Covalent Structural Changes in Unfolded GroES That Lead to Amyloid Fibril Formation Detected by NMR

INSIGHT INTO INTRINSICALLY DISORDERED PROTEINS

Received for publication, February 4, 2011, and in revised form, April 15, 2011 Published, JBC Papers in Press, April 20, 2011, DOI 10.1074/jbc.M111.228445

Hisanori Iwasa, Shunsuke Meshitsuka, Kunihiro Hongo, Tomohiro Mizobata, and Yasushi Kawata

From the Department of Chemistry and Biotechnology, Graduate School of Engineering, and the Department of Biomedical Science, Institute of Regenerative Medicine and Biofunction, Graduate School of Medical Science, Tottori University, Tottori 680-8552, Japan

Co-chaperonin GroES from Escherichia coli works with chaperonin GroEL to mediate the folding reactions of various proteins. However, under specific conditions, i.e. the completely disordered state in guanidine hydrochloride, this molecular chaperone forms amyloid fibrils similar to those observed in various neurodegenerative diseases. Thus, this is a good model system to understand the amyloid fibril formation mechanism of intrinsically disordered proteins. Here, we identified a critical intermediate of GroES in the early stages of this fibril formation. Mutation of these asparagines to alanines resulted in delayed fibril formation. The finding provides a novel insight into the structural process of amyloid fibril formation from a disordered state, which may be applicable to intrinsically disordered proteins in general.

Intrinsically disordered proteins are commonly defined as proteins that do not adopt a well defined structure in solution (1), and they can fold into ordered structures only upon binding to their cellular targets (2). They are abundant in eukaryotic proteins and play a significant role in biological functions associated with signaling and regulation events (3). Some intrinsically disordered proteins, exemplified by α-synuclein (4–6) and poly(Q) (7–9) proteins, are capable of forming insoluble aggregates referred to as amyloid fibrils. Understanding the detailed mechanisms in which intrinsically disordered proteins self-assemble into amyloid fibrils is a very important issue because they associate closely with amyloid-related degenerative diseases (10).

The prevalent model to explain the mechanisms of amyloid fibril formation in vitro is nucleation-dependent fibril formation (11). Fibril formation kinetics consists of two phases, that is, nucleation and extension, as traced with Thioflavin-T-binding fluorescence or turbidity of incubated samples. Nucleus formation requires a series of associations between protein monomers that are thermodynamically unfavorable, and this step represents the rate-limiting step in amyloid fibril formation. Once the nucleus has been formed, the further addition of monomers becomes thermodynamically favorable, resulting in a rapid extension of amyloid fibrils (12, 13). It has generally been assumed until recently that the cytotoxicity in amyloid-related degenerative diseases was due to mature amyloid fibrils, but now attention has shifted to various intermediate species formed during nucleation, such as an oligomeric but soluble state (14–16) or even some monomeric states (17, 18). Therefore, details regarding the structural characteristics of these species that appear in the early stage of the fibril formation are very important.

Here, we set out to investigate the structural changes that occur during the amyloid fibril formation of molecular chaperone GroES, a member of hsp10 from Escherichia coli. Native GroES forms heptameric oligomers of subunits abundant in β-strands (Fig. 1A) and acts as the co-chaperonin of GroEL that mediates various protein folding reactions in vivo and in vitro. We have previously found that GroES may form typical amyloid fibrils from the guanidine hydrochloride (Gdn-HCl)2 unfolded state (19) and determined the core sequence of the mature GroES amyloid fibrils that are resistant to protease (Fig. 1B) (20). Our findings proved that even a molecular chaperone may form amyloid fibrils under certain conditions, supporting the recent finding that amyloid fibrils are an inherently common structure of many proteins under certain conditions in vitro (21–23). As typical amyloid fibrils of GroES are formed from an extensively disordered (unfolded) state (19), this system is a good model for elucidating the common mechanism of amyloid fibril formation of intrinsically disordered proteins.

In this study, we found very early, initial conformational changes in unfolded GroES by using solution NMR spectroscopy, a powerful approach to obtain insights into soluble species formed during fibril formation at an atomic level. We first

*This work was supported in part by grants-in-aid for Scientific Research from the Ministry of Education, Culture, Sports, Science, and Technology of Japan, the Global COE Program, and a grant from the Organization for Regional Industrial Academic Cooperation of Tottori University. Mass spectrometry measurements were performed using the facilities of the Research Center for Bioscience and Technology (Venture Business Laboratory) of Tottori University. The on-line version of this article (available at http://www.jbc.org) contains supplemental Fig. S1.

‡ To whom correspondence should be addressed: Dept. of Chemistry and Biotechnology, Graduate School of Engineering, Tottori University, Koyama-Minami, Tottori 680-8552, Japan. Tel./Fax: 81-857-31-5271; E-mail: kawata@bio.tottori-u.ac.jp.

2 The abbreviations used are: Gdn-HCl, guanidine hydrochloride; HSQC, heteronuclear single quantum coherence; SSP, secondary structure propensity; TOCSY, total correlation spectroscopy.
Asn-Gly Transition Leads to Fibrillation of Unfolded GroES

The successful construction of each mutant was confirmed by DNA sequence analysis of the entire GroES coding region. Both wild-type and mutant proteins were expressed in E. coli BL21(DE3) (Novagen) and purified as described previously (24). Chlorella medium (Chlorella Industry) uniformly labeled with stable isotope (15N or 15N and 13C) was used for the expression of wild-type GroES, and M9 minimal medium supplemented with [15N]NH4Cl (Shoko) was used for the expression of N51A and N45A/N51A for NMR measurements. The cells were resuspended in 10-fold ice-cold buffer A (50 mM Tris-HCl (pH 8.0), 1 mM EDTA, 1 mM DTT, 0.1 mM PMSF) and lysed by sonication on ice. The crude extracts were cleared by centrifugation and nucleic acids were removed by the addition of 2% streptomycin. After addition of 55% ammonium sulfate and centrifugation, precipitates were resuspended in buffer A and heated at 80 °C for 20 min. After the heat treatment, the mixture was quickly cooled in ice-cold water for 20 min. To remove heat-denatured proteins, the heated fraction was centrifuged at 13,000 × g for 40 min at 4 °C. The supernatant was applied to a Q-Sepharose anion-exchange column (ϕ2.7 cm × 18 cm) equilibrated in buffer B (50 mM Tris-HCl (pH 8.0), 1 mM EDTA, 1 mM DTT). Proteins were eluted by a 0–0.5 M NaCl gradient (total 1000 ml), and eluted GroES fractions were dialyzed against Milli-Q water. The GroES thus obtained was filtered through a 0.22-μm cellulose-acetate membrane and stocked at 4 °C after lyophilization. The purity of GroES was checked by SDS-PAGE. The concentration of GroES protein was determined by using either an extinction coefficient at 280 nm, ε1% = 2.7 cm mol⁻¹, or a protein dye reagent (Protein Assay kit; Bio-Rad Laboratories) using bovine serum albumin (Sigma) as a standard.

