Korean J. Chem. Eng., **29**(6), 693-702 (2012)
DOI: 10.1007/s11814-012-0060-x

INVITED REVIEW PAPER

Techniques for monitoring protein misfolding and aggregation in vitro and in living cells

Simpson Gregoire*, Jacob Irwin*, and Inchan Kwon*,,†**

*Department of Chemical Engineering, University of Virginia, Charlottesville, Virginia 22904, USA
**Institutes on Aging, University of Virginia, Charlottesville, Virginia 22904, USA
(*Received 7 April 2012 • accepted 4 May 2012*)

Abstract−Protein misfolding and aggregation have been considered important in understanding many neurodegenerative diseases and recombinant biopharmaceutical production. Various traditional and modern techniques have been utilized to monitor protein aggregation in vitro and in living cells. Fibril formation, morphology and secondary structure content of amyloidogenic proteins in vitro have been monitored by molecular probes, TEM/AFM, and CD/FTIR analyses, respectively. Protein aggregation in living cells has been qualitatively or quantitatively monitored by numerous molecular folding reporters based on either fluorescent protein or enzyme. Aggregation of a target protein is directly correlated to the changes in fluorescence or enzyme activity of the folding reporter fused to the target protein, which allows non-invasive monitoring aggregation of the target protein in living cells. Advances in the techniques used to monitor protein aggregation in vitro and in living cells have greatly facilitated the understanding of the molecular mechanism of amyloidogenic protein aggregation associated with neurodegenerative diseases, optimizing culture conditions to reduce aggregation of biopharmaceuticals expressed in living cells, and screening of small molecule libraries in the search for protein aggregation inhibitors.

Key words: Protein Misfolding, Aggregation, Neurodegenerative Diseases, Amyloid Fibrils

INTRODUCTION

Protein misfolding and aggregation have attracted great attention in recent years. These phenomena are implicated in the onset of numerous human diseases and are also critical issues in the production and formulation of biopharmaceuticals. Misfolding and aggregation of proteins are known to cause numerous neurodegenerative diseases, such as Alzheimer's (AD), Parkinson's (PD), Amyotrophic Lateral Sclerosis (ALS), and Huntington's [1-6]. AD is the most common form of dementia. Currently, 5.3 million people in US are affected, with the number projected to rise to 13.5 million by 2050 [7]. PD is the second most common neurodegenerative disease with nearly 1 million people in the US affected. PD is a brain disorder leading to shaking and difficulty with walking and movement. Amyotrophic lateral sclerosis (ALS), also known as Lou Gehrig's disease, is a rapidly progressing and fatal neuromuscular disease that attacks the nerve cells responsible for controlling voluntary muscles [8]. A common pathological hallmark of these neurodegenerative diseases is the accumulation of insoluble protein aggregates in the central nervous system. AD, PD, and ALS are closely associated with aggregation of toxic amyloid-beta peptide (Aβ), α-synuclein, human copper/zinc superoxide dismutase (SOD1), respectively. Aβ and α-synuclein are accumulated outside of the cells, whereas misfolded SOD1 tends to form insoluble toxic protein aggregates inside cells [9,10]. Therefore, monitoring protein aggregation in vitro and in living cells is essential to understanding the molecular mechanism of neurodegenerative diseases and in identifying drug candidates capable of modulating protein aggregation.

Protein aggregation is also one of the major issues in biopharmaceutical production, storage, and formulation. Numerous recombinant protein drugs have been produced from *E. coli* (reviewed in [11,12]). However, over-expression of recombinant proteins usually results in aggregates (such as, inclusion body) formation inside *E. coli* [13]. Many environmental factors can perturb correctly folded conformation or inhibit correct protein folding; leading to protein misfolding and aggregation (Fig. 1). Once misfolded proteins are formed inside the cells, they can be returned to their correctly folded state with the aid of chaperones or even removed via degradation. However, abnormally increased misfolded proteins create a damaged or over-loaded protein degradation system leading to intracellular aggregate formation. Misfolded or unfolded recombinant proteins in inclusion bodies require an extra refolding process in order to be functional. Furthermore, many proteins that are structurally complex and require multiple disulfide bonds cannot easily refold [14]. In those cases, formation of a soluble target protein is desirable. To

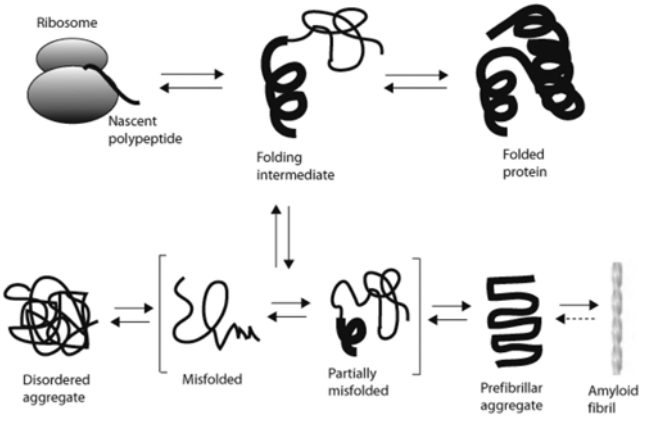

Fig. 1. Scheme of protein folding and misfolding/aggregation.

†To whom correspondence should be addressed.
E-mail: ik4t@virginia.edu

facilitate protein folding by chaperone proteins as well as post-translational modification, more protein drugs have been produced in cultured yeast and mammalian cells in recent years. However, intracellular aggregate formation and misfolding of biopharmaceuticals, such as α-galactosidase, α-glucosidase, antithrombine III, and angiopoietin-1 in mammalian cells remain obstacles in achieving high yield production of biopharmaceuticals [15-19]. To reduce intracellular aggregate formation, cell culture medium composition and environmental factors have been optimized [20]. With advancements in molecular biology techniques and high-throughput screening, less aggregation-prone variants of proteins have been identified by screening mutants [21]. Various methods have been developed to enable monitoring protein aggregation in cultured cells, including bacteria, yeast and mammalian cells.

This review focuses on recent advances in protein aggregation monitoring methods in vitro and in cultured cells. Considering space limitations, it is not possible to review all methods available or all target proteins. Therefore, this review mainly focuses on 1) biophysical and immunological monitoring methods of protein aggregation leading to amyloid fibril formation directly linked to neurodegenerative diseases, and 2) molecular reporter systems monitoring intracellular protein aggregation in bacteria and mammalian cells.

1. Monitoring Protein Aggregation In Vitro

Partially misfolded or misfolded proteins self-assemble to form protein aggregates in vitro. Protein aggregates have diverse structures including disordered aggregates, prefibrillar aggregates, and amyloid fibrils (Fig. 1). Originally, insoluble fibrils were thought to be the principal conformer conferring neurotoxicity in diseases such as AD [22]. As such, the scientific community committed much time and resources into the development and optimization of many traditional in vitro techniques to characterize this particular form of protein aggregate. However, through recent in vivo and in vitro discovery, smaller, prefibrillar aggregates (Fig. 1) have now emerged as the primary toxic species in several of these diseases [23-26]. Prefibrillar aggregates include the following conformers: prefibrillar oligomers (globular aggregates lacking the ordered cross-stacked β-sheet structure) and fibrillar oligomers/protofibrils (aggregates with the cross-stacked β-sheet structure). Because of this paradigm shift, many limitations and potential drawbacks of using traditional fibril monitoring techniques to study prefibrillar aggregates have become apparent. In addition, new techniques (or new uses of traditional techniques) have emerged to better characterize prefibrillar/oligomeric aggregates and the effectiveness of proposed aggregation inhibition therapies. The focus of this section is to review recent trends in the use of several traditional techniques (small molecule probes, TEM/AFM, and CD/FTIR) and a new technique (dot-blot assay using conformational-specific antibodies) for routine monitoring of prefibrillar aggregates in the study of protein aggregation.

