Unlocked capacity of proteins to attack membranes characteristic of aggregation: the evil for diseases and aging from Pandora's box

Liangzhong Lim†, Yimei Lu†, and Jianxing Song[*]

Department of Biological Sciences, Faculty of Science, National University of Singapore; 10 Kent Ridge Crescent, Singapore 119260;

Short title: Transformation of proteins from functional players to evils

†The first two authors contribute equally.
: corresponding author; dbssjx@nus.edu.sg.

Key words: Protein aggregation; Aging; *Escherichia coli*; S1 ribosomal protein; Membranes; Bicelle; Liposome; Amphiphipathic helix; Circular dichroism spectroscopy; NMR spectroscopy.

Proteins, the most important functional players for all forms of life, can transform into evils triggering diseases and aging, if their hydrophobic/amphiphilic regions are unlocked by genetic, pathological or/and environmental factors to attack membranes, which is characteristic of severe aggregation.

Abstract

Aggregation of specific proteins is characteristic of a large spectrum of human diseases including all neurodegenerative diseases, while aggregation of non-specific proteins has been now identified to be a biomarker for cellular aging down to *Escherichia coli*. Previously, as facilitated with our discovery in 2005 that "completely insoluble" proteins could be all solubilized in unsalted water [Song (2009) *FEBS Lett*. 583: 953], we found that the TDP-43 prion-like domain in fact contains an intrinsic membrane-interacting subdomain [Lim et al. [2016] *PLoS Biol.* 14, e1002338]. We decrypted that ALS-causing mutations/cofactor-depletion act to render the well-structured folds of cytosolic VAPB-MSP domain and SOD1 into highly disordered states, thus becoming buffer-insoluble. Most surprisingly, this also unlocks the amphiphilic/hydrophobic regions universally exiting in proteins, which thus acquire a novel capacity in abnormally interacting with membranes [Qin et al. (2013) *F1000Res* 2-221.v2; Lim (2016) *BBA-Biomembranes*. 1858: 2223]. Here we aimed extend our discovery to address two fundamental questions: 1) why many *E. coli* proteins become aggregated in aging; and 2) whether aggregation-prone proteins can also acquire a novel capacity in interacting with membranes; by dissecting the 557-residue S1 ribosomal protein into 7 fragments to disrupt its 6 S1 folds, followed by extensive CD and NMR characterizations. The results reveal that we have successfully eliminated all 6 S1 folds and fragment 4 becomes highly disordered and thus buffer-insoluble. Most strikingly, F4 does acquire a capacity in transforming into a helical conformation in membrane environments. Here, for the first time, our study deciphers that like ALS-causing mutants, the disruption of a well-folded *E. coli* cytosolic protein also unlocks its amphiphilic/hydrophobic regions which are capable of abnormally interacting with membranes. Therefore, proteins, the most important functional players for all forms of life, can transform into membrane-toxic forms triggering diseases and aging, if their hydrophobic/amphiphilic regions are unlocked by genetic, pathological or/and environmental factors, which is characteristic of severe aggregation.

Introduction

Proteins are the most important functional players for all forms of life we know of so far. They are linear heteropolymers composed of 20 common α-amino acids, which amazingly are all in L-mirror-image [1]. Although 20 amino acids have distinctive characteristics, they can be briefly divided into two groups: hydrophobic (or non-polar) and hydrophilic (or polar). A portion of proteins spontaneously self-organizes into unique three-dimensional structures via protein folding processes [2-4], while many are fully functional, but lack well-defined structures, and are thus called intrinsically disordered proteins (IDPs) [5]. It is widely recognized that the folding of cytosolic proteins is mainly resulting from solvophobic interactions of polar water molecules with the hydrophobic side chains of proteins [3,4].

One intriguing phenomenon associated with proteins is their insolubility in aqueous buffers. Protein aggregation/insolubility is not only problematic for *in vitro* protein research and industry applications, but is commonly characteristic of a large spectrum of human diseases [7-14], which include all human neurodegenerative diseases [7-11], such as Parkinson's disease (PD), Alzheimer's disease (AD), Huntington's disease (HD), spinocerebellar ataxias (SCA), amyotrophic lateral sclerosis (ALS); as well as diabetes [12] and cardiac dysfunction [14]. Amazingly, aggregation of non-specific proteins have been now identified to be associated with aging of all organisms [15,16]; and in particular the cellular aging and rejuvenation of ***Escherichia coli*** cells has been found to be characteristic of asymmetric segregation of protein aggregates [17].

It has been widely established that aggregation of specific proteins leads to human diseases by "loss of functions" or/and "gain of functions". However, as even for the human neurodegenerative diseases, the involved proteins are so functionally diverse and their knockout (such as SOD1) does not always result in the corresponding diseases. Consequently

4

it is most likely there is no common mechanism underlying "loss of functions". On the other hand, as proteins involved in the diseases are all prone to aggregation, it was widely thought that the formation of protein aggregates represents a common mechanism to trigger the diseases. However, some emerging evidence do not support this hypothesis as: 1) many neurodegenerative diseases such as amyotrophic lateral sclerosis (ALS) triggered by SOD1 mutants were initiated without any detectable protein aggregates [6,9,18]. 2) More radically, it has been increasingly found that inclusion body formation in fact reduces neuronal deaths [19,20].

Previously, partly-soluble proteins involved in human diseases have been found to have capacity to interact with membranes [8,12,18,21-30]. However, many disease-causing mutants were found to be "completely insoluble", thus could not be studied before. However, in 2005, we discovered that all insoluble proteins, even the most hydrophobic integral membrane protein fragment in nature, are all soluble in pure water at concentrations at least up to 100 μM [31- 52]. With this powerful tool, we have been recently focused on characterizing "completely insoluble" mutants/proteins causing ALS [40-51]. We deciphered that the TDP-43 prion-like domain in fact contains a hidden region which is capable of interacting with membranes [50]. We also decoded that ALS-causing mutation, or truncation, or cofactor-depletion acts to disrupt the well-structured native β-folds and consequently they become highly disordered in unsalted water but rapidly aggregated in buffers. Most surprisingly, the mutation/truncation/cofactor-depletion unlocks the amphiphilic/hydrophobic regions universally existing in all proteins, which thus acquire a novel capacity in abnormally interacting with membranes, although these proteins are all well-folded cytosolic proteins and their native functions have nothing to do with membrane interactions [43-45,52].

In the present study, we aimed to extend our discovery to address a fundamental question: whether aggregation of non-specific proteins associated with *Escherichia coli* aging

5

also lead to unlocking the amphiphilic/hydrophobic regions of these proteins, thus acquiring a novel capacity to abnormally interact with membranes. To achieve this, we have examined the aggregation of proteins associated with cellular aging of *E. coli* cells as asymmetric segregation of protein aggregates is the only biomarker characteristic of its cellular aging and rejuvenation [17]. We decided to select a protein for detailed studies which satisfies three criteria: 1) the protein should be extensively found in the *E. coli* aggregation lists. 2) It contains well-structured folds rich in β-sheets. 3) It owns no membrane-interacting region and its native functions have nothing to do with membrane-interaction. As such, the S1 ribosomal protein was selected as it was identified to aggregate severely in an *E. coli* strain carrying the mutant of YajL protein [53], a prokaryotic homolog of parkinsonism-associated protein DJ-1 which functions to repair proteins in response to the global stress particular oxidative stress with multi-functions such as acting as a covalent chaperone [53-56]. Furthermore, the aggregate of S1 ribosomal protein was even identified in healthy *E. coli* cells [57].