NMR Measurement—Lyophilized GroES that are labeled with single 15N or double 15N-13C stable isotopes were dissolved in 10 mM sodium phosphate buffer (pH 6.5), 1.6 mM Gdn-HCl, 90% (v/v) H2O/10% D2O, and NMR measurements were performed at 1 mM (10 mg/ml) GroES concentration and 25 °C. Incubation for fibril formation was done in a Shigemi tube (Shigemi) without agitation at 25 °C. Samples were recorded using a Varian Unity Inova 500 spectrometer operating at a 1H resonance frequency of 500 MHz. All NMR data were processed with NMRPipe (26) and analyzed in NMRView (27) and Sparky (28). Proton chemical shifts were referenced to 4,4-dimethyl-4-silapentane-1-sulfonate as 0.00 ppm. The assignments of the resonances in 1H-15N HSQC spectrum of GroES WT were carried out by using 15N-edited TOCSY-HSQC, CBCANH, CBCA(CO)NH, HN(CA)CO, and HNCO. The 1H-15N HSQC peaks of Gly36 and Gly52 of the intermediate species that appeared after a 28-day incubation in Gdn-HCl were assigned by performing time-lapse measurements of 15N-labeled GroES N51A and N45A/N51A mutants. The secondary structure propensity (SSP) program has been developed by Marsh et al. and gives a measure of the secondary structure populated (29). The SSP score of GroES was calculated using the Cα and Cβ chemical shifts of GroES sample against those of random coil. The relative peak intensity between spectra is given as δI = I/Io, where I and Io represent the peak intensities at 28 days and 2 days, respectively.

EXPERIMENTAL PROCEDURES

Preparation of Wild-type and Mutant GroES Proteins—Genes encoding the GroES mutants, N45A, N51A, and N45A/N51A, were constructed by using the QuikChange site-directed mutagenesis kit (Stratagene) with pETES (24) as a template.

assigned the 1H-15N resonances of monomeric unfolded GroES in Gdn-HCl and next detected structural changes within a specific region during fibril formation. Interestingly, this region was adjacent to the fibril core sequence that we determined previously (20). It was also found that formation of this intermediate hinged on an Asn-Gly rearrangement reaction, which yielded β-aspartic acid as a consequence of deamidation of the asparagine. The relevance of this reaction was further examined by site-directed mutagenesis where the asparagine residue in the Asn-Gly sequence was substituted with an alanine, which showed that the period required for this reaction to occur, our findings the formation of GroES amyloid fibril. Considering the minimal criterion that is required for this reaction to occur, our findings

The successful construction of each mutant was confirmed by DNA sequence analysis of the entire GroES coding region. Both wild-type and mutant proteins were expressed in E. coli BL21(DE3) (Novagen) and purified as described previously (24). Chlorella medium (Chlorella Industry) uniformly labeled with stable isotope (15N or 15N and 13C) was used for the expression of wild-type GroES, and M9 minimal medium supplemented with [15N]NH4Cl (Shoko) was used for the expression of N51A and N45A/N51A for NMR measurements. The cells were resuspended in 10-fold ice-cold buffer A (50 mM Tris-HCl (pH 8.0), 1 mM EDTA, 1 mM DTT, 0.1 mM PMSF) and lysed by sonication on ice. The crude extracts were cleared by centrifugation and nucleic acids were removed by the addition of 2% streptomycin. After addition of 55% ammonium sulfate and centrifugation, precipitates were resuspended in buffer A and heated at 80 °C for 20 min. After the heat treatment, the mixture was quickly cooled in ice-cold water for 20 min. To remove heat-denatured proteins, the heated fraction was centrifuged at 13,000 × g for 40 min at 4 °C. The supernatant was applied to a Q-Sepharose anion-exchange column (ϕ2.7 cm × 18 cm) equilibrated in buffer B (50 mM Tris-HCl (pH 8.0), 1 mM EDTA, 1 mM DTT). Proteins were eluted by a 0–0.5 M NaCl gradient (total 1000 ml), and eluted GroES fractions were dialyzed against Milli-Q water. The GroES thus obtained was filtered through a 0.22-μm cellulose-acetate membrane and stocked at 4 °C after lyophilization. The purity of GroES was checked by SDS-PAGE. The concentration of GroES protein was determined by using either an extinction coefficient at 280 nm, ε1% = 2.7 cm mol⁻¹, or a protein dye reagent (Protein Assay kit; Bio-Rad Laboratories) using bovine serum albumin (Sigma) as a standard.

