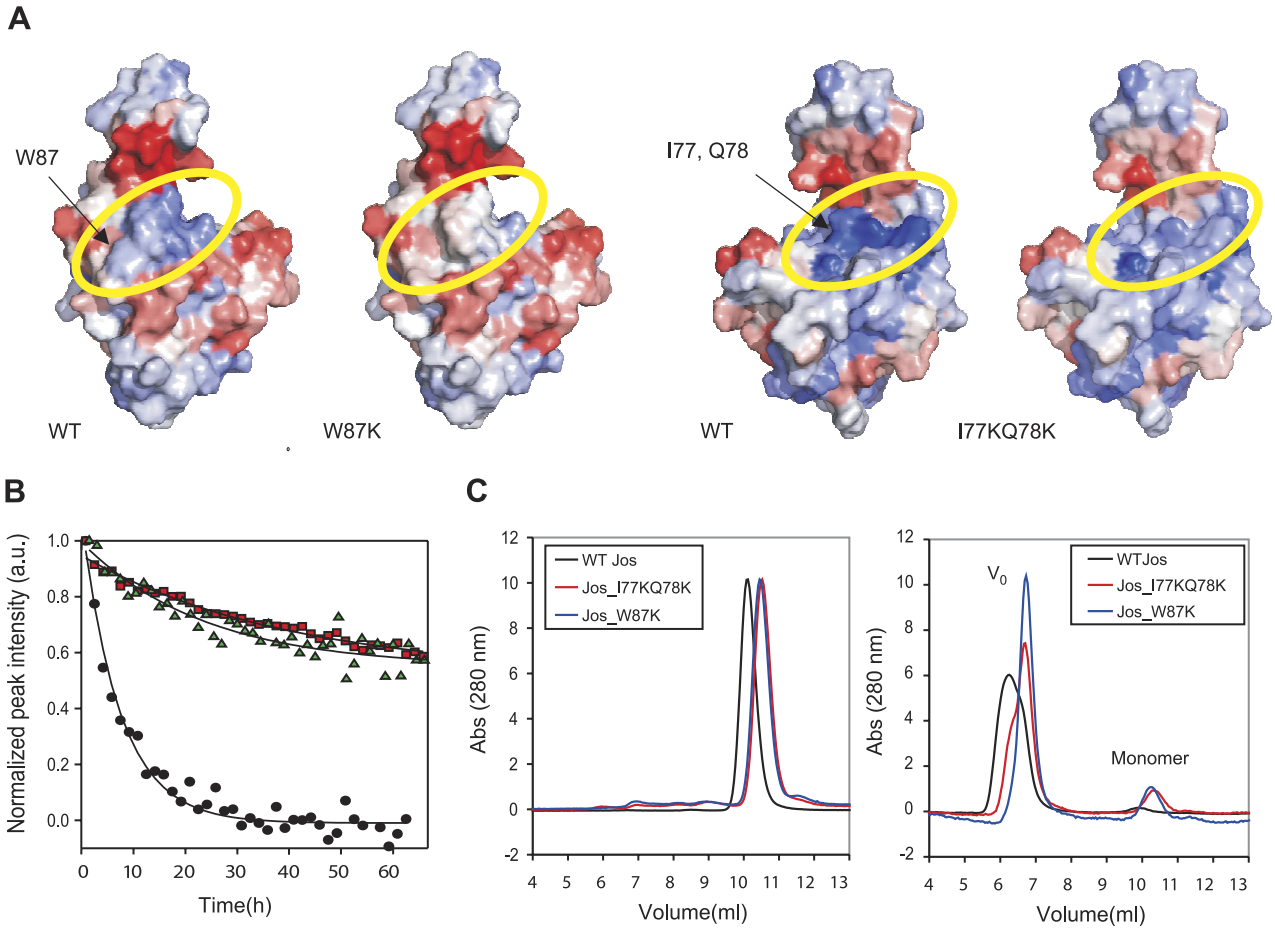

Figure 3. Design and characterization of Josephin mutants. *A*) Predicted changes in surface aggregation propensity of Josephin on mutation. Comparison between wild-type Josephin (PDB entry: 1yzb) and the mutants Jos_W87K and Jos_I77KQ78K. Aggregation-prone and aggregation-resistant regions as defined in ref. 14 are indicated in blue and red, respectively, with arbitrary gradations of color. Arrows indicate mutated positions. Yellow ellipses indicate regions in which aggregation propensity is most affected by the mutations. *B*) Comparison of the aggregation kinetics of the Josephin mutants with those of the wild-type protein, monitored by NMR at 38°C. Sample concentration was 250 μM. As in Fig. 2*A*, kinetics are shown for a representative residue (S111). Solid circles denote wild-type Josephin, red squares denote the Jos_I77KQ78K mutant, and green triangles denote the Jos_W87K mutant. *C*) Comparison of SEC profiles measured for 14 μM wild-type Josephin (black), Jos_I77KQ78K (red), and Jos_W87K (blue). Left panel: fresh proteins prior to incubation. Right panel: proteins incubated at 37°C for 72 h.

The half-life times measured for Jos_I77KQ78K and Jos_W87K are much longer (9246 and 6716 min, respectively) than that of the wild-type (510 min). SEC experiments carried out on the mutant Josephin domains confirm that the aggregation rates of wild-type Josephin are faster than those of Jos_I77KQ78K and Jos_W87K (Fig. 3*C*).

