The Protein Journal, Vol. 24, No. 4, May 2005 (© 2005)
DOI: 10.1007/s10930-005-6713-4

An Unstructured Region is Required by GAV Homologue for the Fibrillization of Host Proteins

Li-Na Ji,[1,2] Hai-Ning Du,[1] Feng Zhang,[2] Hong-Tao Li,[1] Xiao-Ying Luo,[1] Jun Hu,[2,3] and Hong-Yu Hu[1,4]

Accumulating evidence shows that some amyloidogenic proteins contain core sequences, which are critical for their fibrillization. Core sequences of α-synuclein, β-amyloid peptide and prion protein usually reside in their unfolded regions and share a conserved consensus (VGGAVVAGV) designated as GAV homologue. Here we investigate the role of unfolded regions in fibrillization after GAV homologue is attached to the C-terminus or inserted into the loop regions of different host proteins, namely α -Syn_{1-65}, γ-synuclein, *E. coli* thioredoxin and immunoglobulin G binding B1 domain of *streptococcal* protein G. The results imply that an unstructured region is required by GAV homologue for the fibrillization of host proteins. A number of amyloidogenic proteins with core sequences located in unstructured regions are summarized and discussed in details. The finding may provide further insight into the elucidating of the molecular mechanism underlying the fibrillization of α-Syn, Aβ and PrP as well as other amyloidogenic proteins.

KEY WORDS: α-synuclein; amyloidogenic protein; fibrillization; GAV homologue; unstructured regions.

1. INTRODUCTION

Deposition of amyloidogenic proteins in the patient brains is closely associated with a number of neurodegenerative diseases (Goedert, 2001; Rochet and Lansbury, 2000; Soto, 2003). The fibrillizaton of a-synuclein (α-Syn)[5] in the form of intracellular proteinaceous inclusions has been implicated as a causative factor in Parkinson's disease (PD) (Spillantini *et al.*, 1997). Similarly, β-amyloid peptide (Aβ) is the major component in amyloid plaques of Alzheimer's disease (AD) patients (Harper and Lansbury, 1997), and prion protein (PrP) deposits in brain tissue of prion diseases (PrD) patients (Prusiner, 1996).

Our previous work has shown that GAV motif (residues 66–74) is the core sequence of α-Syn, which is required for its fibrillization and cytotoxicity. Introduction of the charged residues into the GAV motif has a great negative impact on the rate of α-Syn fibrillization, and deletion of the GAV motif

[1] Key Laboratory of Proteomics, Institute of Biochemistry and Cell Biology, Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences, Shanghai, 200031, P.R. China.
[2] Shanghai Institute of Applied Physics, Chinese Academy of Sciences, Shanghai, 201800, P.R. China.
[3] Bio-X Research Center, Shanghai Jiaotong University, Shanghai, 200030, P.R. China.
[4] To whom correspondence should be addressed. E-mail: hyhu@sibs.ac.cn

[5] Abbreviations: Aβ, β-amyloid peptide; AD, Alzheimer's disease; AFM, atomic force microscopy; Ala, alanine; CD, circular dichroism; FPLC, fast protein liquid chromatography; GdmCl, guanidium chloride; Gly, glycine; HD, Huntington's disease; IAPP, islet amyloid polypeptide; NAC, non-Aβ component; OPMD, oculopharyngeal muscular dystrophy; PABP2, poly(A) binding protein 2; PBS, phosphate-buffered saline; PD, Parkinson's disease; PGBD, immunoglobulin G binding B1 protein of *streptococcal* protein G; PrD, prion diseases; PrP, prion protein; SBMA, spinal and bulbar muscular atrophy; SCA3/MJD, spino-cerebellar ataxia/Machado-Joseph disease; SDS-PAGE, sodium dodecyl sulfate polyacrylamide gel electrophoresis; α-Syn, human α-synuclein; γ-Syn, human γ-synuclein; ThT, thioflavin T; TRX, thioredoxin; UV, ultra-violet; Val, Valine; WT, wild type.

completely abolishes α-Syn fibrillization and cytotoxicity (Du *et al*., 2003). An earlier report by Giasson *et al*. suggests that the stretch of residues 71–82 is associated with aggregation of α-Syn, though it has never shown any results about the cytotoxicity of protein aggregates (Giasson *et al*., 2001). The hydrophobic stretch shares an overlapping sequence of ^{71}VTGV^{74} with our identified GAV motif. Our further studies showed that ^{71}VTGV^{74} might play an important role in α-Syn fibrillization, but the entire GAV motif of nine residues is essential for the formation of mature regular fibrils and the cytotoxicity of α-Syn (Du *et al*., 2003). The amyloidogenic carboxyl terminus (residues 36–44) of Aβ is critical for its fibrillization. Fluorescence quenching analysis reveals that a conformation change in the C-terminal region of Aβ accompanies its transition from dimer to fibril (Garzon-Rodriguez *et al*., 2000). The C-terminally truncated variants of Aβ form fibrils slowly *in vitro*,while the longer ones aggregate much more rapidly and they are the primary pathogenic proteins in AD (Iwatsubo *et al*., 1994; Jarrett *et al*., 1993). In addition, a peptide comprising the C-terminal nine residues of Aβ forms fibrils *in vitro* (Halverson *et al*., 1990). A similar amyloidogenic sequence is also present in the central region of the PrP. PrP (106–126) is well-known for its amyloidogenic and neurotoxic property (Forloni *et al*., 1993; Lin *et al*., 1997; Rymer and Good, 2000) and now the shorter PrP (113–127) is found to has a marked tendency to form amyloid fibrils and its N-terminal and C-terminal sequences are coordinately involved in the self-assembly (Satheeshkumar and Jayakumar, 2003).

The fibril formation of amyloidogenic proteins is usually nucleation-dependent and addition of preformed fibrils immediately initiates the aggregation by eliminating the long lag phase (Come *et al*., 1993; Jarrett and Lansbury, 1993; Wood *et al*.; 1999), Generally, the seeding event is sequence-specific, namely homogeneous seeding. However, non-Aβ component of AD amyloid (NAC), namely the residues 61–95 of α-Syn seeds Aβ amyloid formation, while Aβ and PrP (106–126) seed NAC amyloidogenesis (Han *et al*., 1995). The heterogeneous seeding among the peptides may be due to the local sequence homology in their core sequences. By comparing the core sequences of α-Syn (residues 66–74, VGGAVVTGV), Aβ (residues 36–44, VGGVVIATV) and PrP (residues 113–127, AGAAAAGAVVGGLGG), a conserved consensus (VGGAVVAGV) was found and designated as GAV homologue (Han *et al*., 1995; Hu, 2001). Our further studies show that the synthetic GAV homologue peptide remarkably self-assembly into fibrils *in vitro* (data not shown).