NMR Measurement—Lyophilized GroES that are labeled with single 15N or double 15N-13C stable isotopes were dissolved in 10 mM sodium phosphate buffer (pH 6.5), 1.6 mM Gdn-HCl, 90% (v/v) H2O/10% D2O, and NMR measurements were performed at 1 mM (10 mg/ml) GroES concentration and 25 °C. Incubation for fibril formation was done in a Shigemi tube (Shigemi) without agitation at 25 °C. Samples were recorded using a Varian Unity Inova 500 spectrometer operating at a 1H resonance frequency of 500 MHz. All NMR data were processed with NMRPipe (26) and analyzed in NMRView (27) and Sparky (28). Proton chemical shifts were referenced to 4,4-dimethyl-4-silapentane-1-sulfonate as 0.00 ppm. The assignments of the resonances in 1H-15N HSQC spectrum of GroES WT were carried out by using 15N-edited TOCSY-HSQC, CBCANH, CBCA(CO)NH, HN(CA)CO, and HNCO. The 1H-15N HSQC peaks of Gly36 and Gly52 of the intermediate species that appeared after a 28-day incubation in Gdn-HCl were assigned by performing time-lapse measurements of 15N-labeled GroES N51A and N45A/N51A mutants. The secondary structure propensity (SSP) program has been developed by Marsh et al. and gives a measure of the secondary structure populated (29). The SSP score of GroES was calculated using the Cα and Cβ chemical shifts of GroES sample against those of random coil. The relative peak intensity between spectra is given as δI = I/Io, where I and Io represent the peak intensities at 28 days and 2 days, respectively.

EXPERIMENTAL PROCEDURES

Preparation of Wild-type and Mutant GroES Proteins—Genes encoding the GroES mutants, N45A, N51A, and N45A/N51A, were constructed by using the QuikChange site-directed mutagenesis kit (Stratagene) with pETES (24) as a template.

assigned the 1H-15N resonances of monomeric unfolded GroES in Gdn-HCl and next detected structural changes within a specific region during fibril formation. Interestingly, this region was adjacent to the fibril core sequence that we determined previously (20). It was also found that formation of this intermediate hinged on an Asn-Gly rearrangement reaction, which yielded β-aspartic acid as a consequence of deamidation of the asparagine. The relevance of this reaction was further examined by site-directed mutagenesis where the asparagine residue in the Asn-Gly sequence was substituted with an alanine, which showed that the period required for nucleus formation was prolonged. These results indicated that the contiguous Asn-Gly sequence before the fibril core region leads to the formation of GroES amyloid fibril. Considering the minimal criterion that is required for this reaction to occur, our findings provide insight into the mechanism of fibril formation related to amyloid-related degenerative diseases as well as the structural characteristics and aggregation propensities of the intrinsically disordered proteins in general.
Asn-Gly Transition Leads to Fibrillation of Unfolded GroES

In $^{15}$N backbone relaxation experiments, $^1$H-$^{15}$N NOE, longitudinal relaxation ($R_1$), and transverse relaxation ($R_2$) measurements of GroES were recorded. $^1$H-$^{15}$N NOE values were measured by recording spectra with or without a $^1$H saturation period of 3 s. The $R_2$ experiments were collected using the following relaxation delay times: 0.01, 0.05, 0.17, 0.49, and 1.80 s. The $R_1$ experiments were collected using the following relaxation delay times: 0.01, 0.03, 0.05, 0.09, and 0.21 s. The relaxation rates and error estimate were calculated by the best fit time constant $T$ (rate constant $R = 1/T$) for well resolved peaks in Sparky using the fitting function (supplemental Fig. S1).

Amyloid Fibril Formation and Thioflavin-T Binding Assay—Experiments of amyloid fibril formation were performed with 1 mg/ml GroES dissolved in 10 mM sodium phosphate (pH 7.4) containing 1.6 M Gdn-HCl with linear agitation (90 min$^{-1}$) at 37 °C in glass test tubes. Fluorescence of Thioflavin-T was measured using a Jasco FP-6300 spectrofluorometer at 25 °C. At appropriate times, aliquots of GroES samples were withdrawn and mixed thoroughly with a staining solution of 25 μM Thioflavin-T, 5 mM sodium phosphate (pH 7.4), 150 mM NaCl (final GroES: 7.5 μg/ml). Fluorescence intensities were monitored at 480 nm with excitation at 440 nm.

Separation of Intermediate Species—Aliquots of samples during fibril formation were withdrawn and ultracentrifuged (150,000 × g, 1 h, 4 °C), and the supernatants were subjected to a desalting PD Spin Trap G-25 (GE Healthcare) to remove Gdn-HCl. Mature amyloid fibrils obtained as a precipitate after ultracentrifugation of samples were denatured and dissolved in 7.5 M Gdn-HCl for 24 h, then Gdn-HCl was removed by a PD Spin Trap G-25 column. These samples were subjected to 10% native-PAGE, and protein bands were stained with Coomassie Brilliant Blue R-250.

Protease Digestion and MALDI-TOF Mass Spectroscopy—In-gel digestion was performed by the method described by Shevchenko et al. (30). Gel slices containing about 5 μg of sample (GroES in various intermediate forms) were destained and dried in a vacuum centrifuge and then digested by adding 10 ng/μl lysyl endopeptidase (enzyme:substrate = 1:17–50 in molar ratio) in 25 mM NH$_4$HCO$_3$ at 37 °C for 16 h. To extract the digests, extraction buffer (50% acetonitrile/5% TFA) was added to the mixture, and the supernatant was collected, followed by further drying of gel pieces by addition of 100% acetonitrile, and the extract was dried completely in a vacuum centrifuge. The dried samples were dissolved in 33% acetonitrile/0.07% TFA and mixed with an equal amount of the matrix solution, 33% acetonitrile/0.07% TFA saturated with α-cyano-4-hydroxy-cinnamic acid. Resultant samples were spotted onto a target plate (MTP 384, Bruker Daltonics) and dried. Measurements were performed on an Autoflex (Bruker Daltonics) in reflection mode with positive ion detection. The mass spectra were calibrated by Peptide Calibration Standard (1000 – 4000 Da, Bruker Daltonics).