Taken together, these data indicate that a reduction in surface hydrophobicity decreases substantially the aggregation propensity of Josephin, though not abolishing it entirely. The hydrophobic regions that mediate partner recognition thus also promote Josephin aggregation.

Ubiquitin binding slows the Josephin aggregation kinetics

To test the effect of an interaction partner binding on self-aggregation kinetics, we repeated the NMR and SEC experiments described above in the presence of mono-Ub. In NMR HSQC spectra, in the presence of a molar excess of Ub (*i.e.*, a [Ub]:[Josephin] ratio of 3:1), a significant residual signal attributable to a fraction of soluble protein that has not yet aggregated is observed after 85 h at 38°C (corresponding to a half-life-time of 2174 min) (**Fig. 4*A***). Thus Ub binding induces a marked decrease in Josephin aggregation rates. Also in this case, the kinetics could be fit by single-exponential functions. A similar effect of Ub was observed for the Jos_I77KQ78K and Jos_W87K mutants, which bind one Ub molecule only.

SEC measurements were performed on wild-type and mutant Josephin >90% saturated by Ub. After 48 h at 37°C, virtually no aggregates are formed in the presence of Ub, whereas aggregation is substantial in the absence of Ub (the ratio aggregate:monomer estimated from peak areas is ~1.8:1) (Fig. 4*B*). These results confirm the role of Ub binding in slowing Josephin aggregation kinetics.

Ubiquitin binding affects Josephin aggregate morphology

As the interaction with Ub delays Josephin aggregation but does not suppress it altogether, we finally explored whether addition of Ub alters formation of the mature