More importantly, the GAV homologous sequences are always located in the unfolded regions of the amyloidogenic proteins. α-Syn (Weinreb *et al*., 1996) and Aβ (Terzi *et al*., 1995) are entirely unfolded in aqueous solutions, while the N-terminus (residues 1–127) of PrP (Zahn *et al*., 2000) displays disordered structures. Core sequences of Tau protein (von Bergen *et al*., 2000), Ataxin-3 (Masino *et al*., 2003), Huntingtin (Li and Li, 1998; Scherzinger *et al*., 1999), yeast prions Sup35p (King *et al*., 1997) and Ure2p (Baxa *et al*., 2003; Taylor *et al*., 1999) also reside in their unstructured regions. These new findings motivate us to further investigate the natively unfolded regions of amyloidogenic proteins that may play an important role in their fibrillization.

Four different proteins are chosen as hosts for GAV homologue to explore the role of natively unfolded regions in fibrillization. α-Syn$_{1\text{-}65}$, the N-terminal 65-residue fragment of α-Syn, is natively unfolded and devoid of the GAV motif. γ-Synuclein (γ-Syn) also displays random coil in solution and significantly differs from α-Syn in their aggregation propensity (Biere *et al*., 2000; Ji *et al*., 1997; Uversky *et al*., 2002). Compactly folded *E. coli* thioredoxin (TRX) (Dyson *et al*., 1989; Holmgren *et al*., 1975) and immunoglobulin G binding Bl domain of *streptococcal* protein G (PGBD) (Alexander *et al*., 1992; Gronenborn *et al*., 1991) are small proteins (109 and 56 residues, respectively) with typical α/β topologies. In the study GAV homologous sequences were attached to the C-terminus or inserted into the loop regions of the host proteins, then the resultant mutants were characterized for their amyloidogenic properties and compared with the wide type (WT) ones. The results show that the unstructured region where the GAV homologue located is required for the fibrillization of host proteins, and may provide insight into the molecular mechanism of protein fibrillization.

2. MATERIALS AND METHODS

2.1. Mutation, Bacterial Expression, and Purification of Proteins

α-Syn$_{1\text{-}65}$ (C-terminally deleted residues 66–140 of α-Syn), α-Syn$_{1\text{-}74}$ (C-terminally attached residues 66–74 of α-Syn to α-Syn$_{1\text{-}65}$), Aβ-CT (C-terminally attached residues 36–44 of Aβ to α-Syn$_{1\text{-}65}$), PrP-CT

(C-terminally attached, residues 117–125 of PrP to α-Syn$_{1-65}$) and GAV-CT (C-terminally attached residues GAV homologue to α-Syn$_{1-65}$) were generated by site-directed mutagenesis using polymerase chain reaction. The γ-GAV mutant was obtained by mutagenesis using the γ-Syn gene as a template. For the sake of spectroscopic detection for proteins, a Phe4Trp mutation was introduced into all the above mutants. TRX-CT or PGBD-CT was constructed by attaching the GAV homologue to the C-terminus of the hosts. TRX-INS and PGBD-INS were constructed by overlap extension and the mutation introduced GAV homologue into the loop regions of the hosts. All the constructs encoding the target proteins were confirmed by DNA sequencing. The recombinant genes were subcloned into the pET3a vector, and the respective proteins were expressed in *Escherichia coli* BL21 (DE3).

Recombinant human α-Syn was prepared as described elsewhere (Du *et al.*, 2003; Hu *et al.*, 2001), α-Syn$_{1-65}$, α-Syn$_{1-74}$, Aβ-CT, PrP-CT and GAV-CT were purified through a CM-Sepharose Fast Flow cation-exchange column. γ-Syn, γ-GAV, WT TRX, TRX-CT, TRX-INS, WT PGBD, PGBD-CT and PGBD-INS proteins were purified through a Q-Sepharose Fast Flow anion-exchange column. Finally, the fractions containing target proteins were further purified with a FPLC Superose-12 column (Pharmacia) with PBS buffer (100 mM phosphate and 100 mM NaCl, pH 7.0). The homogeneity of the proteins was identified by SDS-PAGE and the concentrations of the proteins were determined by measuring the absorbance at 280 nm using extinction coefficients calculated according to the amino acid sequences.

2.2. Time Course of the Aggregation Process

All the protein samples were concentrated by using Centricon YM-3 spin filters (Millipore) in 0.1 M PBS buffer and then sterilely filtered through 0.22 μm filters to remove any granular matter. All the protein samples (200 μM) were incubated separately in sterile eppendorf tubes with continuous shaking at 37°C. With the incubation progressing, aliquots of each protein were taken at various time points and stored at −20°C. The time course of the aggregation process was monitored by a ThT (Aldrich) fluorescence assay. The enhancement of ThT fluorescence was used to semiquantitatively estimate the relative rate of filament formation. Fluorescence measurements were performed on a Hitachi F-4010 fluorophotometer. Each 20 μl of the incubated samples was added to 980 μl of 5 μM ThT in 50 mM glycine-NaOH buffer (pH 9.0). The emission intensities at 482 nm were recorded immediately after addition of the aliquot to the ThT solution with excitation at 446 nm.

2.3. Circular Dichroism Measurement

Far-UV CD spectra were measured by a JASCO-715 spectropolarimeter as described previously (Hu *et al.*, 2001). The incubated protein samples were diluted to 0.2 mg/ml in doubly distilled water and transferred to a 0.1-cm quartz cuvette. The fresh sample was used as a comparison. For each spectrum, an average of two scans was obtained, and the data were presented as relative ellipticity (milli deg) or mean residue molar ellipticity (deg cm^2/dmol). The GdmCl-induced unfolding was performed by monitoring the fluorescence change, and the thermodynamic parameters were estimated from the relative fractions of folded and unfolded states according to the two-state assumption (Pace, 1989).