1-1. Molecular Probes for Aggregate Characterization

1-1-1. Thioflavin T Binding Assay

The use of a small molecule dye, Thioflavin T (ThT), and its derivatives is arguably the simplest and most widely used method to monitor aggregation of amyloidogenic proteins. ThT displays a fluorescence emission maximum at around the 485 nm wavelength upon binding to the β-sheet groove structure of fibrillar protein aggregates [27]. The use, binding sites, and binding mechanisms of ThT with protein aggregates have been reviewed extensively [28-35].

Traditionally, ThT has been used to detect amyloid fibrils only due to the characteristic sigmoidal increase in fluorescence that occurs between the monomer and end fibril state [36-40]. However, ThT has also been found to bind prefibrillar aggregates that contain the β-sheet groove binding site (fibrillar oligomers and protofibrils). Prefibrillar aggregates which contain the β-sheet structure have fewer binding sites than fibrils and so their fluorescence is lower but still observable (1.5 fold increase for prefibrillar aggregates vs. 100 fold for fibrils) [32,41]. Therefore, ThT could prove indicative of the presence of toxic protofibrils and fibrillar oligomers, but not prefibrillar oligomers that lack defined β-sheet structure [42,43]. Lastly, caution should be taken when using ThT to monitor aggregation in the presence of aggregation inhibitors due to potential spectral interference with ThT fluorescence [31,36].

1-1-2. Congo Red Binding Assay

Congo Red (CR) is another small molecule probe which, similar to ThT, has traditionally been used to identify amyloid fibrils, specifically in the form of deposits in the brain tissue or in vitro (reviewed in [28,30,32,44]). CR binds to β-sheet rich structures present in amyloid fibrils [31] and demonstrates a characteristic shift in the absorbance maximum from 490 to 540 nm and green birefringence with crossed polarized light [29].

More recent studies have shed light on the use of CR toward examining prefibrillar aggregates. Walsh et al. observed a change in CR absorbance when applied to protofibrils, albeit the change was less marked than that of amyloid fibrils [45]. Additionally, Maezawa et al. applied surface plasmon resonance to identify the binding affinity/dissociation constant of CR towards prefibrillar oligomers [30]. However, a shift in absorbance maximum upon CR binding was not observed for prefibrillar oligomers. So, as was the case with ThT, recent studies demonstrate that CR could be applicable to characterize protofibrils and fibrillar oligomers, but not prefibrillar oligomers that do not have defined stacked β-sheet structure.

1-1-3. ANS Fluorescence Assay

Another small molecule, 1-anilinonaphthalene 8-sulfonate (ANS), and its closely related derivative, 4,4′-bis-1-anilinonaphthalene 8-sulfonate (Bis-ANS), have served since their discovery in the 1950s as one of the most frequently used fluorescent probes for characterizing a diverse array of proteins [35,38,46]. ANS provides an assessment of surface hydrophobicity by showing an increase in fluorescence intensity and a blue shift (decrease in wavelength) in the fluorescence maximum upon being exposed to hydrophobic regions on the surface of proteins [46].

Interaction and subsequent perturbation of hydrophobic lipid membrane bilayers is believed to be one of the primary mechanisms by which prefibrillar aggregates confer toxicity to cells in associated diseases (for an in-depth review of possible toxicity mechanisms, see reference [47]). Thus, to gain insight into the potential for toxicity through hydrophobic interactions, an assessment of surface hydrophobicity of protein aggregates through ANS could be potentially very useful in the study of prefibrillar protein aggregates. Indeed, recent studies have shown that 1) prefibrillar oligomers of Aβ42 peptide exhibited an increase in fluorescence and a blue shift when exposed to ANS compared to fibrils and monomers [48,49] and 2) a correlation between increased ANS fluorescence and toxicity [48-50]. In the specific area of monitoring prefibrillar protein aggregates, ANS has been used less frequently than CR and ThT. Ladi-

wala et al. used ANS to characterize Aβ42 fibrils, freshly disaggregated low molecular weight species, and soluble prefibrillar oligomers and reported that prefibrillar oligomers showed the largest increase in fluorescence and blue shift, characteristic of ANS interactions with hydrophobic regions [37]. These oligomers were also reported to be more toxic than fibrils and low molecular weight species.

1-1-4. Antibody Dot Blot Assay

Because of the difficulty in obtaining high resolution crystal structures of protein aggregates (prefibrillar aggregates, in particular), conformational-specific antibodies have been developed over the past 10 years to help identify and monitor the state of amyloidogenic protein aggregation and screen potential aggregation inhibitors/modulators.

Glabe et al. developed three conformational-specific antibodies that are important to detect physiologically-relevant prefibrillar aggregates: A11 (recognizing prefibrillar oligomers but not fibrillar conformers [23]; OC (recognizing the cross-stacked β-sheet structure of fibrillar oligomers, protofibrils, and fibrils [51]; and αAPF (recognizing annular protofibrils [52]). These conformational-specific antibodies have been extensively used and reviewed [24,30,37,53-55], and a summary of their respective reactivity to the various aggregation conformers is shown in Fig. 2(a). As an example of the A11 antibody to evaluate the inhibitory capacity of Aβ aggregation modulator, inhibition of toxic A11-reactive Aβ aggregate formation by brilliant blue G (BBG), a small molecule inhibitor, is illustrated in Fig. 2(b) [56].

Even though the application of these antibodies has provided important insight into the properties of prefibrillar protein aggregates and the effects of potential therapeutics, recent studies have shown that care must be taken when using and interpreting the results. First, due to the transient nature of prefibrillar aggregates compared to end-state conformers, preparing a homogeneous sample of exclusively A11, OC, or αAPF reactive (no cross reactivity) prefibrillar aggregates in vitro has proven quite challenging, though preparation of homogeneous prefibrillar aggregates was reported by a few research groups [41,43,52]. Despite their success, non-physiological aggregation conditions (low pH) [41] or incomplete removal of pre-treatment disaggregation agent [41,43] were employed in order to generate these homogeneous populations. Other groups who have applied the OC and A11 conformational-specific antibodies have observed more non-homogeneous populations, either in the form of fibrillar conformers populations containing significant A11-reactivity [57-59] or prefibrillar oligomer populations possessing significant OC-reactivity [36,60]. Because of the cross reactivity demonstrated by these populations, it is difficult to discern which species is dominant. Second, in rare but present cases [61], false positive antibody reactivity has been observed when testing the inhibitory/modulatory activity of extrinsic compounds on protein aggregates. Because of these two factors, care must be taken when designing experiments and interpreting the results of these antibodies.

Because of the insight that the three primary conformational-specific antibodies discussed above have given and due to the high degree of polymorphism that exists in prefibrillar aggregates, additional conformational-dependent antibodies are being developed to further characterize these conformers. Wang et al. developed four single-chain variable fragment antibodies that specifically recognize Aβ oligomers, but not monomers or fibrils [62]. However, the use of this antibody does not appear to have caught on in a widespread manner throughout the community. Furthermore, Kayed et al. developed six new monoclonal antibodies that recognize immunologically distinct preparations and sub-variants of prefibrillar oligomers also recognized by the more general polyclonal antibody, A11 [63]. Even though the sub-classes of prefibrillar oligomers identified by the monoclonal antibodies showed different molecular weight size distributions through Western blot analysis, it is not yet clear whether or not these variants are more pathologically relevant than A11 antibody.