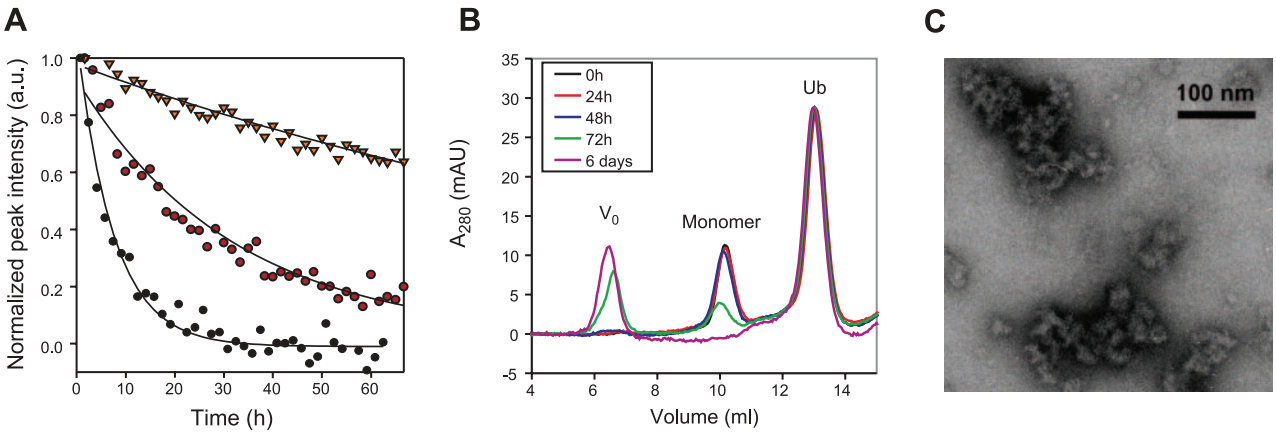

Figure 4. Effects of Ub on Josephin aggregation properties. *A*) Aggregation kinetics of Josephin in the presence of an excess of Ub ([Ub]:[Josephin] ratio of 3:1), monitored by NMR as in Figs. 2*A* and 3*B*. Wild-type and mutant Josephin concentration was 250 μM. Red circles denote wild-type Josephin; yellow triangles denote Jos_I77KQ78K mutant. Kinetics for wild-type Josephin in the absence of Ub (black circles) are reported for comparison. *B*) SEC profiles of 14 μM wild-type Josephin in the presence of 593 μM Ub at different time points during incubation at 37°C. *C*) Electron micrograph of 14 μM wild-type Josephin incubated at 37°C for 6 d in the presence of 593 μM Ub.

fibers and/or changes their morphology. Analysis of EM images of the Josephin samples incubated in the presence of saturating Ub concentrations still reveals the formation of aggregates, in agreement with NMR and SEC data. After 72 h of incubation at 37°C, a heterogeneous population of aggregates is observed. While sharing some of the features of wild-type Josephin fibrils, the aggregates seem overall less ordered and more amorphous, with dimensions varying between 12 and 100 nm (Supplemental Fig. S2). As aggregation proceeds, these differences in the morphology of Josephin aggregates formed in the absence and in the presence of Ub are consistently observed, as shown by EM micrographs obtained after 6 d at 37°C (Fig. 4*C*). In these samples, part of the aggregates appear to originate from assemblies of the same fibrillar species observed for wild-type Josephin, that are stuck together to form larger objects, with overall widths up to 200 nm. Ub only samples were incubated at 37°C and analyzed under the same conditions used for Josephin samples, as controls. EM micrographs acquired at different time points show the presence of highly concentrated protein, but the absence of aggregates after 6 d of incubation (Supplemental Fig. S3).

Thus, while not inhibiting completely the formation of the ordered fibrillar structures, Ub binding has an effect on Josephin aggregate morphology.

DISCUSSION

Functional interactions as a protection mechanism against aggregation

Substantial efforts have been made to characterize the mechanisms of protein aggregation and fiber formation because of their implication in neurodegeneration (1, 2, 12). Proteins known to cause misfolding diseases can be intrinsically disordered, locally unfolded or have a stable globular fold. In the latter case, proteins generally display low propensity to aggregate spontaneously, as the elements that commonly promote aggregation are not exposed to the solvent (10). Thus, solution conditions that reduce the thermodynamic stability (for instance by heat, pressure, or solvent perturbation) have been widely used to analyze the properties of globular proteins *in vitro*. While providing a large plethora of valuable knowledge of the mechanisms of protein misfolding, these studies have, however, told us relatively little about the very early events which trigger aggregation in the absence of destabilizing conditions. Yet, recent evidence has indicated at least a few examples of proteins that are able to aggregate also under native conditions (12).