As shown in Fig 1, S1 ribosomal protein contains 6 well-structured S1 domains linked by flexible regions [57,58]. To mimic fragmentation resulting from radical-mediated protein oxidation associated with aging [60-63], we dissected S1 ribosomal protein into 7 fragments with the cleavage sites all located within the well-structured S1 folds (Fig 1), in an attempt to eliminate the folded structures of all 6 S1 domains. We obtained recombinant proteins of the full-length S1 ribosomal protein and its 7 dissected fragments, which were then extensively characterized by CD and NMR spectroscopy as we previously conducted on other proteins [30-52,64,65]. The results reveal that the successful elimination of the S1 folds indeed leads to unlocking the amphiphilic/hydrophobic regions which acquire a novel capacity in abnormally interacting with membranes. Therefore, proteins, the most important functional players for all forms of life, can transform into membrane-toxic forms triggering diseases and

aging, if their hydrophobic/amphiphilic regions are unlocked by genetic, pathological or/and environmental factors to attack membranes, which is characteristic of severe aggregation.
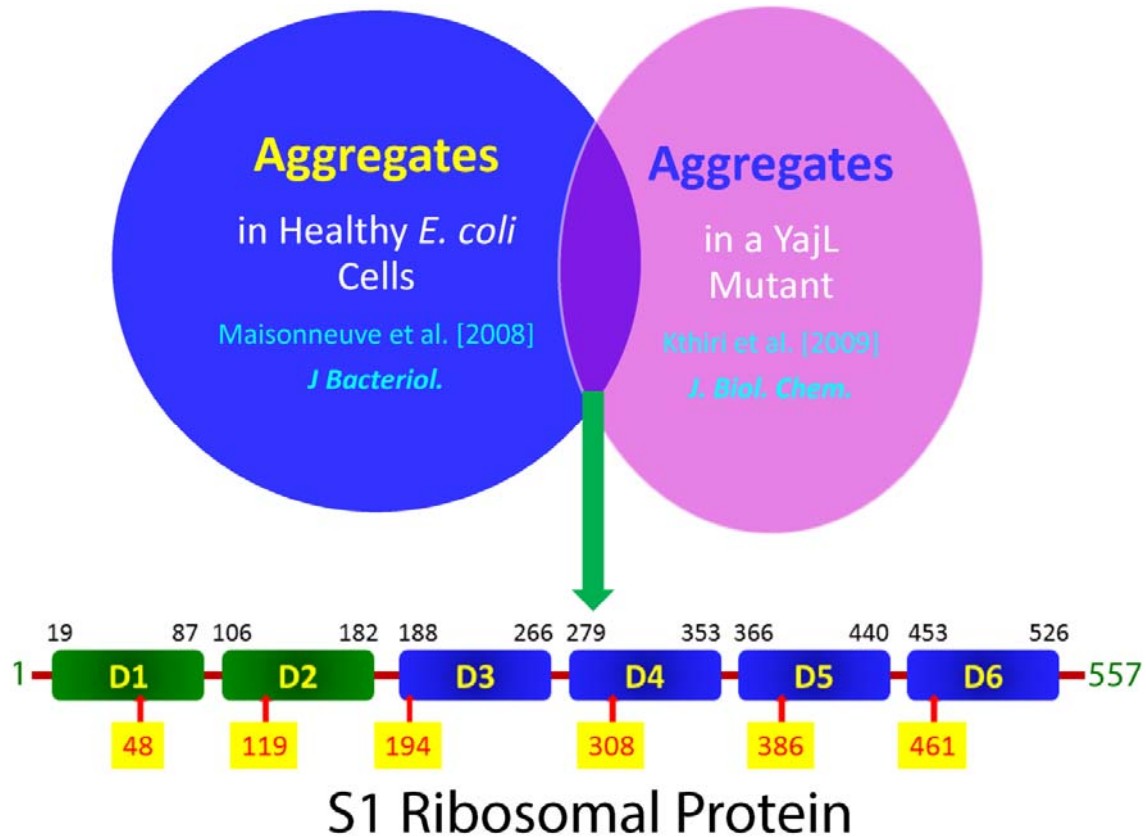

Fig 1. Selection and domain organization of *E. coli* S1 ribosomal protein.
The S1 ribosomal protein was found in lists of protein aggregates of both healthy *E. coli* cells and of a YajL mutant cell strain. *E. coli* S1 ribosomal protein consists of 6 well-structured S1 domains. The starting residues for dissection are labeled, which were all located within the 6 S1 domains.

Results

1. Dissection successfully eliminated the folded structures of six S1 domains.

To mimic the fragmentation due to protein damages such as by radical-mediated over-oxidation, we dissected the *E. coli* S1 ribosomal protein into 7 fragments with all 6 cleavage sites located within 6 S1 domains (Fig 1), which aimed to abolish the well-structured S1 fold. Subsequently we have successfully expressed and purified recombinant proteins of the full-length and its 7 dissected fragments. Interestingly, except for the forth fragment (F4) which was only soluble in Milli-Q water, all the rest could be first dissolved in Milli-Q water and subsequently diluted into 1 mM phosphate buffer at pH 6.8 to reach a concentration of at least 100 µM. Fig 2 presents far-UV CD spectra of the full-length and 7 fragments at a concentration of 15 µM in both Milli-Q water (pH 4.0) and 1 mM phosphate buffer (pH 6.8). However, F4 became completely precipitated upon dilution into the buffer and consequently no CD signal was detected in the buffer.
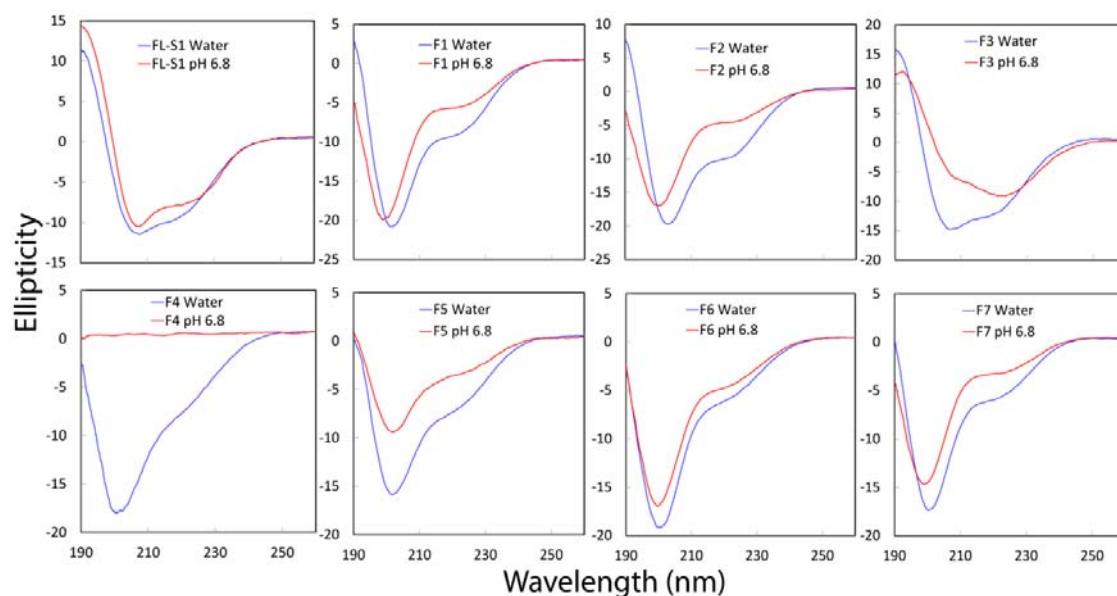