1-2. Methods for Obtaining Visual Morphological (Quaternary Structure) Information on Protein Aggregate Species

1-2-1. Electron Microscopy & Atomic Force Microscopy

Transmission electron microscopy (TEM) and atomic force microscopy (AFM) are the two techniques most commonly used to visualize the morphology of amyloidogenic protein aggregate samples [30, 31,36,37,39-41,48,50,58,60,64-67]. Both TEM and AFM provide information (both qualitative and quantitative) at the nanometer level of quaternary structure characteristics, including the length, width,

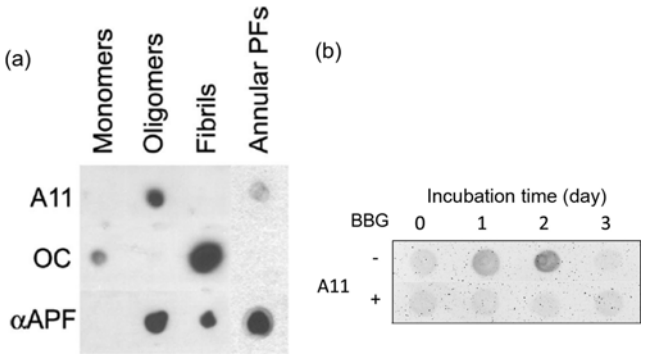

Fig. 2. (a) Summary of reactivity of A11, OC, and α-APF conformational-specific antibodies to various aggregate species [52]. (b) Inhibition of A11-reactive Aβ aggregate formation by brilliant blue G (BBG), a small molecule inhibitor, for 3 days [56].

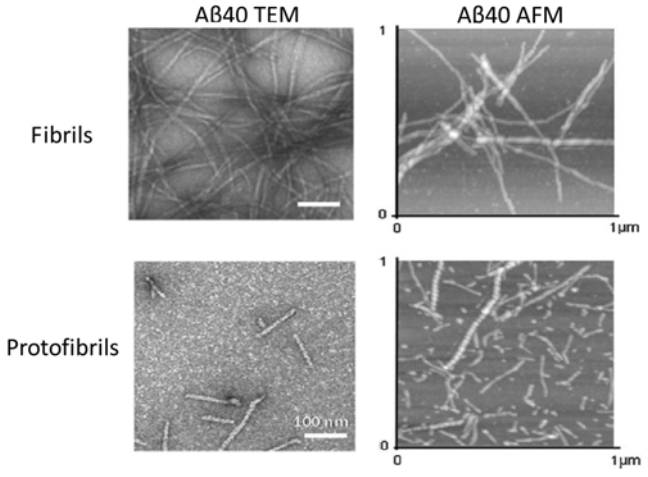

Fig. 3. Aβ40 fibrils (top) and protofibrils (bottom) aggregates visualized using TEM (left) and AFM (right) [56,58]. TEM (left side) scale bar is 100 nm.

curvature, and surface features of protein aggregates (reviewed in [28,35,68-72]). Several examples of Aβ40 fibril (top) and protofibril (bottom) aggregates visualized using TEM (left) and AFM (right) are shown in Fig. 3.

However, there are several significant differences and/or limitations to consider when using these methods to study prefibrillar protein aggregates. Although TEM has the advantage of being a direct method (i.e., crystallization of the sample is not required) [68] and can be performed fairly quickly [70], the degree of resolution that can be obtained for smaller, prefibrillar aggregates less than ~20 nm is fairly limited compared to higher resolution techniques [72]. Because of this limitation, TEM is useful in verifying formation, inhibition, and/or disaggregation of larger protofibrils/fibrils and providing an approximate gauge for morphology of smaller aggregates, but not for yielding high-resolution, low-error details of these small prefibrillar conformers. However, despite this drawback, TEM has been used to provide a numerical estimation of the length or width of smaller aggregates less than 30 nm [42,60,73].

Conversely, AFM provides sub-nanometer three-dimensional (including height) resolution of protein sample characteristics [74] and is thus well-suited for studying smaller prefibrillar aggregates with low expected error. However, the sample preparation for AFM often takes longer and involves more steps (for example, freezing, adsorption to mica surface) than a simple negative-stain TEM sample preparation [70].

1-3. Methods for Obtaining Secondary Structure Information on Protein Aggregate Species

1-3-1. Circular Dichroism & Fourier Transform Infrared Spectroscopy

Circular dichroism (CD) and Fourier transform infrared spectroscopy (FTIR) are two well-established techniques that are used most frequently to assay the secondary structure (β-sheet, α-helix, β-turn, and disordered content) of protein aggregates in vitro [38, 39,66,67]. CD measures the differential absorption of right and left polarized light, and FTIR analyzes molecular bond vibration frequencies [71]. Although it has been reported to be theoretically possible to determine antiparallel vs. parallel β-sheet secondary structure using CD [75,76], FTIR is used more readily to obtain resolution on secondary structures within the β-sheet group.

Secondary structure content is important for the study of protein aggregates because specific secondary structures are characteristic of different stages in the aggregation pathways. For example, in the Aβ peptide associated with AD, monomers and small oligomers have been found to consist of mainly unordered/α-helical structures, whereas the intermediate fibrillar oligomers, protofibril, annular protofibril, and ending fibril conformers contain mostly β-sheet secondary structure [41,52,54,77]. Unlike the other conformers, there is a considerable amount of disagreement in the literature regarding the secondary structure of soluble prefibrillar oligomers. On one hand, several research groups have reported that Aβ oligomers contain mostly random coil/disordered secondary structure [42,43], while others have reported these oligomers possessing prevailing parallel or antiparallel β-sheet content [57,58,73]. Given the transient nature of this aggregate species and the difficulty in preparing a homogeneous conformer sample, it is not surprising that different research labs obtained different results.

Despite extensive general reviews of the techniques [28,35,71,72, 78-80] and the vast amount of primary works employing CD and FTIR to assess protein aggregates, a single, straight-forward protocol does not exist that details how to use these methods to estimate secondary structure content percentages. Instead, researchers employ a host of scientifically acceptable approaches to estimate secondary structure content from the raw CD and FTIR data generated. Often, a qualitative-based analysis is used by comparing general features of the CD or FTIR spectra with controls or "expected" results, which are used to correlate secondary structure changes or overall content. For example, the CD or FTIR spectra of an unknown protein aggregate sample could be compared to a predominantly β-sheet fibril

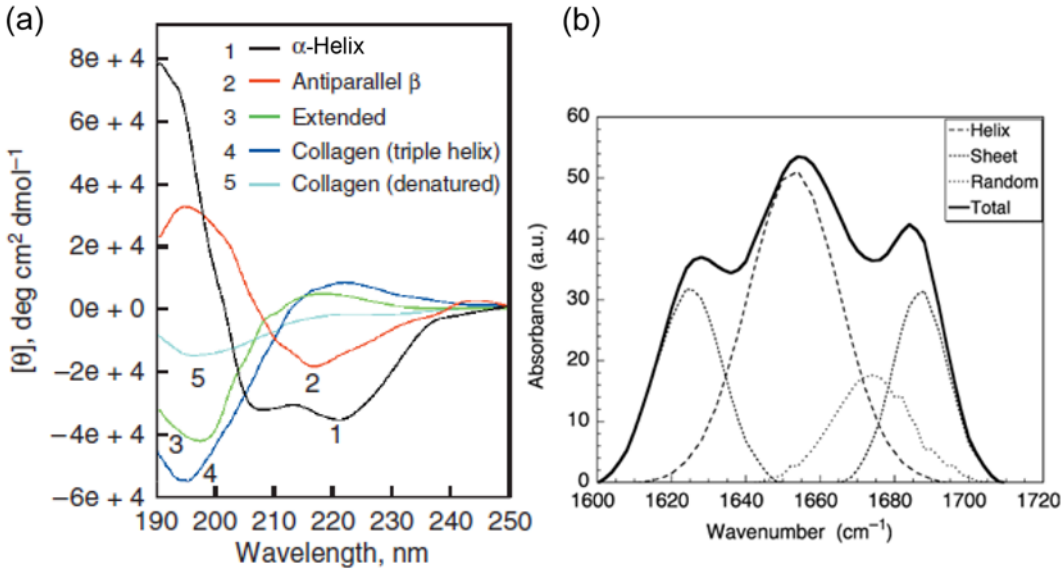

Fig. 4. (a) Characteristic β-sheet (2-red line), α-helix (1-black line), and unordered/random coil (5-light blue line) secondary structure CD spectra [85]. (b) Decomposition of total FT-IR spectrum (solid line) in to β-sheet (short dash line), α-helix (long dash line), and unordered/random coil (dotted line) secondary structure components [111].