One such case is Josephin, as shown here by a combination of NMR, SEC, and EM methods. Josephin, however, differs from previous examples, which seem to misfold through transient local perturbations, exposure of regions due to structural fluctuations, subunit dissociation, or *cis-trans* prolyl peptide isomerization (12). Josephin is a relatively rigid, globular α-rich domain with no disordered regions. SAXS and NMR data support the presence in solution of a well-defined open conformation in which a more flexible α-hairpin can rotate around a hinge but remains folded (25, 29, 42). Although not excluding that structural fluctuations or rearrangements could occur during aggregation and could even have a strong impact on fibrillization (we do observe a structural transition at later stages of the aggregation pathway), the early events of aggregation do not seem to require a structural opening: solvent-exposed aggregation-prone patches formed by residues contiguous in space but not necessarily in sequence are the initiation signal for self-assembly.

A key result of our study, which provides a mechanism for how aggregation under native conditions is triggered, is the observation that the multiple aggrega-

tion sites map onto surfaces involved in the normal function of the protein (26). This observation is supported by the dramatic effects that Ub binding has on Josephin aggregation kinetics *in vitro*: even though not preventing aggregation altogether, monoUb (which has relatively low affinity for Josephin; ref. 26) significantly reduces Josephin self-association rates. In the cell, where ataxin-3 is known to bind preferentially polyUb chains of ≥4 subunits, inhibition of aggregation may be amplified by the tighter affinity of the interaction. This suggests a direct link between structure, "normal" function, and aggregation.

Our results bear important consequences for our understanding of protein misfolding and developing ways to prevent it. First, considerable interest has been paid in recent years to predicting fibrillogenic regions. So far (and in this work), aggregation-prone regions have been predicted on the basis of sequence information. The "signal" that promotes aggregation of Josephin is, instead, on the protein surface and is not sequential, suggesting a completely different perspective to predict aggregation-prone regions. Second, our data may clarify how proteins have evolved to avoid misfolding. A widely accepted evolutionary hypothesis is that proteins have utilized negative design principles to generate sequences that can not only reach the final desired structure but can also avoid unproductive pathways (3, 10, 14). We show here how normal protein function can have an active role in protecting against aberrant aggregation and we suggest how this molecular mechanism occurs in ataxin-3. Analysis of the effects of the studied mutations on the aggregation process indicates that solvent-exposed aggregation-prone patches that promote Josephin aggregation have been carefully designed by evolutionary selection to enable strong intermolecular interactions while maintaining specificity and selectivity of target recognition, as also indicated by the extremely high conservation of the residues in these regions. Josephin aggregation and function are, therefore, encoded in the same sequence patterns and represent competing pathways, so that aggregation may be regarded as the "dark side" of function **(Fig. 5)**.

This competition could be exploited for the design of therapeutic strategies. Ligand binding has been shown to be effective in stabilizing the native protein conformation and decrease its propensity to aggregate. The use of small molecules, peptides, molecular chaperones, or intrabodies that could stop or reverse the aggregation process, or redirect it toward the formation of nontoxic species has been tested and suggested for polyglutamine diseases and other conformational disorders (43–45). Among these approaches, targeting interactions with physiological partners and modulating the signaling pathways regulating these interactions seem to represent an effective route to treatment (45, 46).

Finally, our study will hopefully trigger more interest in studying the nonpathologic functions of proteins involved in misfolding diseases as a general approach to these pathologies. Normal function has often been badly overlooked: the nonpathologic role of several proteins associated with misfolding pathologies (*e.g.*, Aβ, prions, α-synuclein) remains unknown. It is instead clear from this study that, far from having only an academic interest, the relationship between normal function and pathology is crucial for understanding