RESULTS

Backbone Assignments of GroES in a Disordered State—As reported previously, we found that GroES formed typical amyloid fibrils from a disordered state formed in 1.6 M Gdn-HCl, where GroES heptamer totally unfolded to a monomeric state (19, 24). To investigate further mechanisms of GroES amyloid fibril formation at atomic resolution, we prepared singly ($^{15}$N) and doubly ($^{15}$N-$^{13}$C)-labeled GroES and performed a series of solution NMR experiments. As shown in Fig. 2A, the $^1$H-$^{15}$N resonances of GroES in HSQC measurements were typical of those of unfolded proteins, showing a limited resonance dispersion of ~1 ppm in the $^1$H dimension, but retaining relatively good dispersion covering ~24 ppm in the $^{13}$NH dimension. The assignment of the backbone resonances of GroES to their respective locations within its sequence was carried out by triple-resonance measurements including CBCANH and CBCA-(CO)NH, as well as HN(CA)CO and HNCO to overcome ambiguities that arise in the standard experiments. The dispersions in $\alpha$ and $\beta$ chemical shifts were poor in the unfolded state, whereas the dispersions of $^{13}$CO resonances were much greater, as in the case of $^{15}$NH chemical shifts, reflecting the sensitivity of these nuclei to the nature of the neighboring amino acid in the primary sequence (31). Although Asn$^2$ and Val$^{19}$ were not identified in the HSQC spectrum due to the narrow chemical shift dispersion, in total, 92 resonances out of 97 residues (excluding the N-terminal Met, Pro$^3$, and Pro$^{56}$) in the sequence have been assigned.

Chemical shifts of $\alpha$ and $\beta$ are often used to calculate the secondary structure propensities of unfolded proteins. Although GroES is totally unfolded in 1.6 M Gdn-HCl, the facts that (i) GroES is still capable of forming amyloid fibrils and that (ii) native GroES heptamer is rich in β-strands, promoted us to examine whether residual β-structure remained in some regions of the sequence, through which GroES may be prone to aggregate. In this study, we used the SSP score (29) to calculate the propensity of each amino acid residue of GroES to assume a certain secondary structure. In Fig. 2B, although a very slight propensity was observed in certain sequence segments (8.3% of β-structure in average for residues 1–80 and 8.2% of α-structure in average for residues 81–97), the overall SSP profile showed that at atomic resolution the conformation of GroES was close to a random coil, in agreement with previous small angle x-ray scattering measurements (24).

Detection of Structural Changes from the Disordered State—To determine the existence of intermediate species during GroES amyloid fibril formation, we performed time-lapse $^1$H-$^{15}$N HSQC measurements of unfolded GroES at 25 °C. Very interestingly, several noticeable changes were observed (Fig. 3). Overlay of sample spectra measured at 2 days and 28 days revealed a significant decrease in the peaks corresponding to Gly$^{46}$ and Gly$^{52}$. In addition, the peak intensities of the adjacent Asn$^{45}$ and Asn$^{51}$ residues were also decreased greatly. Correspondingly, two new peaks were detected in the 28-days spectra (Fig. 3A). Identification of these two new peaks was performed by three-dimensional TOCSY and time-lapse measurement analysis for wild-type GroES, GroES N51A, and N45A/N51A mutants (see below), and it was shown that they corresponded...
to Gly\textsuperscript{46} and Gly\textsuperscript{52}. Slight chemical shift perturbations were also observed for Gly\textsuperscript{44}, Glu\textsuperscript{50}, Lys\textsuperscript{55}, and Leu\textsuperscript{57}. Samples incubated for more than 28 days showed very low peak intensities due to amyloid fibril formation (data not shown). We analyzed the relative peak intensities of each peak using the data of Fig. 3A (Fig. 3B). Although the peak intensities of all of the amino acid residues in general gradually decreased to ~76% at 28 days, the peak intensities within the region Val\textsuperscript{43}-Leu\textsuperscript{57} decreased remarkably (to 30–50%). Because after these changes formation of amyloid fibrils was confirmed by Thioflavin-T binding.
Asn-Gly Transition Leads to Fibrillation of Unfolded GroES

assay in separate experiments, these data strongly suggested the appearance of soluble intermediate species formed during nucleus formation. These residues were adjacent to the fibril core region (Asp58-Lys74) previously identified (20) in GroES amyloid fibrils.

We next performed $^{15}$N relaxation experiments to explore the dynamics of the Gdn-HCl denatured and the intermediate species. NOE values are very sensitive to motions in the picosecond to nanosecond time scales and microsecond to millisecond time scales, respectively (32–34). As seen in the left panels of Fig. 3, C–E, each amino acid residue displayed different values in these measurements, but no significant changes of the overall dynamics between the initial form and the intermediate species ensemble were seen. The average values of $^{1}$H-$^{15}$N NOE, $R_{1}$, and $R_{2}$ of the unfolded species at 2 days were $−0.8$, $2.2$ s$^{-1}$, and $4.5$ s$^{-1}$, and those of the intermediate species ensemble at 28 days were $−0.8$, $2.1$ s$^{-1}$, and $4.2$ s$^{-1}$, respectively, and the lack of large differences in these values indicated that the soluble intermediate species were in a monomeric state. Also, the NOE values were negative across the entire primary sequence, demonstrating a highly dynamic motion throughout the sequence in the presence of Gdn-HCl. This character was especially pronounced in the mobile loop region (around position 22). A highly dynamic motion in the C terminus was also observed in the $R_{1}$ and $R_{2}$ measurements.

When we looked at the specific residues of Gly46 and Gly52 in the intermediate species (peaks in the dotted square in Fig. 3A), it was revealed that all values of NOE, $R_{1}$, and $R_{2}$ of the intermediate were slightly but significantly smaller than those of the disordered structure (the right panels of Fig. 3, C–E). This finding suggested that flexibilities of Gly46 and Gly52 in the intermediate species were increased compared with the initial state.