control spectra, and deviations could be visualized [41]. Similarly, spectral features of an unknown sample, such as curve minima and maxima between 190-250 nm for CD and 1,600-1,700 cm^{-1} for FTIR are often compared to established "expected" correlations to obtain a qualitative assessment of the major secondary structure features [43,59,60,64]. Correlations of CD and FTIR curve characteristics to secondary structure have been reviewed [28,81]. Examples of the characteristic β-sheet, α-helix, and unordered/random coil secondary structure CD spectra and FTIR decomposition component curves are shown in Fig. 4. Alternatively, a more complex, but quantitative analysis has been developed and utilized in order to estimate the percentage of secondary structure content of a sample from CD spectral data using one or more single value decomposition, regression, or neural network algorithms available [82-84]. Several examples of these methods/algorithms include CDSSTR, SELCON, CONTIN, and K2D. For a review of these numerical methods, see [85]. It is important to note that the majority of these algorithms require that the sample CD spectra is 'matched' to a protein with known secondary structure content within various protein reference sets [81]. Because amyloidogenic proteins (or their parent proteins) and disordered/random coil proteins are often not included in these reference sets, caution should be taken in interpreting CD spectra of disordered prefibrillar aggregates. Additionally, the numerical results obtained from the analysis should agree (at least directionally) with what is seen upon visual inspection of the sample spectra, and the error value (normalized root mean square deviation) should not be the only criterion used to select the best fit.

2. Monitoring Protein Misfolding/Aggregation in Living Cells

To date, numerous methods to monitor protein misfolding/aggregation in living cells have been developed. The protein-based folding reporters can be categorized based on either phenotype (fluorescence vs. enzyme activity) or folding reporter principle (C-terminus fusion vs. complementation) (Fig. 5).

2-1. Monitoring Protein Misfolding/Aggregation Using Fluorescent Reporter Proteins

Fluorescent proteins including green fluorescent protein (GFP) and its variants are widely used tools in cell biology. Fluorescent proteins can be fused to a target protein to visualize the target protein expressed inside cells. Furthermore, by fusing the fluorescent protein to a target protein, the extent of target protein misfolding and aggregation can be correlated to cellular fluorescence intensity; typically measured with a flow cytometer or fluorescence plate reader. The focus of this section is to summarize recent developments of fluorescence reporters of protein aggregation used in living cells.

2-1-1. Fluorescence Reporter Fusion to C-terminus of a Target Protein

Waldo et al. first developed a folding reporter GFP (frGFP) fused to the C-terminus of a target protein to examine the aggregation and folding status of the target protein in *E. coli* [21]. Cellular fluorescence intensity of cells expressing frGFP fused to the target protein is directly proportional to how correctly the target protein has folded. As the extent of the target protein misfolding and aggregation increases, solubility of the target protein inside cells decreases and then cellular fluorescence intensity decreases (Fig. 6). Using the frGFP fusion technique, Waldo et al. successfully evolved two insoluble proteins, C33T of gene V protein and bullfrog H-subunit ferritin to highly soluble variants via several rounds of directed evolution [21]. Kim et al. fused frGFP to Aβ40 and Aβ42 peptides and screened Aβ peptide mutant libraries expressed in *E. coli* in order to identify Aβ variants with enhanced or reduced aggregation propensity [86,87]. These findings clearly demonstrate that frGFP fusion could be a convenient tool to design/identify soluble variants of other insoluble proteins.

A derivative of GFP, venus yellow fluorescent protein (vYFP) has been used as a folding reporter [88]. In comparison to GFP, vYFP forms its fluorophore at a faster rate, has increased folding efficiency and remains stable in harsh cellular environments. Arslan et al. reported the use of vYFP as a reporter protein to monitor the aggregation of Aβ42 peptide in cell-free systems. Arslan et al. screened a library of peptides and identified peptides that facilitate the folding of Aβ42 in a *cis*-manner in *E. coli* [88,89].

Similar to GFP and its derivatives, Heddle et al. tested the reef

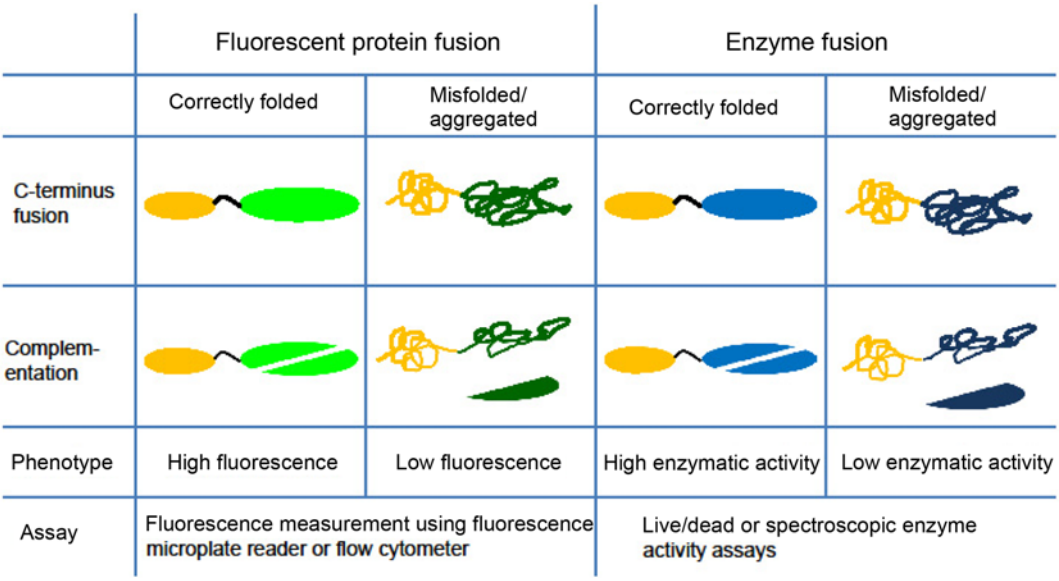

Fig. 5. Comparison of protein-based folding reporters.

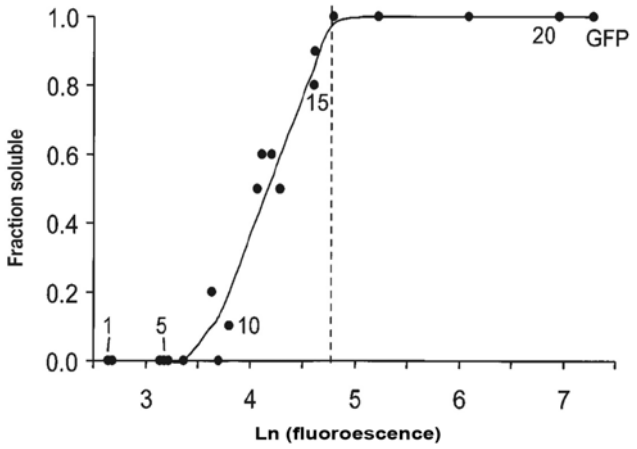

Fig. 6. Correlation of solubility of proteins expressed in *E. coli* with fluorescence of cells expressing corresponding proteins fused to GFP [21].

coral fluorescent protein (RCFP) as a aggregation reporter [90]. The tertiary structures of the RCFPs and GFP are similar and RCFP emits fluorescence at similar levels as GFP, making RCFPs an alternative potential folding reporter. ZsGreen, (an RCFP variant), shows the most promise as a folding reporter, because it can differentiate between poorly folded and well folded target proteins without altering the aggregation propensity of the target protein.