Fig 2. CD characterization.
Far-UV CD spectra of the full-length S1 protein and its 7 dissected fragments in Milli-Q water (blue) and in 1 mM phosphate buffer at pH 6.8 (red).

8

We further utilized NMR HSQC spectroscopy to characterize the solution conformations of the full-length and 7 fragments at a concentration of 100 µM in both Milli-Q water (pH 4.0) and 1 mM phosphate buffer (pH 6.8). As shown in Fig 3, probably due to the µs-ms conformational exchanges or/and dynamic aggregation, even in Mill-Q water the full-length has an HSQC spectrum lacking of well-dispersed NMR peaks typical of a well-folded proteins. In the buffer, most NMR peaks became too broadened to be detected (Fig 3), implying that salt ions can significantly enhance µs-ms conformational exchanges or/and dynamic aggregation as we previously deciphered on other proteins [31-52], although the secondary structures remain similar in Milli-Q water and the buffer as reflected by CD spectra (Fig 2). Most strikingly, all 7 dissected fragments had HSQC spectra which are also absent of any well-dispersed NMR peaks which were previously observed on isolated S1 domains [57,58]. This clearly indicates that our dissection successfully eliminated the ability of all 7 fragments to fold into the native S1 structures.
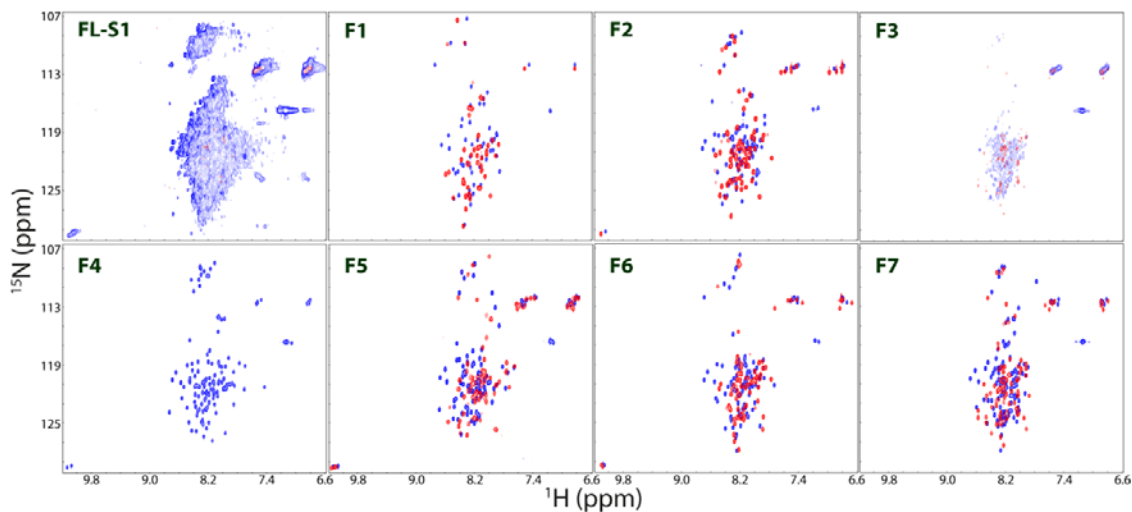

Fig 3. NMR HSQC characterization.
NMR HSQC spectra of the full-length S1 protein and its 7 dissected fragments in Milli-Q water (blue) and in 1 mM phosphate buffer at pH 6.8 (red).

9

2. Interactions with DMPC/DHPC bicelle and liposome.

To assess whether the full-length and 7 fragments have potential to interact with membranes, we titrated them with DMPC/DHPC bicelle as monitored by CD spectroscopy as we previously conducted on ALS-causing P56S-MSP mutant and TDP-43 prion-like domain. As shown in Fig 4, the full-length only had very small changes of the CD spectra upon adding the bicelle at different ratios. Out of 7 fragments, only F4 and F5 showed significant changes of the CD spectra upon adding the bicelle. In particular, as judged by CD spectra, F4 had a significant conformational transition from a highly disordered state in aqueous solution to helical conformation in bicelle. Furthermore, F4 also underwent a similar conformational transition upon interacting with liposome prepared from the total extract of *E. coli* lipids.
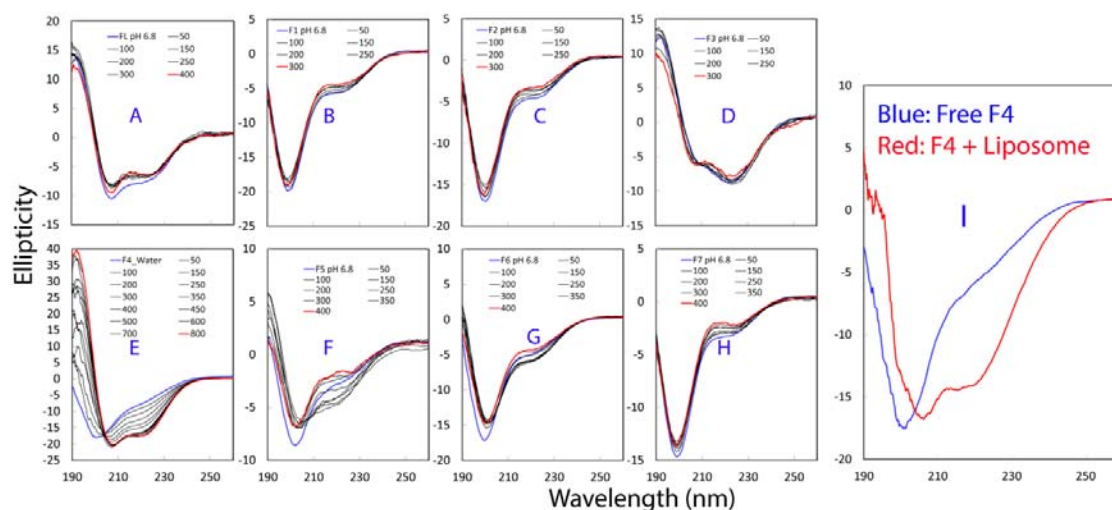

Fig 4. CD characterization of interactions with bicelle and liposome.
(A-H) Far-UV CD spectra of the full-length S1 protein and its 7 dissected fragments in the presence of DMPC/DHPC bicelle at different molar ratios. (I) Far-UV CD spectra of F4 in the absence (blue) and in the presence (red) of liposome prepared from the total extract of *E. coli* lipids.