2.4. Atomic Force Microscopy Imaging

Each sample was prepared by depositing 5 μl of the protein solution on freshly cleaved mica (Alfa Aesar). After adsorption for 5 min, the mica surface was gently washed with deionized water to remove redundant buffer and the protein that was not firmly attached to the surface. Excess water was removed with condensed air. Images were obtained at a commercial AFM facility (nanoscope III, Digital Instruments) equipped with a 130 μm × 130 μm scanner (J-scanner) by tapping mode imaging. The cantilevers with a nominal force constant of 20–100 N/m were used for the experiments. Every protein was imaged for at least three times independently and each sample was observed in more than five regions to avoid experimental error.

3. RESULTS

3.1. GAV Homologous Sequences Originating from Different Amyloidgenic Proteins Can Trigger Fibrillization of Natively Unfolded α-Syn$_{1-65}$

To explore the role of natively unfolded regions in driving filament formation, the GAV-motif

deficient α-Syn$_{1-65}$ was chose as host protein for GAV homologous sequences. α-Syn$_{1-65}$, α-Syn$_{1-74}$, Aβ-CT, PrP-CT and GAV-CT were investigated for their aggregation propensities, and the 66–74 residues of the latter four were shown in Fig. 1a. Upon incubation as monitored by ThT fluorescence assay, α-Syn$_{1-65}$ does not form aggregates even after incubation for several days at 37°C with continuous shaking (Fig. 1b). Consequently, AFM fails to reveal the presence of any filaments in the incubated α-Syn$_{1-65}$ sample (Fig. 1c). When the GAV homologous sequences are attached to the C-terminus of nonamyloidogenic α-Syn$_{1-65}$, the resultant proteins α-Syn$_{1-74}$, Aβ-CT, PrP-CT and GAV-CT become amyloidogenic. ThT fluorescence and AFM imaging suggest that all the hybrid proteins readily aggregate into fibrils (Fig. 1b, c).

The results demonstrate that GAV homologue acts as a core sequence for some amyloidgenic proteins, moreover, it can induce the natively unfolded and non-amyloidogenic α-Syn$_{1-65}$ to form fibrils.

3.2. GAV Homologue Can Initiate Fibrillization of Natively Unfolded γ-Syn

γ-Syn is the homologous protein of α-Syn and shares 60% similarity with α-Syn (Ji *et al.*, 1997). Although γ-Syn is also abundant in brain, it has not been linked to PD genetically. When γ-Syn and α-Syn were incubated at the same concentration under the same conditions *in vitro*, it does not form fibrils as α-Syn (Biere *et al.*, 2000; Du *et al.*, 2003). Furthermore, the fibrillization of α-Syn is inhibited in the presence of γ-Syn (Uversky *et al.*, 2002), and the preformed α-Syn fibrils cannot cross-seed γ-Syn (Biere *et al.*, 2000). Comparing the primary sequences of α-Syn and γ-Syn indicates that there is no GAV motif in the central region of γ-Syn. Thus, we constructed γ-Syn mutant (γ-GAV) by replacing the residues 66–74 (VSEAVVSSV) with GAV homologue (VGGAVVAGV). As previously reported (Biere *et al.*, 2000; Du *et al.*, 2003), aggregation of γ-Syn is rather slow as determined by ThT fluorescence and no fibrils can be detected by AFM imaging (Fig. 2a,b). However, the GAV-containing mutant can rapidly aggregate as WT α-Syn with a very short lag period during incubation (Fig. 2a). AFM imaging reveals that γ-GAV generates short fibrils after 4-day incubation and longer ones after 6-day incubation (Fig. 2b). CD studies show that WT γ-Syn retains its unfolded structure during incubation, while γ-GAV undergoes a structural transformation from random coil into β-sheet (Fig. 2c). The result further implies that the natively unfolded structure may play an important role in assisting γ-GAV with its fibrillization.

3.3. GAV Homologue Cannot Induce Fibrillization of the Compactly Folded Hosts

To further examine whether natively unfolded region is required by GAV homologue for the initiation of fibrillization, we adopted TRX and PGBD as model hosts. Firstly, two chimera proteins based on TRX were prepared, namely TRX-CT (GAV homologue attached to the C-terminus of TRX) and TRX-INS (GAV homologue inserted into the loop region between the α_3-helix and β_4-strand). Both chimeras retain the similar CD spectra as the WT TRX, suggesting their secondary structures remain unchanged (Fig. 3a). ThT fluorescence detection shows that WT protein and the chimera ones cannot bind ThT and enhance the fluorescence intensity (Fig. 3b). All the proteins can only form small oligomers upon incubation as revealed by AFM imaging (Fig. 3c). This is an example that the GAV sequence is incapable of inducing a well-folded protein to form fibrils.

As for the PGBD host, we also constructed two chimeras, PGBD-CT (GAV homologue attached to the C-terminus of PGBD) and PGBD-INS (GAV homologue inserted into the loop region between the α-helix and β_3-strand). The chimera proteins have secondary structures similar to WT PGBD as indicated by CD spectra (Fig. 4a). Upon incubation, PGBD-INS shows significant ThT fluorescence enhancement while PGBD-CT does not (Fig. 4b). AFM imaging reveals that both the chimeras fail to form regular fibrils, though PGBD-INS forms abundant amorphous aggregates while PGBD-CT forms only oligomers after incubation for 6 days (Fig. 4c). To understand why the insertion chimera binds ThT but only forms amorphous aggregates, we performed GdmCl-induced unfolding experiments of the three proteins for estimating the thermodynamic parameters (Pace, 1989). The Gibbs free energy change (ΔG^{H_2O}) for PGBD-CT is 3.63 kcal/mol and the midpoint value (C_m) is 1.88 M, similar to those of WT PGBD (3.82 cal/mol,1.89 M). However, PGBD-INS has been destabilized by insertion of the GAV homologue with a ΔG^{H_2O} only half of WT PGBD (1.71 kcal/mol) and a very small C_m. value (0.38 M). This indicates that