The utility of the GFP as folding reporter has been extended to mammalian cells. Gregoire et al. utilized enhanced GFP (EGFP) as a folding reporter of human copper/zinc superoxide dismutase (SOD1) in mammalian cells [91]. Using flow cytometry to quantify cellular fluorescence intensity of HEK293T and NSC-34 cells expressing the SOD1-EGFP fusion protein, they accurately determined deviations from the correctly folded wild-type SOD1 conformation. It was the first report directly correlating the average cellular fluorescence intensity to the extent of misfolding/aggregation of a target protein in mammalian cells.

The frGFP fusion has also been used to identify inhibitors of protein aggregation. Using the frGFP fused to Aβ42 peptide expressed in *E. coli*, Kim et al. screened small molecule libraries and successfully identified Aβ42 aggregation inhibitors [92]. Geng et al. utilized Aβ-enhanced cyan fluorescent protein fusion protein as a folding reporter of Aβ peptide in *E. coli* in order to screen inorganic drug candidates, polyoxometalates (POMs) in the search for Aβ aggregation inhibitors [93]. Several POMs identified inhibited Aβ aggregation in vitro and also reduced Aβ-associated cytotoxicity. Searching through compounds in such a manner allows for parallel experiments and the ability to screen a vast library of compounds in a timely fashion.

Another quantitative measure of protein misfolding/aggregation using GFP is fluorescence resonance energy transfer (FRET) from GFP to blue fluorescent protein, a blue-shifted form of GFP [94]. Tagging each of these fluorescent proteins to the N- and C-terminus of the target protein, the folding status of the target protein could be determined by the FRET signal from the two proteins.

2-1-2. Complementation of Split Fluorescent Reporters

Many of the monitoring systems of protein misfolding/aggregation involve indirect C-terminus fusions using a fluorescent protein or antibiotic resistance proteins. However, these systems increase the probability of "false positives" due to proteolytic cleavage or increased aggregation of a target protein caused by the soluble nature of the reporter protein. As an alternative to frGFP fused to the C-terminus of a target protein, Cabantous et al. developed self-complementary GFP fragments (split GFP) to monitor aggregation of a target protein [95-97]. A short, non-fluorescent 15-amino acid portion (GFP 11) of GFP is tagged to the C-terminus of a target protein. To complete the fluorophore formation of GFP, the other portion of GFP (GFP 1-10) must complement the GFP11 fragment fused to the target protein. Target proteins that are misfolded or insoluble will reduce the accessibility of this portion of the GFP, making self-complementation difficult and reduce the fluorescence of cells expressing both split GFP fragments, whereas highly soluble and well folded target proteins will allow complementation of split GFP fragments leading to high fluorescence.

The split GFP complementation has been successfully used to detect aggregation of target proteins both in bacteria and mammalian cells. Johnson et al. reported that split GFP complementation can be used to monitor aggregation of mutant tau proteins, which is associated with Alzheimer's disease, in mammalian cells [98]. This was the first time the split GFP system was used to quantify protein aggregation in mammalian cells. Listwan et al. modified this system to a high-throughput format, utilizing an automated Biomek FX liquid-handling robot coupled with Cytomat incubators, DTX plate reader, Rotanta centrifuge, and ORCA transportation rail [99].

Many proteins in living cells form protein complexes including heterodimers. Heterodimer formation can increase or decrease stability of each protein subunit. There are potential issues in examining the aggregation of a heterodimer using the split GFP system. If the proteins in a complex are soluble when expressed separately, but insoluble once the complex is formed, the split GFP system cannot accurately determine the level of aggregation of the protein complex. To address this issue, Lockard et al. added an additional hexa-histidine tag to the N-terminus of the target protein while keeping the GFP11 tag on the C-terminus of the target protein [100]. Once cells expressing the fusion protein are lysed, heterodimers that are not aggregated have the hexa-histidine tag available for binding to affinity resin. However, if heterodimers are aggregated, limited accessibility of the hexa-histidine tag would prevent binding of the heterodimer aggregates to affinity resin. By incorporating an affinity tag to the detection system as well as split GFP complementation, a DNA fragment library of the human protein p85 and breakpoint cluster region-homology were screened to identify stable, soluble heterodimeric protein complexes.

2-2. Monitoring Protein Misfolding/Aggregation Using Enzyme Reporters

Numerous enzyme-based protein misfolding and aggregation reporters have been developed. Similar to the GFP reporter, the enzyme reporter activity is strongly correlated to enzyme folding. Any deviations from the correct structure will reduce the ability of the enzyme to catalyze a particular reaction. Here, the most recent developments in monitoring protein misfolding/aggregation using enzymatic activity are summarized.

2-2-1. Enzyme Reporter Fusion to C-terminus of a Target Protein

The first successful enzyme-based reporter of protein aggregation is chloramphenicol acetyltransferase (CAT) [101]. CAT is a

homo-trimeric protein that retains its enzymatic activity even when fused to the C-terminus of another protein. Maxwell et al. reported that mutants of an insoluble HIV integrase fused to CAT were screened to identify soluble mutant of HIV integrase by testing chloramphenicol antibiotic resistance of *E. coli* expressing the mutant HIV integrase fused to CAT [101]. This result indicates that the extent of aggregation of a target protein fused to CAT is directly correlated to chloramphenicol antibiotic resistance.

Murine dihydrofolatereductase (mDHFR) is an enzyme that is essential for reducing dihydrofolic acid to tetrahydrofolic acid and is imperative for *E. coli* survival. Trimethoprim (TMP) is also known to inhibit the activity of *E. coli* DHFR, but has no effect on mDHFR. As a folding reporter, mDHFR is tagged to the N-terminus of a target protein. Similar to the frGFP, the folding of mDHFRis highly dependent upon the folding of the target protein fused to mDHFR. When the fusion protein is over-expressed in *E. coli*, only cells that express correctly folded mDHFR will survive when in the presence of TMP. Dyson et al. were able to utilize this system to find protein constructs that are soluble in *E. coli* [102,103]. In accordance with the split GFP system, it was also used to screen a library of fragmented genes of two proteins to determine domains of proteins that are recalcitrant. Minimal perturbation of the target protein upon mDHFR fusion is an attractive characteristic for monitoring protein aggregation as well as its ability to be expressed in high quantities in *E. coli*.

Human DHFR (hDHFR) has also been used as a folding reporter of protein aggregation. As detailed in the work by Morell et al., hDHFR was tagged to the C-terminus of proteins of interest and expressed in *erg6Δ* yeast cells in the presence of methotrexate (MTX) [104]. Similar to the aforementioned mDHFR assay, misfolding/aggregation of the target protein was inversely proportional to the cell viability in the presence of MTX. Since the cell line used is also drug-permeable, this system can be used as a screen for potential aggregation modulators of recalcitrant proteins.

2-2-2. Complementation of Split Enzyme Reporter

Similar to its split-GFP counterpart, researchers have created assays that rely on the self-complementation of enzyme fragments. Wigley et al. developed an assay for assessing protein aggregation propensity using complementation of β-galactosidase (β-Gal) [105]. A smaller β-Gal fragment was fused to the C-terminus of a target protein. Aggregation of a target protein limits accessibility of the target protein-fused β-Gal fragment to a larger β-Gal fragment inhibiting complementation of two β-Gal fragments, which leads to reduction of β-Gal activity. Proteins with various levels of aggregation propensity were examined to determine whether the enzyme activity of β-Gal fusion protein correlates directly with the solubility of the target protein. When expressed in *E. coli*, β-Gal fusion protein allows for direct quantification of the color intensity of the colonies that are grown on plates containing the larger portion of the β-Gal protein.

Foit et al. developed a screening system to monitor aggregation of a target protein using a tripartite fusion system [106]. This fusion protein is designed such that a target protein is flanked by two fragments of the penicillin resistance gene, TEM-1-β-lactamase. TEM-1-β-lactamase is a small, monomeric protein that hydrolyzes the β-lactam ring of penicillin (Fig. 7). When the two fragments are separated, similar to split GFP, the antibiotic resistance protein has no activity. However, when the enzyme is self-assembled, it regains its enzymatic activity.