As F4 was found to become rapidly precipitated in buffers, as well as to undergo significant conformational changes upon interacting with bicelle, we decided to further characterize its interaction with bicelle by NMR HSQC spectroscopy. As seen in Fig 5, HSQC peaks of F4 shifted significantly upon adding bicelle and the shift became largely

saturated at a molar ratio of 1:200 (F4:bicelle). This clearly reveals that F4 was able to interact with bicelle to undergo a significant conformational change, consistent with the CD results (Fig 4).
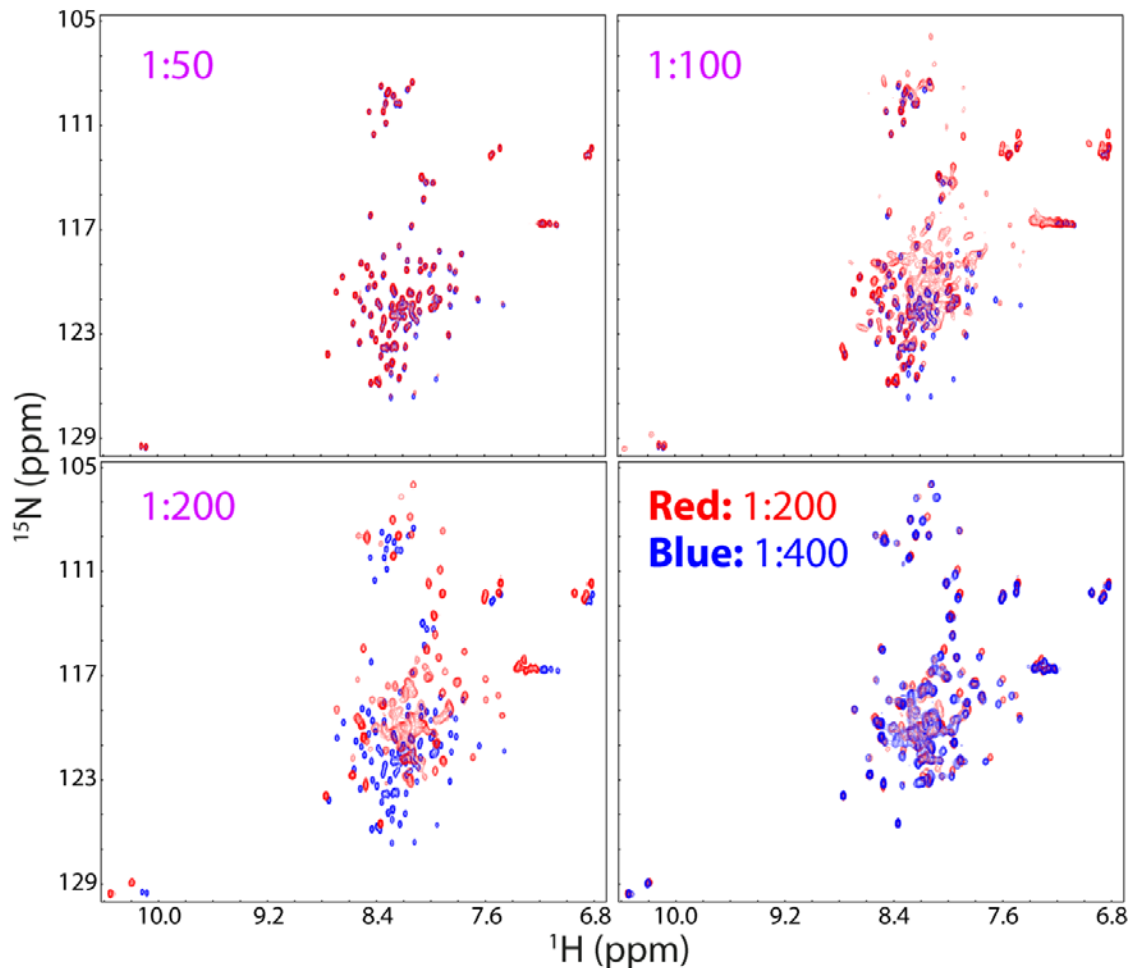

Fig 5. NMR HSQC characterization of interactions of F4 with bicelle.
NMR HSQC spectra of the fragment 4 in the presence of DMPC/DHPC bicelle at different molar ratios.

3. Residue-specific conformations of F4 in aqueous solution and bicelle.

To gain insights into the conformational changes of F4 upon interacting with bicelle, by analyzing three-dimensional NMR spectra including CCC(CO)NH, HN(CO)CACB, here we successfully achieved NMR sequential assignments of F4 in aqueous solution and in bicelle. Fig 6 presents their ($\Delta C\alpha$-$\Delta C\beta$) chemical shifts, which represent a sensitive indicator

of the residual secondary structures in disordered proteins [66]. In aqueous solution, F4 has small absolute values of (ΔCα-ΔCβ) chemical shifts over the whole sequence (Fig 6A), which clearly indicates that it is indeed lacking of any stable secondary structure, completely consistent with the CD results (Fig 4). To gain quantitative insights into the populations of different secondary structures, we further analyzed its NH, N, Hα, Cα and Cβ chemical shifts by SSP program [67]. As seen in Fig 6B, all residues have the absolute values of SSP scores less than 0.4, confirming that the whole domain has no stable secondary structure. Interestingly, the C-terminal residues over 279-306 of F4 adopt three β-strands in the forth S4 domain as previously determined by NMR spectroscopy [58,59]. This suggests that the abolishment of the well-structured folds leads to a complete loss of the β-stranded secondary structures, as we extensively observed on SH3, P56S-MSP and SOD1 mutants [31-52].
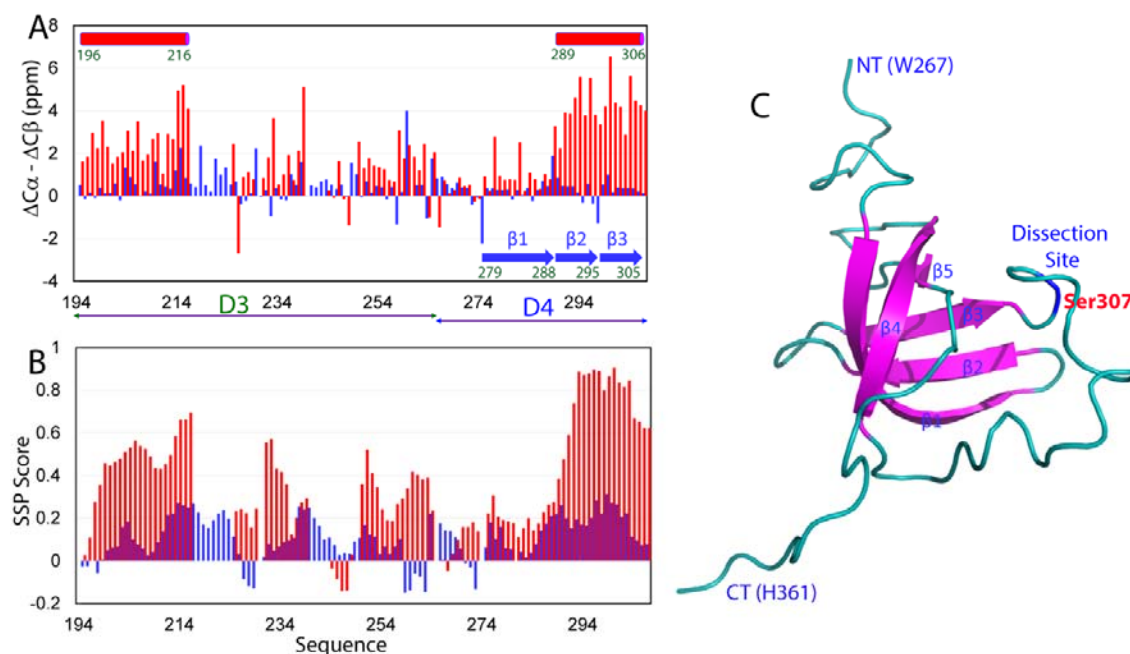