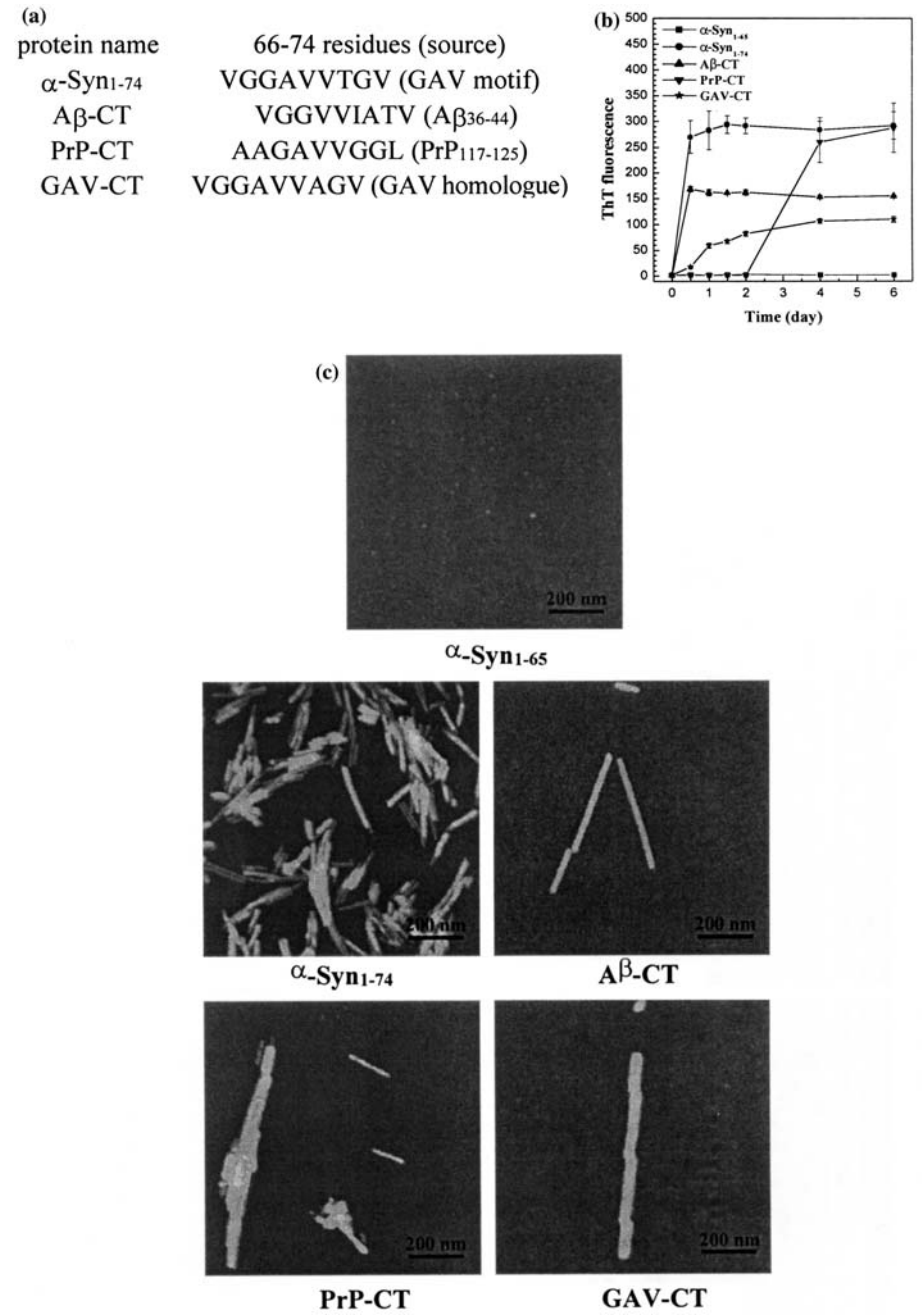

Fig. 1. Amylodgenic propensities of α-Syn$_{1-65}$, α-Syn$_{1-74}$, Aβ-CT, PrP-CT and GAV-CT. (a) The 66–74 residues of α-Syn$_{1-74}$, Aβ-CT, PrP-CT and GAV-CT. (b) Time course of aggregation of α-Syn$_{1-65}$ (square), α-Syn$_{1-74}$ (circle), Aβ-CT (uptriangle), PrP-CT (down triangle) and GAV-CT (star) as determined by ThT fluorescence assay. Data are the average of triplicate incubations (mean ± SEM, next as same). (c) AFM images of α-Syn$_{1-65}$, α-Syn$_{1-74}$, Aβ-CT, PrP-CT and GAV-CT. The photographs were obtained from the samples incubated at 37°C for 6 days with continuous shaking. All graphs are topographical height images 1 μm^2 in area and the brightness represents the heights of areas. The scale bar denotes 200 mm (next as same).

the insertion mutant forms amorphous precipitates but not regular filaments is mainly caused by protein destabilization, demonstrating that the well-folded PGBD is unfavorable to the GAV-induced fibril formation. This study provides further evidence demonstrating that the flexible unstructured regions harboring the GAV homologue are required for fibrillization of host proteins.