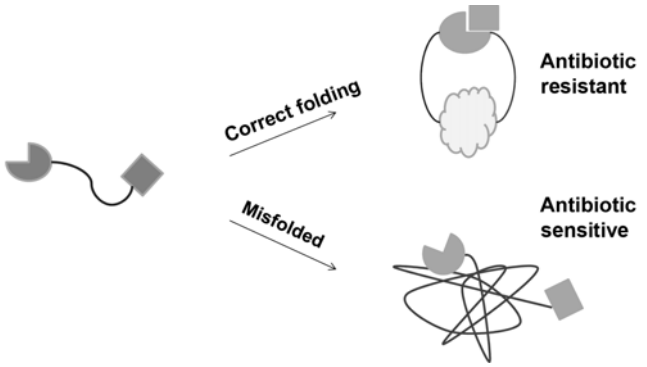

Fig. 7. Tripartite split enzyme assay for monitoring protein aggregation in living cells [112].

The basis for the fusion method is that if a protein is misfolded or aggregated, the two fragments that are fused to the N- and C-termini of the target protein will not be able to self-assemble, leading to reduction of enzymatic activity. Using penicillin as the antibiotic to select bacteria that express well folded fusion proteins, the system is able to give insight on in vitro thermodynamic stability. The benefits of the tripartite fusion protein system are that it reduces the number of false positives that occur frequently in other systems, can be applied to proteins regardless of size and the species the protein is typically expressed in and gives a quantitative output on the thermodynamic stability of the protein. Also, the fusion protein is expressed within the periplasm of the cell, whereas other fusion protein systems are typically expressed within the cytoplasm of the cell. The system has been applied to examining protein folding in periplasmic space, antibody-antigen interactions, and single chain antibody aggregation [107,108].

Yumeredendi et al. developed an aggregation assay where a target protein is fused to a short linear biotin acceptor peptide as well as a hexa-histidine tag at the N-terminus of a target protein [109]. Soluble proteins are more likely to have their peptide accessible for biotin binding, making detection of biotinylation possible using streptavidin. However, aggregated or misfolded proteins will reduce the ability for the peptide to become biotinylated and detected using streptavidin. Using anti-hexa-histidine antibodies, intact and soluble proteins were successfully distinguished from truncated or poorly soluble proteins. This system was first used to determine soluble domains of PB2, the influenza polymerase. The system was also expanded to examining the aggregation of protein complexes, called the Co-ESPRIT system [110]. The fragmented library is cloned between an intein-based open reading frame selection plasmid. The incorporation of the intein-based plasmid will remove many out of frame genes that were present in the initial form of the system. This led to a 9-fold enhancement in screening power, and led to only in frame genes being screened after the first selective pressure was applied to the system [110].

CONCLUSION AND FUTURE PROSPECTS

Various techniques to monitor protein aggregation in vitro and in living cells have been developed. To monitor amyloid fibril formation in vitro, traditional methods, dye binding assay, TEM, CD, and FTIR analysis, have been widely used. More recently, AFM

and dot blot assay using conformation-specific antibodies have been used to characterize physiologically important prefibrillar protein aggregates. Regarding dye binding assay, poor detection of prefibrillar oligomers by current dyes is a critical limitation. Therefore, identification of dyes that preferentially bind to prefibrillar oligomers will allow quantitative monitoring of prefibrillar oligomers. Although dot blot assay using conformation-specific antibodies has been used to monitor formation of prefibrillar aggregates, the binding mechanism of each antibody to the corresponding conformer remains unclear. To enhance the utility of CD and FTIR analyses for monitoring amyloidogenic protein aggregation, two issues need to be resolved. First, reference sets including diverse conformations of amyloidogenic protein aggregates need to be developed to increase the accuracy of quantitative analysis of CD spectra. In particular, addition of more spectra of disordered proteins to existing reference sets will facilitate more accurate analysis of spectra of disordered amyloidogenic aggregates. Second, standard protocols to quantitatively interpret CD and FTIR spectra and to evaluate quality of the analysis need to be established.

To monitor protein aggregation in living cells, various molecular folding reporters using either fluorescent protein or enzyme have been developed. A target protein fusion to either N-terminus of a folding reporter or one of the split reporter proteins directly correlates the target protein aggregation to the loss of either fluorescence or enzyme activity of the reporter, respectively. Despite numerous applications of the molecular folding reporters in *E. coli*, fewer applications have been reported in mammalian cells. Considering the enhanced interest in protein aggregation in mammalian cells, development of molecular folding reporters optimized in mammalian cells is needed.

ACKNOWLEDGEMENT

This work was supported by a National Institutes of Health Grant (NS069946) and a KSEA Young Investigator Grant (I.K.). This work was also supported by an NSF PAGES Scholarship/UVEF graduate Fellowship (S.G) and an NSF Graduate Research Fellowship Program (J. I.).