Separation and Mass Analysis of the Intermediate: Implication of Covalent Structural Changes in Asn-Gly Sequence—It is evident that the intermediate species observed during the nucleus formation process were formed from a Gdn-HCl disordered state, as described above. This result raised an interesting question: What causes the structural changes in the region of Val13-Leu57 adjacent to the fibril core sequence? To address the question, a series of biochemical experiments were performed. GroES is capable of refolding reversibly from 1.6 m Gdn-HCl as described previously (35), which is detectable by native-PAGE (band I in Fig. 4A). In GroES samples incubated for a prolonged interval in Gdn-HCl (7–35 days), however, new species (bands II, III, and IV in Fig. 4A) were detected that could not properly refold to heptamer. The time scale of the formation of these nonheptameric species also corresponded to the time required for the formation of intermediates observed in the NMR measurements (Fig. 3). When mature amyloid fibrils of GroES were dissolved by 7.5 m Gdn-HCl and then Gdn-HCl was removed to allow refolding, we found a fraction of GroES formed intermediate species, which could not form the hepta-

FIGURE 3. Changes in NMR spectra of GroES after prolonged incubation at 25 °C. A, $^{1}$H-$^{15}$N HSQC spectra after a 2-day (black) and 28-day (red) incubation in 1.6 m Gdn-HCl without agitation at 25 °C. Notable changes: peak intensities of Asn45, Gly46, Asn51, and Gly52 decreased significantly, and the peaks of Gly46 and Gly52 were detected partially at new positions (dotted square), indicating that intermediate species were formed. B, relative peak intensity of resonances at 2 days (black circles), 14 days (blue circles), and 28 days (red circles), given as $I_{d}$/$I_{p}$, where $I$ represents the peak intensities at 14 or 28 days and $I_{p}$ is equal to the intensity at 2 days. C–E, $^{15}$N relaxation parameters of GroES. The $^{1}$H-$^{15}$N NOE (C), relaxation data $R_{1}$ (D), and $R_{2}$ (E) of disordered species (black circles) and intermediate species (red triangles) are shown in each panel. The $^{15}$N relaxation parameters of Gly46 and Gly52, detected at new positions in the intermediate species, are also plotted (red columns) along with those in the disordered species (black columns). The data of the disordered and intermediate species were recorded between day 0 and day 4, and between day 21 and 25, respectively.
Asn-Gly Transition Leads to Fibrillation of Unfolded GroES

The intermediate species observed in the native-PAGE gels were further analyzed by MALDI-TOF mass spectroscopy after in-gel digestion by lysyl endopeptidase. Measurements of the mass spectrum focused on a peptide corresponding to residues 35–55 (STRGEVLAVGNGARILGENGGEVK \(M+H^+\) = 2198.18), a region that includes the site of the initial structural changes identified in Fig. 3. A peak of this peptide derived from native heptamer GroES (band I in Fig. 4A) displayed a parent ion mass of 2198.03, together with several isotopic peaks (Fig. 4B). Interestingly, for the corresponding peptide obtained from the intermediate species of bands II and III in Fig. 4A, a 1-Da increase in mass was observed, where the parent ion mass peaks were 2199.04 and 2199.17, respectively (Fig. 4C and D). For the corresponding peptide from band IV in Fig. 4A, a 2-Da increase in mass (a parent ion mass = 2200.17) was observed (Fig. 4E). Taken together with our results showing significant NMR peak changes for Asn\(^{45}\), Gly\(^{46}\), Asn\(^{51}\), and Gly\(^{52}\) (Fig. 3A), our results are consistent with a covalent Asn-Gly rearrangement reaction, which yields \(\beta\)-aspartic acid due to deamination of the asparagine. Rearrangement of the \(\alpha\)-peptide bond at Asn-Gly to a \(\beta\)-peptide bond renders the entire protein incapable of refolding to native heptamer (Fig. 4A). Thus, this structural change, that presumably occurs within the sequences Asn\(^{45}\)-Gly\(^{46}\) and Asn\(^{51}\)-Gly\(^{52}\), results in formation of intermediate species, which then trigger or accelerate fibril nucleus formation. Once the fibril nucleus forms, fibril extension reaction follows by incorporating various forms of intact GroES monomer.

**Amyloid Fibril Formation by GroES Mutants**—To confirm that the rearrangement reaction in Asn-Gly sequence triggers or accelerates GroES amyloid fibril formation, we generated single N45A and N51A and double N45A/N51A GroES mutants and performed fibril formation experiments (Fig. 5A). All of the GroES mutants were found to form amyloid fibrils eventually, most likely because we selected experimental conditions that greatly encouraged fibril formation. However, the times required for nucleus formation were prolonged in N45A/ N51A (12 h) and N51A (22 h) mutants compared with that of wild-type GroES (WT) (6 h), and the time required for the N45A (5 h) mutant was almost the same as the WT. The effect of a single N51A mutation was the greatest, demonstrating that the rearrangement at Asn\(^{51}\)-Gly\(^{52}\) plays an important role in accelerating the fibril formation of GroES. Fibril extension rates were only mildly affected by these mutations, suggesting that these Asn-Gly rearrangements mainly affect nucleus formation.

The incubated samples and the mature amyloid fibrils for each mutant were then analyzed by native-PAGE. As shown in Fig. 5B, bands ii and iv, which were observed for WT, became undetectable in the N51A and N45A mutants, respectively, according to their mutation position. No intermediate bands were observed at all for N45A/N51A. In-gel digestion followed by mass spectroscopy measurements for the mutants was also performed. From the peptides derived from native heptamer N45A and N51A mutants (bands i and iii marked in Fig. 5B), peaks of parent ion mass 2155.10 and 2155.16 were observed, respectively (theoretical mass = 2155.18) (Fig. 5, C and E). The peptides from the nonheptameric species of bands ii and iv (marked in Fig. 5B) showed a mass increase of 1 Da, where peaks of parent ion mass 2156.21 and 2156.15 were observed, respectively (Fig. 5, D and F). The results confirmed that each mutation suppressed the \(\beta\)-rearrangement reaction of the corresponding peptide bond, and as a result, prevented the accumulation of intermediate species. This demonstrates that the rearrangement reaction at Asn\(^{45}\) and/or Asn\(^{51}\) significantly affects successful refolding of GroES to native heptamer. Notably, mutation of these two Asn residues does not hamper GroES refolding in a measurable manner.