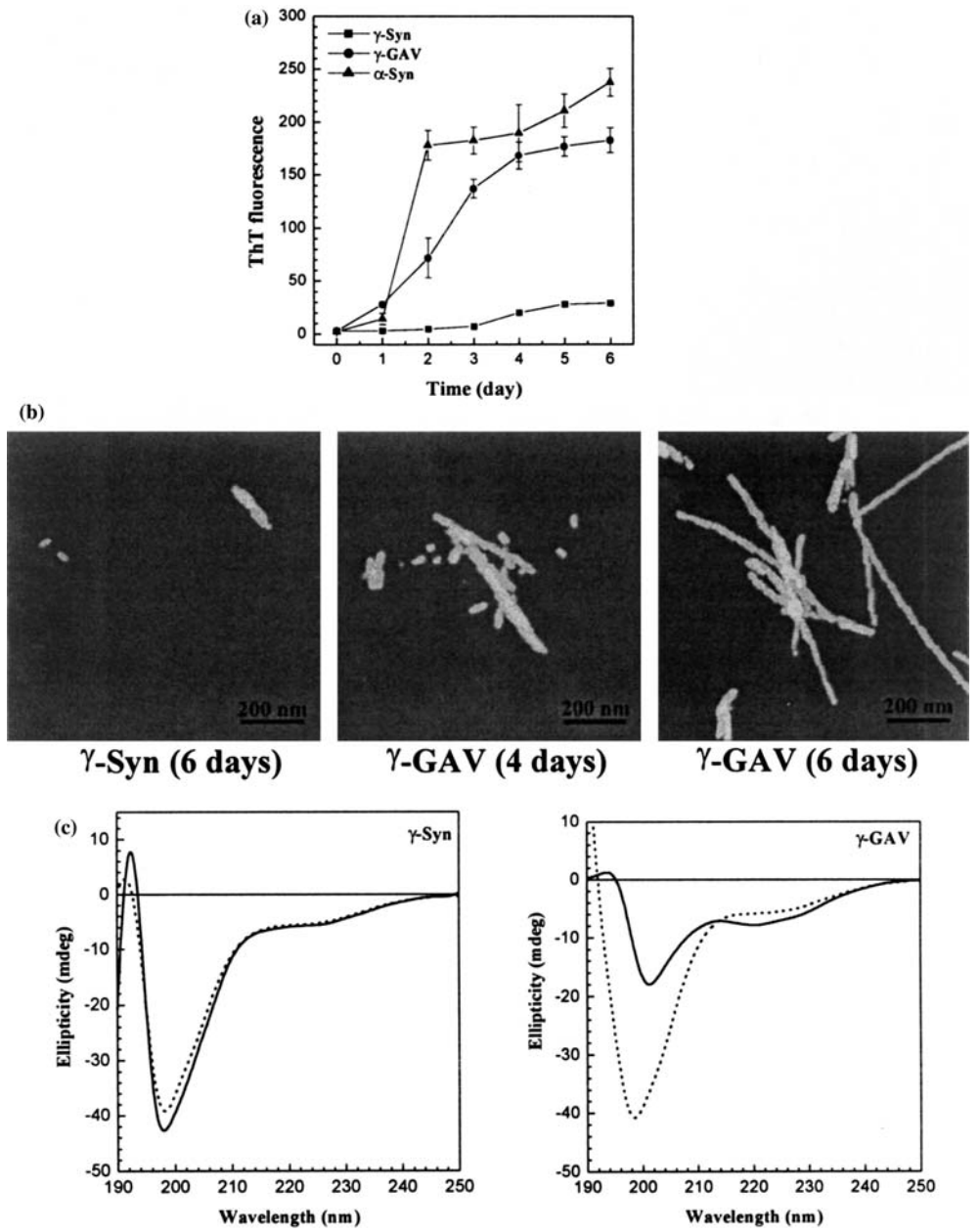

Fig. 2. Amylodgenic propensities of γ-Syn, γ-GAV and α-Syn. (a) Time course of aggregation of γ-Syn (square), γ-GAV (circle) and α-Syn (triangle) as determined by ThT fluorescence assay. (b) AFM images of γ-Syn (left), γ-GAV incubated for 4 days (middle) and γ-GAV incubated for 6 days (right). (c) Circular dichroic spectra of γ-Syn (left) and γ-GAV (right) before (dashed) and after (solid) incubation for 6 days.

4. DISCUSSION

Amyloidogenic aggregation or fibrillization, the possible causative pathogen of abnormal cellular degeneration, is the intrinsic feature of a protein relevant to its primary sequence and secondary or tertiary structure (Clayton and George, 1998; Dobson, 1999). There are diverse amyloidogenic proteins with different sequences and structures, and many of them contain a core sequence and a large flexible unstructured region (Uversky and Fink, 2004). Accumulating evidence show that there is a critical core for initiating the fibrillization of some amyloidogenic proteins that are associated with neurode-

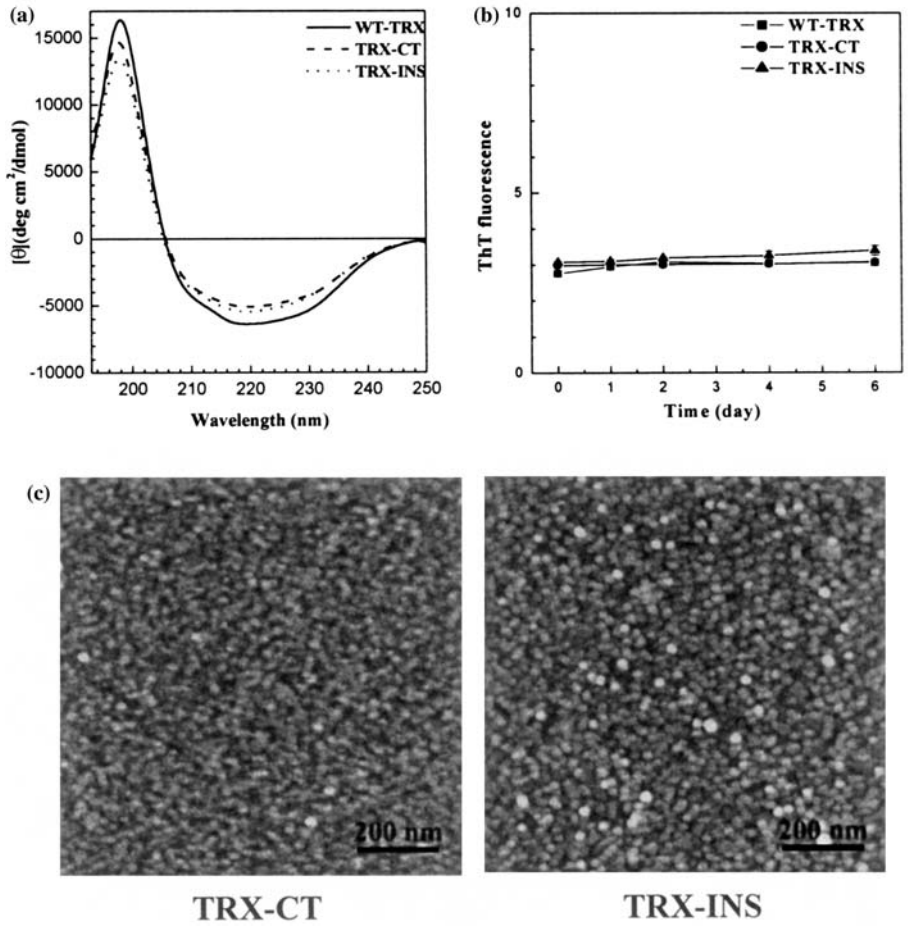

Fig. 3. Aggregation of the GAV homologue-harboring mutants based on the TRX host. (a) Far-UV spectra of WT-TRX (solid), TRX-CT (dashed) and TRX-INS (dotted), showing indistinguishable secondary structures of the three proteins. (b) Time course of aggregation of WT-TRX (square), TRX-CT (circle) and TRX-INS (triangle) as measured by ThT fluorescence. (c) AFM images of TRX-CT (left) and TRX-INS (right) after incubation for 6 days.