Fig 6. Residue-specific conformations of F4.
(A) Residue specific (ΔCα-ΔCβ) chemical shifts of F4 in in aqueous solution (blue) and in DMPC/DHPC bicelle (red). (B) Secondary structure scores of F4 in in aqueous solution (blue) and in DMPC/DHPC bicelle (red), which were obtained by analyzing their chemical shifts with the SSP program. A score of +1 is for the well-formed helix while a score of -1 for the well-formed extended strand. (C) NMR structure of the forth S1 domain with the dissection site labeled.

By a sharp contrast, upon interacting with bicelle, many residues of F4 suddenly had large and positive (ΔCα–ΔCβ) chemical shifts (Fig 6A), indicating that they became highly helical (Fig 6B). In particular, the N-terminal residues Val195-Val214 and C-terminal residues Tyr290-Val306 became highly helical as quantitatively reflected by their positive and large SSP scores (Fig 6B). The results together provide residue-specific evidence that upon interacting with bicelle, many residues of F4, particular the N- and C-terminal regions undergo a conformational transition from a predominantly disordered state to highly helical conformation.

4. Principle mediating the formation of membrane-embedded amphiphilic helices.

To decode the principle mediating the formation of the helical conformations of the N- and C-terminal regions of F4 upon interacting with bicelle, we plotted their helical wheel diagrams for residues Val195-Val214 (Fig 7A) and Tyr290-Val306 (Fig 7B). Indeed, both of them are predicted to have capacity in adopting "amphiphilic helix", which is characterized by spatial segregation of polar and non-polar amino acids that are located on opposing faces as oriented along the long axis of the helix. Amphiphilic helix was first identified in myoglobin and haemoglobin in 1965 [68], and its importance in mediating protein-lipid interactions was first recognized in 1974 [69]. Hydrophobic moment, or amphiphilicity, reflecting the periodicity in protein hydrophobicity, was proposed to detect the ability of a protein sequence to form amphiphilic helix [70,71], which has been now demonstrated to mediate various protein-membrane interactions [70-74]. Amazingly, bioinformatics studies revealed that all proteins universally have segments with high intrinsic amphiphilicity even including randomly generated sequences, regardless of their native structures [75,76].

Indeed, based on dihedral angle constraints derived from the NMR chemical shifts, we modeled the structures of the two fragments and both of them form amphiphilic helice

13

(Fig 7C and 7F), with one side being largely non-polar (Fig 7D and 7G); and another side highly polar (Fig 7E and 7H).
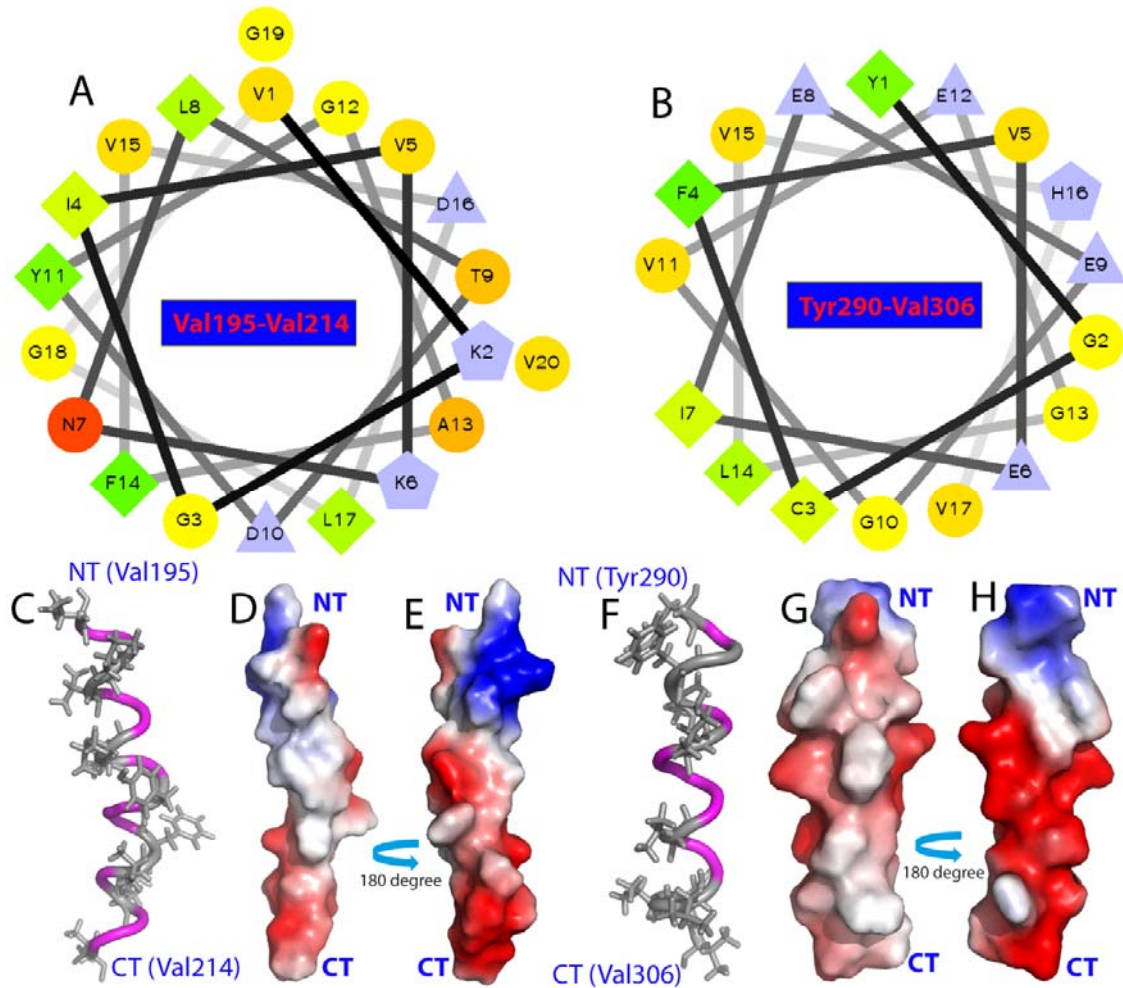

Fig 7. Visualization of two amphiphathic helices of F4 formed in bicelle.
Helical wheel diagrams of the segments over residues Val195-Val214 (A) and Tyr290-Val360 (B); which are characterized by spatial segregation of polar and non-polar amino acids that are located on opposing faces as oriented along the long axis of the helix. Three-dimensional models and their electrostatic potential surfaces respectively for Val195-Val214 (C-E) and Tyr290-Val360 (F-H) in bicelle.