In NMR measurements of N45A/N51A mutant incubated for 28 days, no resonance peaks of Gly\(^{46}\) (Fig. 5G) and Gly\(^{52}\) (Fig. 5H) were observed. From these results, we concluded that rearrangement reactions at Asn\(^{45}\)-Gly\(^{46}\) and/or Asn\(^{51}\)-Gly\(^{52}\) occur during nucleus formation, resulting in \(\beta\)Asp\(^{45}\) and/or \(\beta\)Asp\(^{51}\) to be formed. The intermediates, especially \(\beta\)Asp\(^{51}\), affect significantly the formation of fibril nuclei, and interestingly, leave the fibril extension rates almost unchanged (Fig. 5A).
FIGURE 5. Amyloid fibril formation and the analysis of intermediates formed by various GroES mutants. A, amyloid fibril formation of GroES mutants monitored by Thioflavin-T fluorescence (1 mg/ml GroES, 10 mM sodium phosphate (pH 7.4), 1.6 M Gdn-HCl with linear agitation (90 min−1) at 37 °C). WT, circles; N45A, squares; N51A, triangles; and N45A/N51A, diamonds. B, native-PAGE analysis of GroES native heptamer, incubated GroES (0, 24 h), and GroES mutants resolubilized and refolded from mature amyloid fibrils. The mass spectra of the target peptides of Ser35-Lys55 from GroES N45A and N51A (both (M+H)+ are 2155.18), which were derived from the bands i–iv in native-PAGE gel, are indicated in C–F, respectively. The parent ion mass of each spectrum is also shown in each panel. G and H, 1H-15N HSQC spectra taken of WT and N45A/N51A incubated for 28 days compared by overlaying 1H slices corresponding to the region Gly46 (G) and Gly52 (H) highlighted by the dotted square in Fig. 3A.
DISCUSSION

Studies of amyloid fibrils have shown that amyloid fibrils are an inherently common structure in proteins and also may be formed under certain conditions in vitro by proteins that are unrelated to diseases (22). We previously found that the cochaperonin GroES of E. coli (Fig. 1) formed amyloid fibrils from the Gdn-HCl unfolded state (19), and the core sequence of GroES that forms a rigid β-structure in this fibril was identified by protease digestion (20). In addition to elucidating the details of fibril formation and fibril morphology, recent studies indicate the importance of classifying intermediate species formed during the nucleation process, due to their cytotoxicity (14–18). Therefore, the elucidation of characteristics of the nucleus-forming molecular species and the mechanism of formation are quite important and critical issues. Furthermore, as the GroES protein formed typical amyloid fibrils similar to those observed in various neurodegenerative diseases from a completely disordered state, this can be a good model system to understand the amyloid fibril formation mechanism of intrinsically disordered proteins.

Although several NMR measurements of native heptameric GroES have been studied so far (36–40), in our study, we performed 1H-15N HSQC measurements and backbone assignments of GroES in 1.6 M Gdn-HCl, where native GroES heptamer totally unfolds to monomers (19, 24). The 1H-15N resonances of GroES in 1.6 M Gdn-HCl were found to be similar to those of disordered proteins with a limited resonance dispersion of ~1 ppm in the 1H dimension, but retaining relatively good dispersion covering ~24 ppm in the 15NH dimension (Fig. 2). From our peak assignments of the unfolded GroES protein, we were able to use GroES as a good model to clarify the mechanism of amyloid fibril formation at an atomic level for the intrinsically disordered proteins.

The detailed peak assignments of GroES in the disordered state allowed us to perform time-lapse measurements of 1H-15N HSQC at the atomic level (Fig. 3, A and B). During the fibrillation process, the peak intensities of the overall residues decreased gradually to 76% after 28 days, indicating that soluble species remained. A notable difference in the plot of the relative peak intensity at 28 day was that the region of Val43-Leu57 showed a significant additional decrease in their intensities. A similar phenomenon was observed for α-synuclein during the initial stage of fibril formation (41), where many cross-peaks were significantly attenuated due to the shortened relaxation times as a consequence of polymerization. In the case of GroES, the decrease in peak intensities within the Val43-Leu57 segment was caused by chemical shift changes in the spectra, which were observed notably for Gly46 and Gly52 and slightly for Gly44, Gly50, Lys55, and Leu57, indicating that structural changes occurred in this region adjacent to the fibril core sequence Asp58-Lys74 (20). Further 15N relaxation experiments indicated that the GroES protein remained as a monomeric state during this time (Fig. 3, C–E).

We found that the NMR resonance changes observed in 1H-15N HSQC (Fig. 3) were due to specific covalent structural changes of Asn46-Gly52 and Asn51-Gly52 (Figs. 4 and 5). In general, asparagine residues with Asn-Gly sequence motifs in peptides and proteins have a tendency to undergo spontaneous deamidation reactions at neutral and alkaline pH. This results in the production of Asp and βAsp (isoAsp) residues through a succinimide intermediate (42). The typical ratio of Asp and βAsp formed after rearrangement is about 1:3, favoring the βAsp form (although the actual ratio is dependent on peptide or protein) (43, 44). These nonenzymatic reactions occur primarily at asparagine and aspartate residues. This deamidation reaction results in a 1-Da increase in the molecular mass. Although the reaction takes place relatively slowly in structured proteins (half-life, 1–500 days) (45), the deamidation rate increases dramatically when the susceptible residues are exposed and flexible (43, 46, 47). After rearrangements had occurred, the flexibilities of Gly46 and Gly52 in the intermediate species were increased substantially compared with the initial state (Fig. 3, C–E, right panels). This was reasonable, because an additional carbon atom is inserted into the polypeptide backbone due to the rearrangement to βAsp. A similar series of reactions is reported to occur also for glutamine and glutamate residues (48). βAsp and βGlu residues have been detected in a variety of proteins, including calmodulin (49), myelin basic protein (50), eye lens crystallin (51, 52), and Aβ peptide (53). Such rearrangements often control not only the biological activity and function of proteins, but also aggregation propensity (54), exemplified by an increased aggregation of α-crystallin (55) and β-sheet structure of Aβ peptide (56).