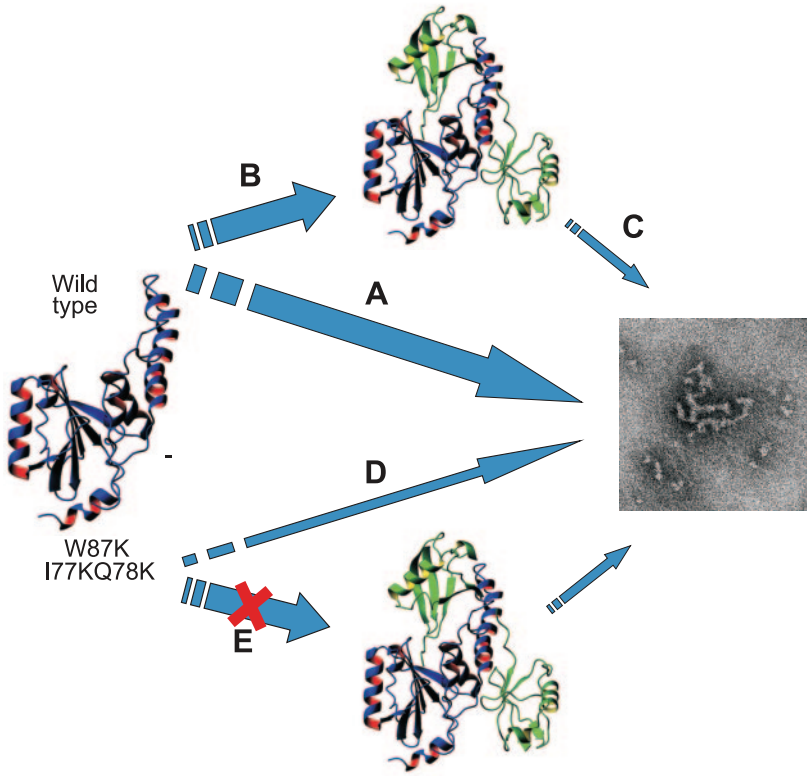

Figure 5. Model of the effects of the competition between Ub binding and aggregation. Josephin (PDB entry: 2jri) is shown in blue; Ub molecules are shown in green. Wild-type Josephin undergoes aggregation under physiological conditions (*A*); Ub binding (*B*) slows aggregation (*C*); targeted mutations (I77KQ78K and W87K) at the 2 Ub-binding interfaces prevent Ub binding (*E*) and also slow aggregation (*D*). Width of arrows indicates reaction rates.

aggregation. This is not the first example: we have recently shown that ataxin-1 (47), another member of the polyQ family, contains a phosphorylation site that acts as the switch that delimits pathology from normal function.

Implications for the role of polyQ expansions in SCA3

The finding that the Josephin domain of ataxin-3 contains highly aggregation-prone nonpolyQ elements opens new questions about the role of polyQ expansions in promoting ataxin-3 aggregation: Why does ataxin-3 cause SCA3 pathology only on polyQ expansion? What prevents it from misfolding when nonexpanded? We previously proposed an "induced misfit" model to explain the aggregation properties of nonexpanded ataxin-3 and the observed hierarchy of aggregation of the two fibrillogenic elements, the polyQ region, and Josephin (29). This model is fully supported by our current data. Although *in vitro* experiments cannot reproduce the vast complexity of the situation *in vivo*, where several pathways and multiple different interactions are present, our results suggest an explanation that may constitute a more general paradigm also for other proteins involved in misfolding diseases: the binding of Josephin to Ub or other cellular partners, or even of other regions of the protein itself, could make the aggregation-prone surfaces of Josephin unable to self-assemble by masking the fibrillogenic sites, thereby stabilizing the soluble form of nonexpanded ataxin-3 (Fig. 5). This hypothesis is independently supported by the reported interaction with another partner of ataxin-3, the valosin-containing protein VCP. Although this interaction was mapped to the C-terminal region and not to Josephin, it also modulates aggregation of ataxin-3 and reduces its toxicity in an *in vivo* model (46). When a polyQ expansion occurs, an additional highly fibrillogenic element is introduced in the bulky and aggregation-prone C terminus, which could destabilize or prevent formation of intermolecular interactions, for instance by mere steric hindrance or by altering folding efficiency (18). As a consequence, aggregation would be triggered, eventually leading to formation of neuronal inclusions.