REFERENCES

1. M. Bartolini and V. Andrisano, *ChemBioChem*, **11**, 1018 (2010).
2. D. W. Colby, J. P. Cassady, G. C. Lin, V. M. Ingram and K. D. Wittrup, *Nat. Chem. Biol.*, **2**, 319 (2006).
3. B. Keshet, J. J. Gray and T. A. Good, *Protein Sci.*, **19**, 2291 (2010).
4. S. Lee, E. J. Fernandez and T. A. Good, *Protein Sci.*, **16**, 723 (2007).
5. R. M. Murphy, *Annu. Rev. Biomed. Eng.*, **4**, 155 (2002).
6. C. J. Pike, A. J. Walencewicz, C. G. Glabe and C. W. Cotman, *Brain Research*, **563**, 311 (1991).
7. A. Alzheimer's, *Alzheimer's & Dementia*, **6**, 158 (2010).
8. D. Majoor-Krakauer, P. J. Willems and A. Hofman, *Clin. Genet.*, **63**, 83 (2003).
9. T. Kabuta, Y. Suzuki and K. Wada, *J. Biol. Chem.*, **281**, 30524 (2006).
10. J. Niwa, S. Yamada, S. Ishigaki, J. Sone, M. Takahashi, M. Katsuno, F. Tanaka, M. Doyu and G. Sobue, *J. Biol. Chem.*, **282**, 28087 (2007).
11. N. Ferrer-Miralles, J. Domingo-Espin, J. Corchero, E. Vazquez and A. Villaverde, *Microbial Cell Factories*, **8**, 17 (2009).
12. K. Graumann and A. Premstaller, *Biotechnol. J.*, **1**, 164 (2006).
13. B. Jurgen, A. Breitenstein, V. Urlacher, K. Buttner, H, Lin, M. Hecker, T. Schweder and P. Neubauer, *Microbial Cell Factories*, **9**, 41 (2010).
14. J. G. Thomas and F. Baneyx, *J. Biol. Chem.*, **271**, 11141 (1996).
15. Y. A. Ioannou, D. F. Bishop and R. J. Desnick, *J. Cell Biol.*, **119**, 1137 (1992).
16. S. Ishii, R. Kase, T. Okumiya, H. Sakuraba and Y. Suzuki, *Biochem. Biophys. Res. Commun.*, **220**, 812 (1996).
17. M. Schröder, R. Schäfer and P. Friedl, *Biotechnol. Bioeng.*, **78**, 131 (2002).
18. T. Okumiya, M. A. Kroos, L. Van Wet, H. Takeuchi, A. T. Van der Ploeg and A. J. J. Reuser, *Mol. Genet. Metab.*, **90**, 49 (2007).
19. S.-J. Hwang, S. Yoon, G. Koh and G. Lee, *Appl. Microbiol. Biotechnol.*, **91**, 305 (2011).
20. N. Jenkins, P. Meleady, R. Tyther and L. Murphy, *Biotechnol. Appl. Biochem.*, **53**, 73 (2009).
21. G S. Waldo, B. M. Standish, J. Berendzen and T. C. Terwilliger, *Nat. Biotechnol.*, **17**, 691 (1999).
22. J. A. H. G. Hardy, *Science*, **256**, 184 (1992).
23. R. Kayed, E. Head, J. L. Thompson, T. M. McIntire, S. C. Milton, C. W. Cotman and C. G. Glabe, *Science (New York)*, **300**, 486 (2003).
24. C. G. Glabe, *J. Biol. Chem.*, **283**, 29639 (2008).
25. M. Bucciantini, E. Giannoni, F. Chiti, F. Baroni, L. Formigli, J. Zurdo, N. Taddei, G. Ramponi, C. M. Dobson and M. Stefani, *Nature*, **416**, 507 (2002).
26. J. P. Cleary, D. M. Walsh, J. J. Hofmeister, G. M. Shankar, M. a. Kuskowski, D. J. Selkoe and K. H. Ashe, *Nat. Neurosci.*, **8**, 79 (2005).
27. A. A. Reinke and J. E. Gestwicki, *Chemical Biology & Drug Design*, **70**, 206 (2007).
28. M. R. Nilsson, *Methods*, **34**, 151 (2004).
29. M. Groenning, *J. Chem. Biol.*, 1 (2009).
30. I. Maezawa, H.-S. Hong, R. Liu, C.-Y. Wu, R. H. Cheng, M.-P. Kung, H. F. Kung, K. S. Lam, S. Oddo, F. M. Laferla and L.-W. Jin, *J. Neurochem.*, **104**, 457 (2008).
31. S. A. Hudson, H. Ecroyd, T. W. Kee and J. A. Carver, *The FEBS Journal*, **276**, 5960 (2009).
32. A. A. Reinke and J. E. Gestwicki, *Chemical Biology & Drug Design*, **77**, 399 (2011).
33. M. Biancalana and S. Koide, *Biochim. Biophys. Acta*, **1804**, 1405 (2010).
34. D. M. Hatters and M. D. W. Griffin, *Amyloid International Journal of Experimental and Clinical Investigation*, **752**, 121 (2011).
35. M. Lindgren and P. Hammarström, *The FEBS Journal*, **277**, 1380 (2010).
36. H. E. Wong and I. Kwon, *Plos One*, **6**, e25752 (2011).
37. A. R. A. Ladiwala, J. S. Dordick and P. M. Tessier, *J. Biol. Chem.*, **286**, 3209 (2011).
38. G Bhak, J.-H. Lee, J.-S. Hahn and S. R. Paik, *Plos One*, **4**, e4177 (2009).
39. J.-H. Lee, G Bhak, S.-G Lee and S. R. Paik, *Biophys. J.*, **95**, L16 (2008).
40. J. Ryu, M. Kanapathipillai, G Lentzen and C. B. Park, *Peptides*, **29**, 578 (2008).
41. J. W. Wu, L. Breydo, J. M. Isas, J. Lee, Y. G Kuznetsov, R. Langen and C. Glabe, *J. Biol. Chem.*, **285**, 6071 (2010).