In the present study, we found that the rearrangement reaction at Asn plays an important role in fibril formation of GroES by accelerating nucleus formation (Fig. 5). The nucleus formation time was significantly prolonged (by 16 h) in the N51A mutant compared with WT, showing that this rearrangement at Asn51-Gly52 contributes greatly to nucleus formation. On the other hand, the nucleus formation time of the N45A mutant was almost the same as WT and that of the N45A/N51A double mutant was shorter than that of the N51A mutant by 10 h. The differences between the various mutants might be attributed to their relative positions with respect to the fibril core region (Asp58-Lys74). Also, we observe that formation of βAsp newly generates a negatively charged side chain along with the rearrangement of the backbone. Because Asn51 is surrounded by negatively charged residues Glu50 and Glu53, the formation of βAsp51 might also give a more significant electrostatic effect that accelerates nucleus formation of the contiguous region.

A summary of our findings is shown in Fig. 6 as a possible schematic model of GroES amyloid fibril formation. The residues of Val43-Leu57 have been identified as the region where structural changes occur in the early stages of the fibril formation. This region is distinct from the fibril core region (residues of Asp58-Lys74 denoted a β-strand in ribbon model). Formation of this intermediate is attributed to the Asn45-Gly46 and Asn51-Gly52 rearrangement reactions, whose rate is slow but significant under the conditions we used, and especially the rearrangement at Asn51 resulting in βAsp51 accelerates fibril nucleus formation. The fibril nucleus region is located just after this structurally changed region. The intermediate species are soluble and are detectable in
Asn-Gly Transition Leads to Fibrillation of Unfolded GroES

![Diagram of GroES fibrillation](image)

NMR measurements. Once the fibril nucleus is formed, very rapid fibril formation occurs by incorporating random and flexible intact species of GroES monomer. This is why the majority of intact GroES monomers were capable of refolding to native heptamer after being incorporated into mature GroES amyloid fibrils (Fig. 4A). The nucleus formation and the nucleus-dependent extension reactions are too fast to be detected reliably by NMR, and also because of the precipitation of mature fibrils.

Finally, in the present experiments, we found that the GroES unfolded in Gdn-HCl formed amyloid fibrils triggered by the rearrangement at an Asn-Gly site. As this rearrangement-triggered fibril formation also occurred in 3 M urea containing 1 M NaCl (data not shown), it is necessary for GroES to be completely unfolded for this rearrangement to occur to a significant extent. From this point of view, our results may be broadly relative to intrinsically disordered proteins in general, which are abundant in eukaryotes (~33%) and play a number of crucial roles in numerous biological processes (3). Recently, it was also reported that some intrinsically disordered proteins are involved in human neurodegenerative diseases (10). It is also noteworthy that many intrinsically disordered proteins related to neurodegenerative diseases, including Aβ peptide (Alzheimer disease) (53), tau-protein (motor neuron disease with neurofilibrillary tangles) (57), and prion protein (prion disease) (58, 59), undergo rearrangement reactions to form βAsp within their sequences. Clarifying the relation between this β-rearrangement reaction in intrinsically disordered proteins and various amyloidoses would be quite important when medical treatments for the diseases are considered at the molecular level. In our study, we have assigned the ¹H,¹⁵N resonances of disordered GroES and found that the contiguous Asn-Gly sequence immediately preceding the fibril core region is a trigger for GroES amyloid fibril formation. These findings may lead to understanding the mechanism of fibril formation related to amyloidoses in general, as well as the structural characteristics and aggregation propensities of the intrinsically disordered proteins, especially those that are retained for long periods in the cell.

REFERENCES
1. Wright, P. E., and Dyson, H. J. (1999) J. Mol. Biol. 293, 321–331
2. Sugase, K., Dyson, H. J., and Wright, P. E. (2007) Nature 447, 1021–1025
3. Ward, J. J., Sodhi, J. S., McGuffin, L. J., Buxton, B. F., and Jones, D. T. (2004) J. Mol. Biol. 337, 635–645
4. Spillantini, M. G., Schmidt, M. L., Lee, V. M., Trojanowski, J. Q., Jakes, R., and Goedert, M. (1997) Nature 388, 839–840
5. Yagi, H., Kusaka, E., Hongo, K., Mizobata, T., and Kawata, Y. (2005) J. Biol. Chem. 280, 38609–38616
6. Uversky, V. N., and Eliezer, D. (2009) Curr. Protein Pept. Sci. 10, 483–499
7. Scherzinger, E., Lurz, R., Turmaine, M., Mangiariini, L., Hollenbach, B., Hasenbank, B., Bates, G. P., Davies, S. W., Lehach, H., and Wanker, E. E. (1997) Cell 90, 549–558
8. Chen, S., Berthelot, V., Hamilton, J. B., O’Nuallain, B., and Wetzel, R. (2002) Biochemistry 41, 7391–7399
9. Ross, C. A. (2002) Neuron 35, 819–822
10. Uversky, V. N. (2009) Front. Biosci. 14, 5188–5238
11. Bhak, G., Choe, Y. J., and Paik, S. R. (2009) BMB Rep. 42, 541–551
12. Jarrett, J. T., and Lansbury, P. T., Jr. (1993) Cell 73, 1055–1058
13. Wood, S. I., Wypych, J., Steavenson, S., Louis, J. C., Citron, M., and Biere, A. L. (1999) J. Biol. Chem. 274, 19509–19512
14. Laurens, J., Gimbel, D. A., Nyard, H. B., Gilbert, J. W., and Strittmatter, S. M. (2006) J. Biol. Chem. 281, 1698, 133–145
15. Sakono, M., and Zako, T. (2010) FEBS J. 277, 1348–1358
16. Kitamura, A., and Kubota, H. (2010) FEBS J. 277, 1369–1379
17. Nagai, Y., Inui, T., Popiel, H. A., Fujikake, N., Hasegawa, K., Uzade, Y., Goto, Y., Naiki, H., and Toda, T. (2007) Nat. Struct. Mol. Biol. 14, 332–340
18. Naiki, H., and Nagai, Y. (2009) J. Biochem. 146, 751–756
19. Higurashi, T., Yagi, H., Mizobata, T., and Kawata, Y. (2005) J. Biol. Chem. 331, 1057–1069
20. Yagi, H., Sato, A., Yoshida, A., Hattori, Y., Hara, M., Shimamura, J., Sakane, I., Hongo, K., Mizobata, T., and Kawata, Y. (2008) J. Mol. Biol. 377, 1593–1609
21. Gisjarro, J. I., Sunde, M., Jones, J. A., Campbell, I. D., and Dobson, C. M. (1998) Proc. Natl. Acad. Sci. U.S.A. 95, 4224–4228
22. Dobson, C. M. (2001) Philos. Trans. R. Soc. Lond. B Biol. Sci. 356, 133–145
23. Uversky, V. N., and Fink, A. L. (2004) Biochim. Biophys. Acta 1698, 131–153
24. Higurashi, T., Hiragi, Y., Ichimura, K., Seki, Y., Soda, K., Mizobata, T., and Kawata, Y. (2003) J. Biol. Chem. 333, 605–620
25. Makio, T., Arai, M., and Kuwajima, K. (1999) J. Mol. Biol. 293, 125–137
Asn-Gly Transition Leads to Fibrillation of Unfolded GroES

26. Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J., and Bax, A. (1995) J. Biomol. NMR 6, 277–293
27. Johnson, B. A., and Blevins, R. A. (1994) J. Biomol. NMR 4, 603–614
28. Goddard, T. D., and Kneller, D. G. (2008) SPARKY 3, University of California, San Francisco, CA
29. Marsh, J. A., Singh, V. K., Jia, Z., and Forman-Kay, J. D. (2006) Protein Sci. 15, 2795–2804
30. Shevchenko, A., Tomas, H., Havlis, J., Olsen, J. V., and Mann, M. (2006) Nat. Protoc. 1, 2856–2860
31. Yao, J., Dyson, H. J., and Wright, P. E. (1997) FEBS Lett. 419, 285–289
32. Jarymowycz, V. A., and Stone, M. J. (2006) Chem. Rev. 106, 1624–1671
33. Bertoncini, C. W., Rasia, R. M., Lamberto, G. R., Binolfi, A., Zweckstetter, M., Griesinger, C., and Fernandez, C. O. (2007) J. Mol. Biol. 372, 708–722
34. Wu, K. P., Kim, S., Fela, D. A., and Baum, J. (2008) J. Mol. Biol. 378, 1104–1115
35. Higurashi, T., Nosaka, K., Mizobata, T., Nagai, J., and Kawata, Y. (1999) J. Mol. Biol. 291, 703–713
36. Landry, S. J., Zeilstra-Ryalls, J., Fayet, O., Georgopoulos, C., and Gierasch, L. M. (1993) Nature 364, 255–258
37. Shewmaker, F., Maskos, K., Simmerling, C., and Landry, S. J. (2001) J. Biol. Chem. 276, 31257–31264
38. Fiaux, J., Bertelsen, E. B., Horwich, A. L., and Wüthrich, K. (2002) Nature 418, 207–211
39. Fiaux, J., Bertelsen, E. B., Horwich, A. L., and Wüthrich, K. (2004) J. Biomol. NMR 29, 289–297
40. Horst, R., Wider, G., Fiaux, J., Bertelsen, E. B., Horwich, A. L., and Wüthrich, K. (2006) Proc. Natl. Acad. Sci. U.S.A. 103, 15445–15450
41. Tashiro, M., Kojima, M., Kihara, H., Kasai, K., Kamishihara, T., Uèda, K., and Shimotakahara, S. (2008) Biochem. Biophys. Res. Commun. 369, 910–914
42. Clarke, S. (1987) Int. J. Pept. Protein Res. 30, 808–821
43. Geiger, T., and Clarke, S. (1987) J. Biol. Chem. 262, 785–794
44. Aswad, D. W., Paranandi, M. V., and Schurter, B. T. (2000) J. Pharm. Biomed. Anal. 21, 1129–1136
45. Robinson, N. E., and Robinson, A. B. (2001) Proc. Natl. Acad. Sci. U.S.A. 98, 12409–12413
46. Bischoff, R., and Kolbe, H. V. (1994) J. Chromatogr. B Biomed. Appl. 662, 261–278
47. Robinson, N. E., Robinson, M. L., Schulze, S. E., Lai, B. T., and Gray, H. B. (2009) Protein Sci. 18, 1766–1773
48. Wright, H. T. (1991) Crit. Rev. Biochem. Mol. Biol. 26, 1–52
49. Johnson, B. A., Shirokawa, J. M., and Aswad, D. W. (1989) Arch. Biochem. Biophys. 268, 276–286
50. Chou, F. C., Chou, C. H., Sakai, M., Nakamura, T., Morimoto, Y., Kinouchi, T., Goto, Y., and Lampi, K. (2007) Amino Acids 32, 87–94
51. Fujii, N., Ishibashi, Y., Satoh, K., Fujiino, M., and Harada, K. (1994) Biochim. Biophys. Acta 1204, 157–163
52. Sharma, K. K., and Santhoshkumar, P. (2009) Biochim. Biophys. Acta 1790, 1095–1108
53. Roher, A. E., Lowenson, J. D., Clarke, S., Wolkow, C., Wang, R., Cotter, R. J., Reardon, I. M., Zürcher-Neely, H. A., Heinrikson, R. L., and Ball, M. J. (1993) J. Biol. Chem. 268, 3072–3083
54. Shimizu, T., Watanabe, A., Ogawara, M., Mori, H., and Shirasawa, T. (2000) Arch. Biochem. Biophys. 381, 225–234
55. Fujii, N., Shimmymo, Y., Sakai, M., Sadakane, Y., Nakamura, T., Morimoto, Y., Kinouchi, T., Goto, Y., and Lampi, K. (2007) Eur. J. Biochem. 221, 959–964
56. Watanabe, A., Takio, K., and Ihara, Y. (1999) J. Biol. Chem. 274, 7368–7378
57. Weber, D. J., McFadden, P. N., and Caughey, B. (1998) Biochem. Biophys. Res. Commun. 246, 606–608
58. Sandmeier, E., Hunziker, P., Kunz, B., Sack, R., and Christen, P. (1999) Biochim. Biophys. Res. Commun. 261, 578–583
59. Koradi, R., Billeter, M., and Wuthrich, K. (1996) J. Mol. Graph. 14, 51–55, 29–32