We calculated the hydrophobic moment or amphiphilicity of F4 and interestingly, even for such a short fragment, four regions have high amphiphilicity, which include Val195-Val214, His219-Leu246, Trp267-Thr285 and Tyr290-Val360 (Fig 8A), consistent with previous reports that amphiphilic fragments universally exist in all proteins [75,76]. However,

why out of the 4 segments, only the N- and C-terminal segments were experimentally shown to have high capacity in interacting with bicelle ?
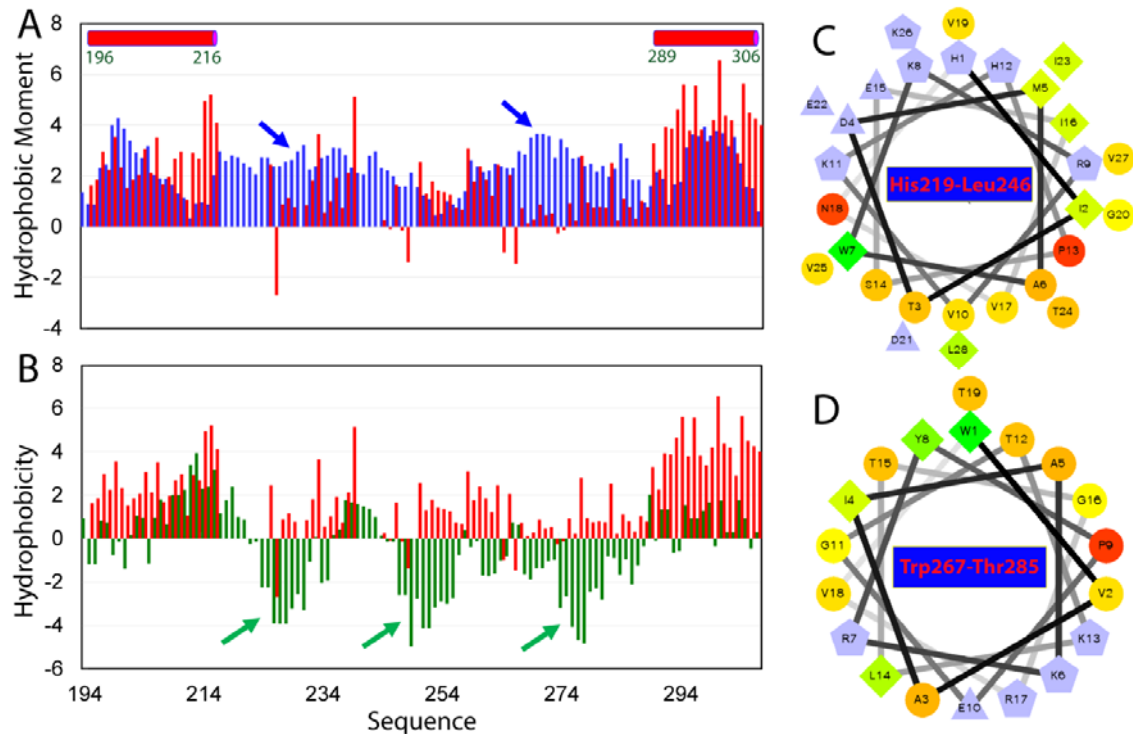

Fig 8. Membrane-interactions mediated by both amphiphilicity and hydrophobicity.
(A) Amphiphilicity (blue) and residue specific (ΔCα-ΔCβ) chemical shifts of F4 in DMPC/DHPC bicelle (red). (B) Hydrophobicity (blue) and residue specific (ΔCα-ΔCβ) chemical shifts of F4 in DMPC/DHPC bicelle (red). Helical wheel diagrams of the segments over residues His219-Leu246 (C) and Trp267-Thr285 (D). To make them comparable with the (ΔCα-ΔCβ) chemical shifts, the values of Amphiphilicity and Hydrophobicity times a factor of 3.

We thus calculated the hydrophobicity scale [77] of F4 (Fig 8B), and interestingly the middle two segments over His219-Leu246 and Trp267-Thr285 have negative hydrophobicity scales for most of their residues. Indeed, examination of the helical wheel diagrams (Fig 8C and 8D) reveals that their hydrophobic surfaces are much smaller than those of Val195-Val214 (Fig 7A) and Tyr290-Val306 (Fig 7A). This observation immediately suggests that both amphiphilicity and hydrophobicity need to be sufficiently high for a protein segment to have strong ability in forming a membrane-induced amphiphilic helix. To confirm this finding, we also calculated the amphiphilicity and hydrophobicity for the rest of the dissected

15

fragments (Fig 9). Indeed, only F5 has two segments with both amphiphilicity and hydrophobicity being relatively large (Fig 9), which is completely consistent with the CD results that F5 also had large conformational changes upon interacting with bicelle (Fig 4).
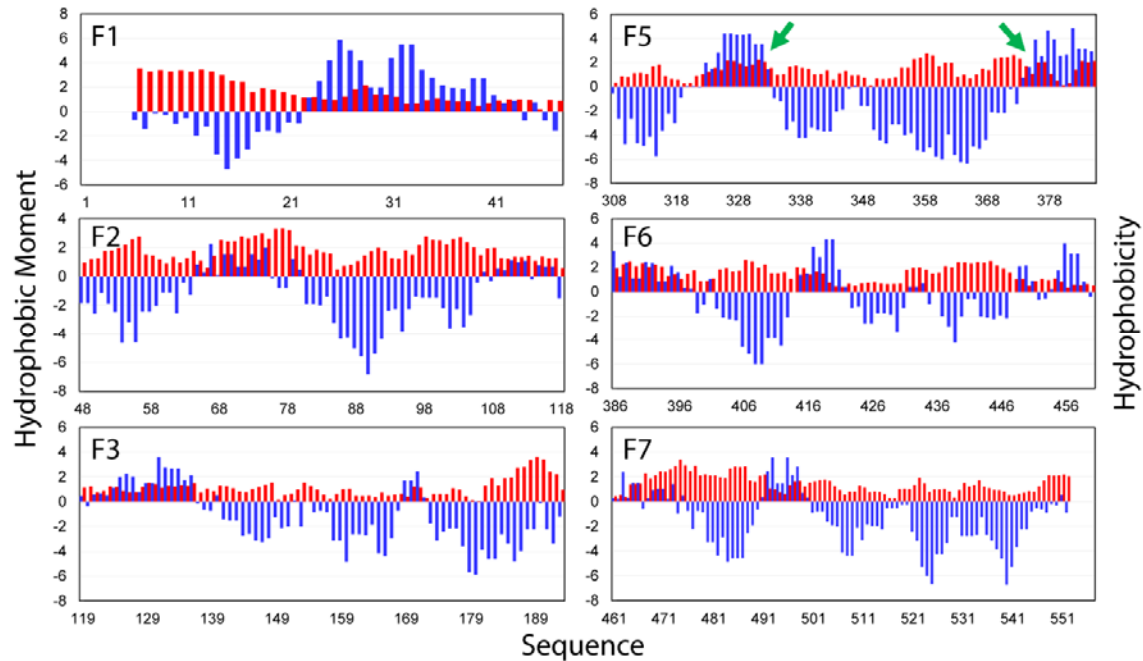
Fig 9. Amphiphilicity and Hydrophobicity of different S1 fragments.
(A) Hydrophobicity (blue) and Amphiphilicity (red) of 6 dissected fragments of the *E. coli* S1 ribosomal protein. The values of Amphiphilicity and Hydrophobicity times a factor of 3.