generative diseases and abnormal phenotypes of yeast or or fungus *Podospora anserina* (Table 1). α-Syn, Aβ and PrP have their own core sequences and share a conserved consensus, the GAV homologue. Besides the GAV homologue, other core sequences have been found to promote fibrillization, such as ^{306}VQIVYK^{311} in Tau protein (von Bergen *et al*., 2000), polyQ in Ataxin-3 (Masino *et al*., 2003), Huntingtin (Li and Li, 1998; Scherzinger *et al*., 1999) and Androgen Receptor (Michalik and Van, 2003; Ross, 2002), Q/N rich region in yeast prions (Baxa *et al*., 2003; King *et al*., 1997; Taylor *et al*., 1999), polyA in OPMD-associated poly(A) binding protein 2 (PABP2) (Brais *et al*., 1998; Shanmugam *et al*., 2000) and ^{20}SNNFGAILSS^{29} in islet amyloid polypeptide (IAPP) (Green *et al*., 2003; Higham *et al*., 2000). The newly identified HET-s protein from *Podospora anserina* contains an N-terminal compact domain and a C-terminal unstructured region. The flexible C-terminal region is responsible for its amyloidogenic fibrillization, which is closely related to the [Het-s] phenotype (Balguerie *et al*., 2003; Dos *et al*., 2002).

Nevertheless, the core sequences are not sufficient for fibrillization of proteins, though some synthetic peptides corresponding to the core sequences have been reported to be capable of self assembly into fibrils at high concentration (Bodles *et al*., 2001; Garzon-Rodriguez *et al*., 2000; Giasson *et al*., 2001; Satheeshkumar and Jayakumar, 2003). For some amyloidogenic proteins, the flexible unstructured regions where the core sequences located might benefit the fibril formation as shown in Table 1, α-Syn (Weinreb *et al*., 1996), Aβ (Terzi

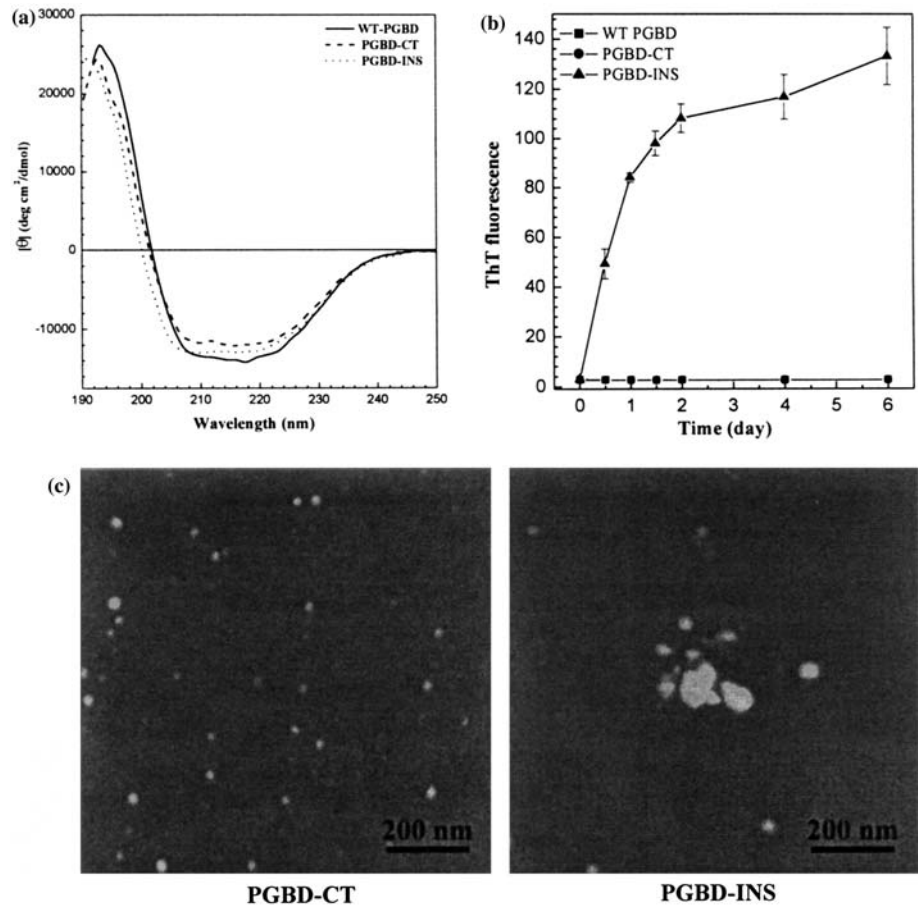

Fig. 4. Aggregation of the GAV homologue-harboring mutants based on the PGBD host, (a) Far-UV CD spectra of WT-PGBD (solid), PGBD-CT (dashed) and PGBD-INS (dotted), showing indistinguishable secondary structures of the three proteins. (b) Time course of aggregation of WT-PGBD (square), PGBD-CT (circle) and PGBD-INS (triangle) as measured by ThT fluorescence. (c) AFM images of PGBD-CT (left) and PGBD-INS (right) after incubation for 6 days.

et al., 1995) and Tau protein (von *et al.*, 2000) are entirely unfolded in aqueous solutions, while the N-terminus (residues 1–127) of PrP (Zahn *et al.*, 2000) displays disordered structure. Other amyloidogenic proteins, such as Ataxin-3 (Masino *et al.*, 2003), Huntingtin (Li and Li, 1998; Scherzinger *et al.*, 1999), Androgen Receptor (Michalik and Van, 2003; Ross, 2002), PABP2 (Brais *et al.*, 1998; Shanmugam *et al.*, 2000) and IAPP (Green *et al.*, 2003; Higham *et al.*, 2000) in human diseases and Sup35p (King *et al.*, 1997) and Ure2p (Baxa *et al.*, 2003; Taylor *et al.*, 1999) in yeast abnormal phenotypes have been characterized to have a flexible unstructured region. Moreover, GAV homologue fail to induce the compact hosts, such as TRX (Fig. 3) and PGBD (Fig. 4), to form fibrils under physiological condition, However, when GAV homologue is attached to or inserted into the flexible unfolded host, such as α-Syn$_{1\text{-}65}$ and γ-Syn, the resultant proteins become the amyloidogenic (for instance, α-Syn$_{1\text{-}74}$ and γ-GAV) as shown in Figs. 1 and 2. This is reminiscent of the previous observation that the flexible unstructured region may readily undergo structural transformation into β-sheet structure and consequently contribute to the fibrillization (Du *et al.*, 2003). With extensive exploration of the mechanism underlying the disease-associated aggregation, more and more amyloidogenic proteins with a core sequence located in the flexible unstructured region will be identified to support this hypothesis.