Sequence analysis and the available structural data show that ataxin-3 is not the only protein causing polyQ diseases that contains folded functional domains flanked by regions that are intrinsically unstructured and include polyQ tracts. Ataxin-1, ataxin-2, and ataxin-7, for example, share similar characteristics (48–50). Ataxin-1, in particular, has striking similarities with ataxin-3. The AXH domain, near the C terminus of the molecule, seems to be the only structured portion of ataxin-1 and is surrounded by extensive regions that lack secondary structure and include the polyQ tract (49). Similarly to Josephin, the AXH domain is prone to aggregation and able to form fibers *in vitro* (51). Although more detailed studies are needed to assess a more general validity of the induced misfit model, the results that we have presented here provide a compelling example of the way in which folding, interdomain interactions, molecular recognition, and aberrant aggregation can be encoded in the same amino acid sequences, and how their balance can determine protein function and dysfunction. FJ

The authors are very thankful to Xavier Daura and Oscar Conchillo Sole for help with the Aggrescan Web server and to the Medical Research Council (MRC) Biomedical NMR Centre for technical support. The authors acknowledge support from the MRC (grant U117584256).

REFERENCES

1. Chiti, F., and Dobson, C. M. (2006) Protein misfolding, functional amyloid, and human disease. *Annu. Rev. Biochem.* **75,** 333–366
2. Powers, E. T., Morimoto, R. I., Dillin, A., Kelly, J. W., and Balch, W. E. (2009) Biological and chemical approaches to diseases of proteostasis deficiency. *Annu. Rev. Biochem.* **78,** 959–991
3. Dobson, C. M. (1999) Protein misfolding, evolution and disease. *Trends Biochem. Sci.* **24,** 329–332
4. Fowler, D. M., Koulov, A. V., Balch, W. E., and Kelly, J. W. (2007) Functional amyloid—from bacteria to humans. *Trends Biochem. Sci.* **32,** 217–224
5. Zhang, S. (2003) Fabrication of novel biomaterials through molecular self-assembly. *Nat. Biotechnol.* **21,** 1171–1178
6. Chiti, F., Stefani, M., Taddei, N., Ramponi, G., and Dobson, C. M. (2003) Rationalization of the effects of mutations on peptide and protein aggregation rates. *Nature* **424,** 805–808
7. Marshall, K. E., and Serpell, L. C. (2009) Structural integrity of beta-sheet assembly. *Biochem. Soc. Trans.* **37,** 671–676
8. Nelson, R., and Eisenberg, D. (2006) Structural models of amyloid-like fibrils. *Adv. Protein Chem.* **73,** 235–282
9. Kelly, J. W. (1998) The alternative conformations of amyloidogenic proteins and their multi-step assembly pathways. *Curr. Opin. Struct. Biol.* **8,** 101–106
10. Tartaglia, G. G., Pawar, A. P., Campioni, S., Dobson, C. M., Chiti, F., and Vendruscolo, M. (2008) Prediction of aggregation-prone regions in structured proteins. *J. Mol. Biol.* **380,** 425–436
11. Uversky, V. N., and Fink, A. L. (2004) Conformational constraints for amyloid fibrillation: the importance of being unfolded. *Biochim. Biophys. Acta* **1698,** 131–153
12. Chiti, F., and Dobson, C. M. (2009) Amyloid formation by globular proteins under native conditions. *Nat. Chem. Biol.* **5,** 15–22
13. Balch, W. E., Morimoto, R. I., Dillin, A., and Kelly, J. W. (2008) Adapting proteostasis for disease intervention. *Science* **319,** 916–919
14. Pechmann, S., Levy, E. D., Tartaglia, G. G., and Vendruscolo, M. (2009) Physicochemical principles that regulate the competition between functional and dysfunctional association of proteins. *Proc. Natl. Acad. Sci. U. S. A.* **106,** 10159–10164
15. Cummings, C. J., and Zoghbi, H. Y. (2000) Trinucleotide repeats: mechanisms and pathophysiology. *Annu. Rev. Genomics Hum. Genet.* **1,** 281–328
16. 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
17. Riess, O., Rub, U., Pastore, A., Bauer, P., and Schols, L. (2008) SCA3: neurological features, pathogenesis and animal models. *Cerebellum* **7,** 125–137
18. Saunders, H. M., and Bottomley, S. P. (2009) Multi-domain misfolding: understanding the aggregation pathway of polyglutamine proteins. *Protein Eng. Des. Sel.* **22,** 447–451
19. 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