However, considering the fact such diverse amphiphiphic regions are involved in protein-lipid interactions [73], it is possible that other fragments other than F4 are also interact with membranes/lipids with specific chemical properties.

Discussion

There are around 300 hypotheses of aging and the best known is the free radical hypothesis. Proteins are highly susceptible to oxidative damages which lead to many irreversible covalent modifications including fragmentation of protein backbones [60-63]. In particular, during aging of *E. coli* cells, oxidative damage represents a major factor in triggering aggregation of many non-specific proteins [53-56,63]. Unfortunately, the mechanism underlying the oxidation-induced protein aggregation still remains highly elusive.

Here, to mimic the oxidation-induced protein fragmentation, we dissected the *E. coli* S1 ribosomal protein into 7 fragments. Detailed CD and NMR studies reveal that the dissection did eliminate all 6 well-structured S1 domains. However, out of 7 fragments, only F4 becomes "completely insoluble" while the other 6 fragments are still soluble in buffer to some extent. As all 7 fragments contain parts of the well-folded regions in 6 all-β S1 domains, this implies that despite the elimination of the tertiary folds, fragments except for F4 can rearrange their secondary structures from β-strands into helices to bury hydrophobic regions in order to avoid "complete insolubility", as we previously demonstrated on a Nck2 SH3 domain [83]. On the other hand, because F4 has longer regions which have both high amphiphilicity and hydrophobicity (Fig 8 and 9), it thus appears impossible for F4 to have them all properly buried. As a consequence, it becomes "completely insoluble" even in 1 mM phosphate buffer. Furthermore, we observed that other fragments also started to precipitate in buffers containing high concentrations of salts. Therefore, our current results with protein fragmentation enforce our previous discovery that modifications such as a point mutation or cofactor depletion in β-rich proteins such as S1, SH3, MSP and SOD1 folds, is sufficient to completely eliminate their tertiary structures, which accounts for reduction of solubility or even becoming "completely-insoluble" [31-52]. This mechanism also rationalizes previous observations that β-rich proteins were significantly over-represented in aggregation-prone

17

proteins relative to the proteomes [81], because the tertiary folds rich in β-sheets are relatively easier to be abolished by sequence modifications [31-52].

Remarkably, the asymmetric segregation of protein aggregates in *E. coli* led to the aged mother cells holding the accumulated aggregates and rejuvenated daughter cells free of aggregates [17]. This implies that the aggregation-prone proteins may also trigger cellular aging of *E. coli* cells by a "gain of cytotoxicity" mechanism. So a fundamental question remaining to be answered is: whether the elimination of the well-folded structures of *E. coli* cytosolic proteins will also lead to gain of a novel capacity to abnormally interact with membrane? Here we experimentally reveal that the fragments like F4 resulting from the oxidation-induced fragmentation of an *E. coli* cytosolic protein do acquire a novel and strong capacity in abnormally interacting with membranes, if the fragments have both high amphiphilicity and hydrophobicity, which are thus prone to severe aggregation in cells or buffers [31-52]. In other words, the proteins/fragments characteristic of severe aggregation also own high capacity to attack membranes, as we previously decoded for ALS-causing mutants [32,43-45,50,52]. Consequently, this suggests that to abnormally interact with membranes may also represent a common mechanism by which aggregation-prone proteins cause aging by a "gain of cytotoxicity" mechanism.

Our current study on a protein of *E. coli*, the lowest-level organism, decrypt that the *E. coli* protein also follows the same mechanism of aggregation and gain of membrane-toxicity as the proteins of human being, the highest organism [31-52]. Therefore. proteins, the most important functional players for all forms of life, have potentials to transform from functional players to evils triggering diseases and aging by abnormally attacking membranes, if their hydrophobic/amphiphilic regions are unlocked by genetic, pathological or/and environmental factors, which is characteristic of severe aggregation. So the proteome appears to be Pandora's box.