Protein amyloidogenic fibrillization is complicated but very important for the understanding of cellular dysfunction (Aguzzi and Haass, 2003). The insolubility and unfolded structure of these proteins make them difficult to be studied by biochemical and

Table 1. Amyloidogenic Proteins Possessing a Core Sequence Located in the Flexible Unstructured Region are Associated with Neurodegenerative Diseases and Abnormal Phenotypes of Yeast or Fungus *Podospora anserina*

Protein or peptide	Core sequence	Unstructured region	Disease or phenotype
α-Syn	Residues 66–74	Full-length	PD
Aβ	Residues 36–44	Full-length	AD
PrP	Residues 113–127	N-terminus (1–127)[a]	PrD
Tau protein	Residues 306–311 (VQIVYK)	Full-length	AD
Ataxin-3	polyQ	C-terminus (292–364)	SCA3/MJD[b]
Huntingtin	polyQ	N-terminus	HD
Androgen Receptor	polyQ	N-terminus	SBMA[c]
Sup35p	Q/N rich	N-terminus (2–114)	[PSI$^+$]
Ure2P	Q/N rich	N-terminus (1–89)	[URE3]
PABP2	polyA	N-terminus	OPMD[d]
IAPP	Residues 20–29 (SNNFGAILSS)	Full-length	Type 2 diabetes
HET-s	218–289	C-terminus (218–289)	[Het-s]

[a]The numeral in the parentheses denotes the flexible unstructured region reported in the literatures.
[b]SCA3/MJD, spino-cerebellar ataxia/Machado-Joseph disease.
[c]SBMA, spinal and bulbar muscular atrophy.
[d]OPMD, oculopharyngeal muscular dystrophy.

structural methods. Our studies by mutagenesis combined with nanoscale imaging techniques will provide further insight into the intrinsic properties and fibril morphologies of the amyloidogenic proteins related to disease pathogenesis (Khurana *et al.*, 2003; Hoyer *et al.*, 2004). Current research focusing on the role of the unstructured region in fibrillization will provide a possibility to elucidate the molecular mechanism underlying the fibrillization of α-Syn, Aβ and PrP as well as other amyloidogenic proteins.

ACKNOWLEDGMENTS

We thank Drs. Lin Tang, Ai-Xin Song and Miss Xiao-Jing Lin for valuable discussion and technical assistance. Financial support was provided by the National Natural Science Foundation of China (NSFC39990600 and NSFC30070165) and the Chinese Academy of Sciences (KSCX2-SW-209, STZ00-07) and Shanghai Commission of Science and Technology (0159-NM078).

REFERENCES

Aguzzi, A., and Haass, C. (2003). *Science* **302:** 814–818.
Alexander, P., Fahnestock, S., Lee, T., Orban, J., and Bryan, P. (1992). *Biochemistry* **31:** 3597–3603.
Balguerie, A., Dos-Reis, S., Ritter, C., Chaignepain, S., Coulary-Salin, B., Forge, V., Bathany, K., Lascu, I., Schmitter, J. M., Riek, R., and Saupe, S. J. (2003). *EMBO J.* **22:** 2071–2081.
Baxa, U., Taylor, K. L., Wall, J. S., Simon, M. N., Cheng, N., Wickner, R. B., and Steven, A. C. (2003). *J. Biol. Chem.* **278:** 43717–43727.
Biere, A. L., Wood, S. J., Wypych, J., Steavenson, S., Jiang, Y., Narhi, L., Anafi, D., Jacobsen, F. W., Jarosinski, M. A., Wu, G. M., Louis, J. C., Martin, F., Narhi, L. O., and Citron, M. (2000). *J. Biol. Chem.* **275:** 34574–34579.
Bodles, A. M., Guthrie, D. J. S., Greer, B., and Irvine, G. B. (2001). *J. Neurochem.* **78:** 384–395.
Brais, B., Bouchard, J. P., Xie, Y. G., Rochefort, D. L., Chretien, N., Tome, F. M., Lafreniere, R. G., Rommens, J. M., Uyama, E., Nohira, O., Blumen, S., Korczyn, A. D., Heutink, P., Mathieu, J., Duranceau, A., Codere, F., Fardeau, M., Rouleau, G. A., and Korcyn, A. D. (1998). *Nat. Genet.* **18:** 164–167.
Clayton, D. E., and George, J. M. (1998). *Trends Neurosci.* **21:** 249–254.
Come, J. H., Fraser, P. E., and Lansbury, P. T. Jr. (1993). *Proc. Nall. Acad. Sci. USA* **90:** 5959–5963.
Dobson, C. M. (1999) *Trends Biochem Sci.* **24:** 329–332.
Dos Reis, S., Coulary-Salin, B., Forge, V., Lascu, I., Begueret, J., and Saupe, S. J. (2002). *J. Biol. Chem.* **277:** 5703–5706.
Du, H. N., Tang, L., Luo, X. Y., Li, H. T., Hu, J., Zhou, J. W., and Hu, H. Y. (2003). *Biochemistry* **42:** 8870–8878.
Dyson, H. J., Holmgren, A., and Wright, P. E. (1989). *Biochemistry* **28:** 7074–7087.
Forloni, G., Angeretti, N., Chiesa, R., Monzani, E., Salmona, M., Bugiani, O., and Tagliavini, F. (1993). *Nature* **362:** 543–546.
Garzon-Rodriguez, W., Vega, A., Sepulveda-Becerra, M., Milton, S., Johnson D., A., Yatsimirsky, A. K., and Glabe, C. G. (2000). *J. Biol. Chem.* **275:** 22645–22649.
Giasson, B. I., Murray, I. V., Trojanowski, J. Q., and Lee, V. M. (2001). *J. Biol. Chem.* **276:** 2380–2386.
Goedert, M. (2001) *Rev. Neurosci.* **2:** 492–501.
Green, J., Goldsbury, C., Mini, T., Sunderji, S., Frey, P., Kistler, J., Cooper, G., and Aebi, U. (2003). *J. Mol. Biol.* **326:** 1147–1156.