20. Evert, B. O., Araujo, J., Vieira-Saecker, A. M., de Vos, R. A., Harendza, S., Klockgether, T., and Wullner, U. (2006) Ataxin-3 represses transcription via chromatin binding, interaction with histone deacetylase 3, and histone deacetylation. *J. Neurosci.* **26,** 11474–11486
21. Li, F., Macfarlan, T., Pittman, R. N., and Chakravarti, D. (2002) Ataxin-3 is a histone-binding protein with two independent transcriptional corepressor activities. *J. Biol. Chem.* **277,** 45004–45012
22. 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
23. Chai, Y., Shoesmith 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
24. Doss-Pepe, E. W., Stenroos, E. S., Johnson, W. G., and Madura, K. (2003) Ataxin-3 interactions with rad23 and valosin-containing protein and its associations with ubiquitin chains and the proteasome are consistent with a role in ubiquitin-mediated proteolysis. *Mol. Cell. Biol.* **23,** 6469–6483
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. U. S. A.* **102,** 10493–10498
26. Nicastro, G., Masino, L., Esposito, V., Menon, R. P., De Simone, A., Fraternali, F., and Pastore, A. (2009) The josephin domain of ataxin-3 contains two distinct ubiquitin binding sites. *Biopolymers* **91,** 1203–1214
27. 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
28. Marchal, S., Shehi, E., Harricane, M. C., Fusi, P., Heitz, F., Tortora, P., and Lange, R. (2003) Structural instability and fibrillar aggregation of non-expanded human ataxin-3 revealed under high pressure and temperature. *J. Biol. Chem.* **278,** 31554–31563
29. 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
30. Shehi, E., Fusi, P., Secundo, F., Pozzuolo, S., Bairati, A., and Tortora, P. (2003) Temperature-dependent, irreversible formation of amyloid fibrils by a soluble human ataxin-3 carrying a moderately expanded polyglutamine stretch (Q36). *Biochemistry* **42,** 14626–14632
31. Chow, M. K., Ellisdon, A. M., Cabrita, L. D., and Bottomley, S. P. (2004) Polyglutamine expansion in ataxin-3 does not affect protein stability: Implications for misfolding and disease. *J. Biol. Chem.* **1,** 1
32. Ellisdon, A. M., Pearce, M. C., and Bottomley, S. P. (2007) Mechanisms of ataxin-3 misfolding and fibril formation: kinetic analysis of a disease-associated polyglutamine protein. *J. Mol. Biol.* **368,** 595–605
33. Ellisdon, A. M., Thomas, B., and Bottomley, S. P. (2006) The 2-stage pathway of ataxin-3 fibrillogenesis involves a polyglutamine-independent step. *J. Biol. Chem.* **281,** 16888–16896
34. Gales, L., Cortes, L., Almeida, C., Melo, C. V., do Carmo Costa, 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.* **26,** 26
35. Nicastro, G., Masino, L., Frenkiel, T. A., Kelly, G., McCormick, J., Menon, R. P., and Pastore, A. (2004) Letter to the editor: assignment of the (1)H, (13)C, and (15)N resonances of the Josephin domain of human ataxin-3. *J. Biomol. NMR* **30,** 457–458
36. Reumers, J., Conde, L., Medina, I., Maurer-Stroh, S., Van Durme, J., Dopazo, J., Rousseau, F., and Schymkowitz, J. (2008) Joint annotation of coding and non-coding single nucleotide polymorphisms and mutations in the SNPeffect and PupaSuite databases. *Nucleic Acids Res.* **36,** D825–D829