42. M. Ahmed, J. Davis, D. Aucoin, T. Sato, S. Ahuja, S. Aimoto, J. I. Elliott, W. E. Van Nostrand and S. O. Smith, *Nat. Struct. Mol. Biol.*, **17**, 561 (2010).
43. A. R. A. Ladiwala, J. C. Lin, S. S. Bale, A. M. Marcelino-Cruz, M. Bhattacharya, J. S. Dordick and P. M. Tessier, *J. Biol. Chem.*, **285**, 24228 (2010).
44. W. E. Klunk, R. F. Jacob and R. P. Mason, *Methods Enzymol.*, **309**, 6879 (1999).
45. D. M. Walsh, D. M. Hartley, Y. Kusumoto, Y. Fezoui, M. M. Condron, A. Lomakin, G. B. Benedek, D. J. Selkoe, D. B. Teplow and D. B. J. Biol, *J. Biol. Chem.*, **274**, 25945 (1999).
46. A. Hawe, M. Sutter and W. Jiskoot, *Pharmaceutical Research*, **25**, 1487 (2008).
47. C. G. Glabe, *The Open Biology Journal*, **2**, 222 (2009).
48. A. R. A. Ladiwala, J. M. Perchiacca, Z. S. Fishman, M. Bhattacharya, A. M. Hickey, B. G. Domigan, J. S. Dordick, and P. M. Tessier, *Biotechnology and Bioengineering*, **109**, 1869 (2012).
49. S. Campioni, B. Mannini, M. Zampagni, A. Pensalfini, C. Parrini, E. Evangelisti, A. Relini, M. Stefani, C. M. Dobson, C. Cecchi and F. Chiti, *Nat. Chem. Biol.*, **6**, 140 (2010).
50. B. Bolognesi, J. R. Kumita, T. P. Barros, E. K. Esbjorner, L. M. Luheshi, D. C. Crowther, M. R. Wilson, C. M. Dobson, G. Favrin and J. J. Yerbury, *ACS Chem. Biol.*, **5**, 735 (2010).
51. R. Kayed, E. Head, F. Sarsoza, T. Saing, C. W. Cotman, M. Necula, L. Margol, J. Wu, L. Breydo, J. L. Thompson, S. Rasool, T. Gurlo, P. Butler and C. G. Glabe, *Mol. Neurodegener.*, **2**, 18 (2007).
52. R. Kayed, A. Pensalfini, L. Margol, Y. Sokolov, F. Sarsoza, E. Head, J. Hall and C. Glabe *J. Biol. Chem.*, **284**, 4230 (2009).
53. H. LeVine, III and L. C. Walker, *Neurobiology of Aging*, **31**, 542 (2010).
54. M. Fandrich, *J. Mol. Biol.*, http://dx.doi.org/10.1016/j.jmb.2012.01.006 (2012).
55. C. C. Glabe, *Subcellular Biochemistry*, **38**, 167 (2005).
56. H. E. Wong, W. Qi, H.-M. Choi, E. J. Fernandez and I. Kwon, *ACS Chem. Neurosci.*, **2**, 645 (2011).
57. E. Cerf, R. Sarroukh, S. Tamamizu-Kato, L. Breydo, S. Derclaye, Y. F. Dufrêne, V. Narayanaswami, E. Goormaghtigh, J.-M. Ruysschaert and V. Raussens, *Biochem. J.*, **421**, 415 (2009).
58. R. Sarroukh, E. Cerf, S. Derclaye, Y. F. Dufrêne, E. Goormaghtigh, J.-M. Ruysschaert and V. Raussens, *Cellular and Molecular Life Sciences: CMLS*, **68**, 1429 (2011).
59. Y. Feng, X.-P. Wang, S.-G. Yang, Y.-J. Wang, X. Zhang, X.-T. Du, X.-X. Sun, M. Zhao, L. Huang and R.-t. Liu, *Neurotoxicology*, **30**, 986 (2009).
60. L. A. Woods, G. W. Platt, A. L. Hellewell, E. W. Hewitt, S. W. Homans, A. E. Ashcroft and S. E. Radford, *Nat. Chem. Biol.*, **7**, 730 (2011).
61. M. Necula, R. Kayed, S. Milton and C. G. Glabe, *J. Biol. Chem.*, **282**, 10311 (2007).
62. X.-p. Wang, J.-h. Zhang, Y.-j. Wang, Y. Feng, X. Zhang, X.-x. Sun, J.-l Li, X.-t. Du, M. P. Lambert, S.-g. Yang, M. Zhao, W. L. Klein and R.-t. Liu, *FEBS Lett.*, **583**, 579 (2009).
63. R. Kayed, I. Canto, L. Breydo, S. Rasool, T. Lukacsovich, J. Wu, R. Albay, A. Pensalfini, S. Yeung, E. Head, J. L. Marsh and C. Glabe, *Mol. Neurodegener.*, **5**, 57 (2010).
64. A. Sood, M. Abid, C. Sauer, S. Hailemichael, M. Foster, B. Török and M. Török, *Bioorg. Med. Chem. Lett.*, **21**, 2044 (2011).
65. D. Lee, E.-K. Lee, J.-H. Lee, C.-S. Chang and S. R. Paik, *Eur. J. Biochem.*, **268**, 295 (2001).
66. C. Ha and C. B. Park, *Biotechnol. Bioeng.*, **90**, 848 (2005).
67. S. H. Ku and C. B. Park, *Langmuir*, **24**, 13822 (2008).
68. R. Friedman, *Biochem. J.*, **438**, 415 (2011).
69. K. A. Dasilva, J. E. Shaw and J. McLaurin, *Exp. Neurol.*, **223**, 311 (2010).
70. S. L. Gras, L. J. Waddington and K. N. Goldie, *Notes*, **752**, 197 (2011).
71. B. H. Toyama and J. S. Weissman, *Annu. Rev. Biochem.*, **80**, 557 (2011).
72. A. E. Langkilde and B. Vestergaard, *FEBS Lett.*, **583**, 2600 (2009).
73. S. Chimon, M. A. Shaibat, C. R. Jones, D. C. Calero, B. Aizezi and Y. Ishii, *Nat. Struct. Mol. Biol.*, **14**, 1157 (2007).
74. W. B. Stine, K. N. Dahlgren, G. a. Krafft and M. J. LaDu, *The Journal of Biological Chemistry*, **278**, 11612 (2003).
75. A. Perczel, K. Park and G. D. Fasman, *Proteins: Structure, Function, and Bioinformatics*, **13**, 57 (1992).
76. L. Whitmore and B. A. Wallace, *Biopolymers*, **89**, 392 (2008).
77. M. Bartolini, C. Bertucci, M. L. Bolognesi, A. Cavalli, C. Melchiorre and V. Andrisano, *Chembiochem: A European Journal of Chemical Biology*, **8**, 2152 (2007).
78. N. J. Greenfield, *Trends in Analytical Chemistry*, **18**, 236 (1999).
79. J. T. Pelton and L. R. McLean, *Anal. Biochem.*, **277**, 167 (2000).
80. W. K. Surewicz, J. H. H. Mantsch and D. Chapman, *Biochemistry*, **32** (1993).
81. D. H. A. Corrêa and C. H. I. Ramos, *J. Biochem.*, **3**, 164 (2009).
82. C. Soto and E. M. Castan, *J. Biochem.*, **314**, 701 (1996).
83. E. P. Vieira, H. Hermel and H. Möhwald, *Biochim. Biophys. Acta*, **1645**, 6 (2003).
84. S. Tomaselli, V. Esposito, P. Vangone, N. a. J. van Nuland, A. M. J. J. Bonvin, R. Guerrini, T. Tancredi, P. a. Temussi and D. Picone, *Chembiochem: A European Journal of Chemical Biology*, **7**, 257 (2006).
85. N. J. Greenfield, *Nature Protocols*, **1**, 2876 (2006).
86. W. Kim and M. H. Hecht, *J. Biol. Chem.*, **280**, 35069 (2005).
87. W. Kim and M. H. Hecht, *J. Mol. Biol.*, **377**, 565 (2008).
88. P. E. Arslan and A. Chakrabartty, *Biochem. Cell Biol.*, **87**, 631 (2009).
89. P. E. Arslan, V. K. Mulligan, S. Ho and A. Chakrabartty, *J. Mol. Biol.*, **396**, 1284 (2010).
90. C. Heddle and S. L. Mazaleyrat, *Protein. Eng. Des. Sel.*, **20**, 327 (2007).
91. S. Gregoire and I. Kwon, *ACS Meeting*, **241**, 234-BIOT (2011).
92. W. Kim, Y. Kim, J. Min, D. J. Kim, Y. T. Chang and M. H. Hecht, *ACS Chem. Biol.*, **1**, 461 (2006).
93. J. Geng, M. Li, J. Ren, E. Wang and X. Qu, *Angew. Chem. Int. Ed. Eng.*, **50**, 4184 (2011).
94. B. Philipps, J. Hennecke and R. Glockshuber, *J. Mol. Biol.*, **327**, 239 (2003).
95. S. Cabantous, T. C. Terwilliger and G. S. Waldo, *Nat. Biotechnol.*, **23**, 102 (2005).
96. S. Cabantous, J. D. Pédelacq, B. L. Mark, C. Naranjo, T. C. Terwilliger and G. S. Waldo, *J. Struct. Funct. Genomics.*, **6**, 113 (2005).
97. S. Cabantous and G. S. Waldo, *Nat. Methods*, **3**, 845 (2006).
98. W. J. Chun, G. S. Waldo and G. V. W. Johnson, *J. Neurochem.*, **103**, 2529 (2007).
99. P. Listwan, T. C. Terwilliger and G. S. Waldo, *J. Struct. Funct. Genomics.*, **10**, 47 (2009).

100. M. A. Lockard, P. Listwan, J. D. Pedelacq, S. Cabantous, H. B. Nguyen, T. C. Terwilliger and G. S. Waldo, *Protein Engineering Design & Selection*, **24**, 565 (2011).
101. K. L. Maxwell, A. K. Mittermaier, J. D. Forman-Kay and A. R. Davidson, *Protein Sci.*, **8**, 1908 (1999).
102. M. R. Dyson, R. L. Perera, S. P. Shadbolt, L. Biderman, K. Bromek, N. V. Murzina and J. McCafferty, *Nucleic Acids Res.*, **36**, e51. Epub 2008 Apr 17. (2008).
103. M. R. Dyson, *Biochem. Soc. Trans.*, **38**, 908 (2010).
104. M. Morell, N. S. de Groot, J. Vendrell, F. X. Aviles and S. Ventura, *Mol. Biosyst.*, **7**, 1121 (2011).
105. W. C. Wigley, R. D. Stidham, N. M. Smith, J. F. Hunt and P. J. Thomas, *Nat. Biotechnol.*, **19**, 131 (2001).
106. L. Foit, G. J. Morgan, M. J. Kern, L. R. Steimer, A. A. von Hacht, J. Titchmarsh, S. L. Warriner, S. E. Radford and J. C. Bardwell, *Mol. Cell.*, **36**, 861 (2009).
107. T. J. Mansell, S. W. Linderman, A. C. Fisher and M. P. DeLisa, *Protein Science*, **19**, 1079 (2010).
108. T. J. Mansell, A. C. Fisher and M. P. DeLisa, *Curr. Protein Peptide Sci.*, **9**, 138 (2008).
109. H. Yumerefendi, F. Tarendeau, P. J. Mas and D. J. Hart, *J Struct Biol.*, **172**, 66 (2010).
110. Y. An, H. Yumerefendi, P. J. Mas, A. Chesneau and D. J. Hart, *J. Struct. Biol.*, **175**, 189 (2011).
111. B. Nölting, Methods in modern biophysics: Ch. 5 - protein infrared spectroscopy, Springer Berlin Heidelberg, 91 (2010).
112. R. L. Wiseman and J. W. Kelly, *Mol. Cell.*, **36**, 730 (2009).

Inchan Kwon is Assistant Professor of the Chemical Engineering at the University of Virginia. Professor Kwon earned his BS and MS in Chemical Engineering at Seoul National University, South Korea (BS and MS). He received his PhD in Chemical Engineering from California Institute of Technology (Caltech) and completed his postdoctoral training in the Bioengineering at the University of California at Berkley. Professor Kwon focuses on the engineering proteins using nonnatural amino acid incorporation and the modulation of protein aggregation associated with neurodegenerative diseases including Alzheimer's disease and Lou Gehrig's disease. Professor Kwon has been recognized in numerous ways including Constantin G. Economou Memorial Award at Caltech, KSEA Young Investigator Grant, and James M. Lee Young Investigator Award.