Gronenborn, A. M., Filpula, D. R., Essig, N. Z., Achari, A., Whitlow, M., Wingfield, P. T., and Clore, G. M. (1991). *Science* **253:** 657–661.
Halverson, K., Fraser, P. E., Kirschner, D. A., and Lansbury, P. T. Jr. (1990). *Biochemistry* **29:** 2639–2644.
Han, H., Weinreb, P. H., and Lansbury, P. T. Jr. (1995). *Chem. Biol.* **2:** 163–169.
Harper, J. D., and Lansbury, P. T. Jr. (1997). *Annu. Rev. Biochem.* **66:** 385–407.
Higham, C., Jaikaran, E., Fraser, P., Gross, M., and Clark, A. (2000). *FEBS Lett.* **470:** 55–60.
Holmgren, A., Soderberg, B. O., Eklund, H., and Branded, C. I. (1975). *Proc. Natl. Acad. Sci. USA* **72:** 2305–2309.
Hoyer, W., Cherny, D., Subramaniam, V., and Jovin, T.M. (2004). *J. Mol. Biol.* **340:** 127–139.
Hu, H. Y. (2001) *Chin. Sci. Bull.* **46:** 1–3.
Hu, H. Y., Li, Q., Cheng, H. Q., and Du, H. N. (2001). *Biopolymers* **62:** 15–21.
Iwatsubo, T., Odaka, A., Suzuki, N., Mizusawa, H., Nukina, N., and Ihara, Y. (1994). *Neuron* **13:** 45–53.
Jarrett, J. T., Berger, E. P., and Lansbury, P. T. Jr. (1993). *Biochemistry* **32:** 4693–4697.
Jarrett, J. T., and Lansbury, P. T. Jr. (1993). *Cell* **73:** 1055–1058.
Ji, H., Liu, Y. E., Jia, T., Wang, M., Liu, J., Xiao, G., Joseph, B. K., Rosen, C., and Shi, Y. E. (1997). *Cancer. Res.* **57:** 759–764.
Khurana, R., Ionescu-Zanetti, C., Pope, M., Li, J., Nielson, L., Ramirez-Alvarado, M., Regan, L., Fink, A. L., and Carter, S. A. (2003). *Biophys. J.* **85:** 1135–1144.
King, C. Y., Tittmann, P., Gross, H., Gebert, R., Aebi, M., and Wuthrich, K. (1997). *Proc. Natl. Acad. Sci. USA* **94:** 6618–6622.
Li, S. H., and Li, X. J. (1998). *Hum. Mol. Genet.* **7:** 777–782.
Lin, M. C., Mirzabekov, T., and Kagan, B. L. (1997). *J. Biol. Chem.* **272:** 44–47.
Masino, L., Musi, V., Menon, R. P., Fusi, P., Kelly, G., Frenkiel, T. A., Trottier, Y., and Pastore, A. (2003). *FEBS Lett.* **549:** 21–25.
Michalik, A., and Van Broeckhoven, C. (2003). *Hum. Mol. Genet.* **12:** 173–186.
Pace, C. N. (1989) *Methods Enzymol.* **131:** 266–280.
Prusiner, S. B. (1996) *Trends Biochem. Sci.* **21:** 482–487.
Rochet, J. C., and Lansbury, P. T. Jr. (2000). *Curr. Opin. Struct. Biol.* **10:** 60–68.
Ross, C. A. (2002) *Neuron* **35:** 819–822.
Rymer, D. L., and Good, T. A. (2000). *J. Neurochem.* **75:** 2536–2545.
Satheeshkumar, K. S., and Jayakumar, R. (2003). *Biophys. J.* **85:** 473–483.
Scherzinger, E., Sittler, A., Schweiger, K., Heiser, V., Lurz, R., Hasenbank, R., Bates, G. P., Lehrach, H., and Wanker, E. E. (1999). *Proc. Natl. Acad. Sci. USA* **96:** 4604–4609.
Shanmugam, V., Dion, P., Rochefort, D., Laganiere, J., Brais, B., and Rouleau, G. A. (2000). *Ann. Neurol.* **48:** 798–802.
Soto, C. (2003) *Nat. Rev. Neurosci.* **4:** 49–60.
Spillantini, M. G., Schmidt, M. L., Lee, V. M., Trojanowski, J. Q., Jakes, R., and Goedert, M. (1997). *Nature* **388:** 839–840.
Taylor, K. L., Cheng, N., Williams, R. W., Steven, A. C., and Wickner, R. B. (1999). *Science* **283:** 1339–1343.
Terzi, E., Holzemann, G., and Seelig, J. (1995). *J. Mol. Biol.* **252:** 633–642.
Uversky, V. N., and Fink, A. L. (2004). *Biochim. Biophys. Acta* **1698:** 131–153.
Uversky, V. N., Li, J., Souillac, P., Millett, I. S., Doniach, S., Jakes, R., Goedert, M., and Fink, A. L. (2002). *J. Biol. Chem.* **277:** 11970–11978.
von Bergen, M., Friedhoff, P., Biernat, J., Heberle, J., Mandelkow, E. M., and Mandelkow, E. (2000). *Proc. Natl. Acad. Sci. USA* **97:** 5129–5134.
Weinreb, P. H., Zhen, W., Poon, A. W., Conway, K. A., and Lansbury, P. T. Jr. (1996). *Biochemistry* **35:** 13709–13715.
Wood, S. J., Wypych, J., Steavenson, S., Louis, J. C., Citron, M., and Biere, A. L. (1999). *J. Biol. Chem.* **274:** 19509–19512.
Zahn, R., Liu, A., Luhrs, T., Riek, R., Schroetter, C.von, Lopez Garcia, F., Billeter, M., Calzolai, L., Wider, G., and Wuthrich, K. (2000). *Proc. Natl. Acad. Sci. USA* **97:** 145–150.