37. Conchillo-Sole, O., de Groot, N. S., Aviles, F. X., Vendrell, J., Daura, X., and Ventura, S. (2007) AGGRESCAN: a server for the prediction and evaluation of "hot spots" of aggregation in polypeptides. *BMC Bioinformatics* **8,** 65
38. Fernandez-Escamilla, A. M., Rousseau, F., Schymkowitz, J., and Serrano, L. (2004) Prediction of sequence-dependent and mutational effects on the aggregation of peptides and proteins. *Nat. Biotechnol.* **22,** 1302–1306
39. Tartaglia, G. G., and Vendruscolo, M. (2008) The Zyggregator method for predicting protein aggregation propensities. *Chem. Soc. Rev.* **37,** 1395–1401
40. Trovato, A., Seno, F., and Tosatto, S. C. (2007) The PASTA server for protein aggregation prediction. *Protein Eng. Des. Sel.* **20,** 521–523
41. Bevivino, A. E., and Loll, P. J. (2001) 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
42. Nicastro, G., Habeck, M., Masino, L., Svergun, D. I., and Pastore, A. (2006) Structure validation of the Josephin domain of ataxin-3: conclusive evidence for an open conformation. *J. Biomol. NMR* **36,** 267–277
43. Cohen, F. E., and Kelly, J. W. (2003) Therapeutic approaches to protein-misfolding diseases. *Nature* **426,** 905–909
44. Ehrnhoefer, D. E., Bieschke, J., Boeddrich, A., Herbst, M., Masino, L., Lurz, R., Engemann, S., Pastore, A., and Wanker, E. E. (2008) EGCG redirects amyloidogenic polypeptides into unstructured, off-pathway oligomers. *Nat. Struct. Mol. Biol.* **15,** 558–566
45. Shao, J., and Diamond, M. I. (2007) Polyglutamine diseases: emerging concepts in pathogenesis and therapy. *Hum. Mol. Genet.* **16,** R115–R123
46. Boeddrich, A., Gaumer, S., Haacke, A., Tzvetkov, N., Albrecht, M., Evert, B. O., Muller, E. C., Lurz, R., Breuer, P., Schugardt, N., Plassmann, S., Xu, K., Warrick, J. M., Suopanki, J., Wullner, U., Frank, R., Hartl, U. F., Bonini, N. M., and Wanker, E. E. (2006) An arginine/lysine-rich motif is crucial for VCP/p97-mediated modulation of ataxin-3 fibrillogenesis. *EMBO J.* **25,** 1547–1558
47. De Chiara, C., Menon, R.P., Strom, M., Gibson, T.J., and Pastore, A. (2009) Phosphorylation of S776 and 14–3-3 binding modulate ataxin 1 interaction with splicing factors. *PLoS ONE* **4**, e8372
48. Albrecht, M., Golatta, M., Wullner, U., and Lengauer, T. (2004) Structural and functional analysis of ataxin-2 and ataxin-3. *Eur. J. Biochem.* **271,** 3155–3170
49. De Chiara, C., Giannini, C., Adinolfi, S., de Boer, J., Guida, S., Ramos, A., Jodice, C., Kioussis, D., and Pastore, A. (2003) The AXH module: an independently folded domain common to ataxin-1 and HBP1. *FEBS Lett.* **551,** 107–112
50. Uversky, V. N. (2009) Intrinsic disorder in proteins associated with neurodegenerative diseases. *Front. Biosci.* **14,** 5188–5238
51. De Chiara, C., Menon, R. P., Dal Piaz, F., Calder, L., 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

Received for publication March 31, 2010.
Accepted for publication August 19, 2010.