It has been shown that in *E. coli*, most cytosolic proteins need to fold into well-defined structures for their functions [84-85], while in eukaryotic organisms there are intrinsically disordered proteins [5], or even intrinsically insoluble proteins [32], which were sent for degradation upon synthesis [86-87]. In cells, intrinsically-insoluble, misfolded, and structure-eliminated proteins are expected to rapidly form aggregates due to a high salt concentration (~150 mM) in cellular environments [31-52]. However, although the chance might be extremely rare, if some of these aggregation-prone proteins/fragments escape from forming "completely insoluble" aggregates, they may transform into evils to trigger diseases and aging by attacking membranes. This also explains the recent observation that human diseases were initiated by the soluble oligomers but not the insoluble aggregates or amyloid fibers, as the formation of soluble oligomers may facilitate these proteins to access membranes. Although it has been recently revealed that the abnormal insertion of SOD1 mutants into ER membranes without forming detectable aggregates is sufficient to initiate ALS pathogenesis [18], the coupling of the membrane-interaction and aggregation/amyloid-formation onto/within membranes may significantly strengthen the capacity of aggregation-prone proteins to cause diseases and aging [45,50,52], in particular for *E. coli* cells which do not have a complex membrane-network within their cellular space.
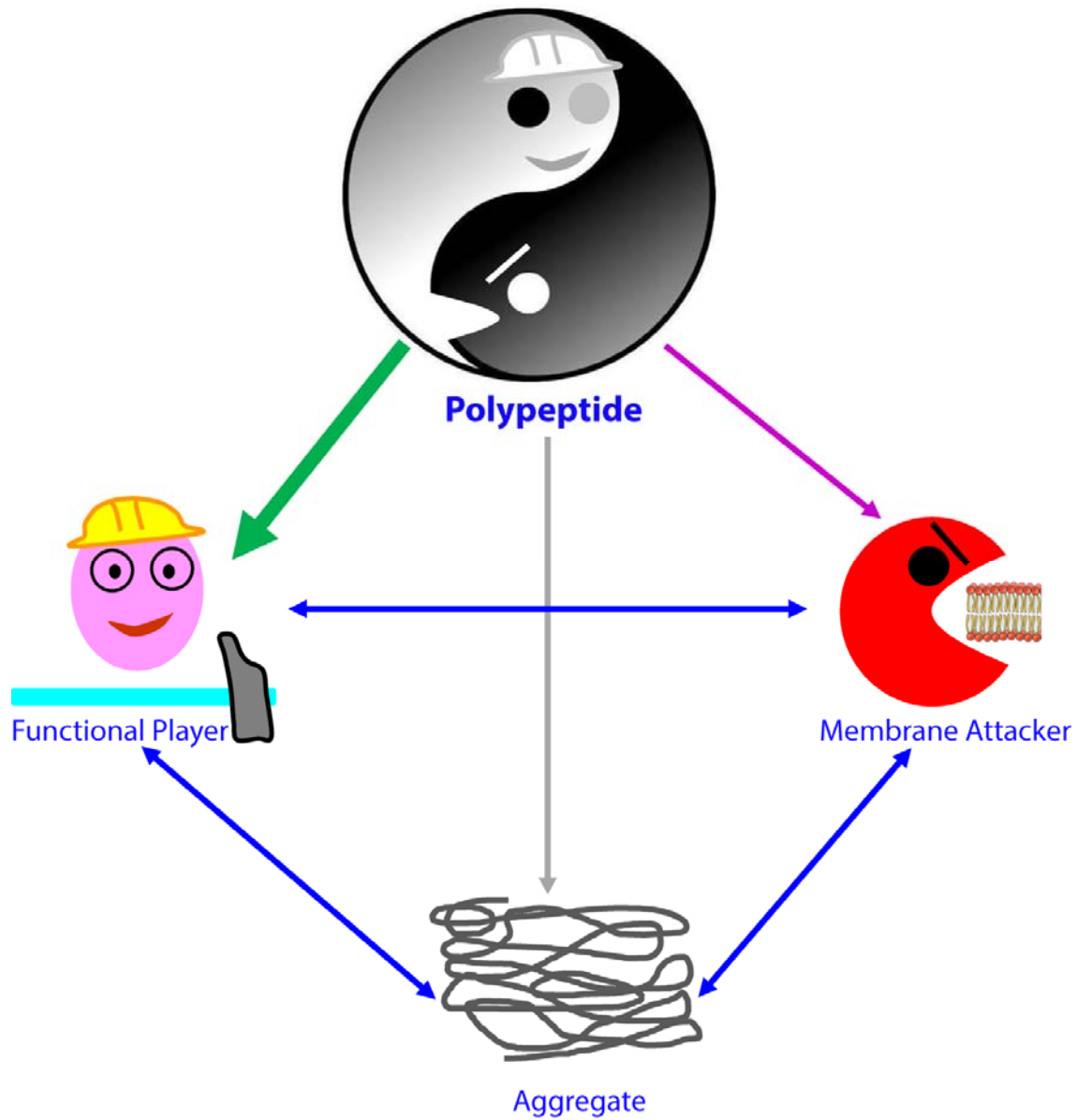
Fig 10. Proposed diagram for transformation of a protein from a functional player to an evil by unlocking the potential capacity to attack membranes, which is universally characteristic of severe aggregation.

Methods

Generation of Recombinant Proteins

The DNA fragment encoding *E. coli* S1 ribosomal protein was amplified by PCR reaction directly on *Escherichia coli* BL21 (DE3) cells (Novagen) with designed primers and subsequently cloned into a modified vector pET28a with 6 His residues at C-terminus as we extensively used for the TDP-43 prion-like domain [51]. Furthermore, DNA fragments encoding 7 dissected fragments (Fig 1) were successfully obtained by PCR reactions with designed primers and subcloned into the same vector. The expression vectors were subsequently transformed into and overexpressed in *Escherichia coli* BL21 (DE3) cells (Novagen). The recombinant proteins were found inclusion body, and consequently were purified by a Ni^{2+}-affinity column (Novagen) under denaturing conditions in the presence of 8 M urea. The fractions containing the recombinant proteins were acidified by adding 10% acetic acid and subsequently purified by reverse-phase (RP) HPLC on a C4 column eluted by water-acetonitrile solvent system. The HPLC elutions containing pure recombinant proteins were lyophilized and stored in -80 °C.

The generation of the isotope-labelled proteins for NMR studies followed a similar procedure except that the bacteria were grown in M9 medium with the addition of $(^{15}NH_4)_2SO_4$ for ^{15}N labeling and $(^{15}NH_4)_2SO_4/[^{13}C]$-glucose for double labelling [31-52]. The purity of the recombinant proteins was checked by SDS–PAGE gels and their molecular weights were verified by a Voyager STR matrix-assisted laser desorption ionization time-of-flight-mass spectrometer (Applied Biosystems). The concentration of protein samples was determined by the UV spectroscopic method in the presence of 8 M urea. Briefly, under the denaturing condition, the extinct coefficient at 280 nm of a protein can be calculated by adding up the contribution of Trp, Tyr and Cys residues [88].

CD and NMR experiments

All circular dichroism (CD) experiments were performed on a Jasco J-810 spectropolarimeter equipped with a thermal controller using 1-mm path length cuvettes. Data from five independent scans were added and averaged. CD samples were prepared by diluting the concentrated samples dissolved in Milli-Q water (pH 4.0) into 1 mM phosphate buffer to reach the final concentration of 15 μM at pH 6.8.

To assess the capacity of the full-length S1 protein and its 7 dissected fragments to interact with membranes, here we used DMPC/DHPC bicelle to mimic the bilayer membrane, which was prepared by mixing up dimyristoylphosphatidylcholine (DMPC) and dihexanoylphophatidylcholine (DHPC) at a q value of 0.25 as previously described [43,51]. The liposome was prepared as we previously described [43] but with the total extract of *E. coli* lipids.

All NMR experiments were acquired on an 800 MHz Bruker Avance spectrometer equipped with pulse field gradient units as described previously [31-52,89,90]. For characterizing the residue-specific conformations of the fragment 4 in both aqueous solutions and bicelle, two pair of triple-resonance experiments HN(CO)CACB, CBCA(CO)NH were collected for the sequential assignment on ^{15}N-/^{13}C-double labelled samples respectively in Milli-Q water at pH 4.0 and in bicelle at protein concentration of 300 μM. NMR data were processed with NMRPipe [91] and analyzed with NMRView [92].

Structure Modelling

To visualize the conformation in bicelle of two helical segments of F4 respectively over residues Val195-Val214 and Tyr290-Val360, the backbone dihedral angles were generated with TALOS+ by inputting backbone ^{1}H, ^{15}N and ^{13}C chemical shifts [93], which

were utilized to generate the model of two segments by CYANA [45,48,50,52,94]. The

lowest target-function CYANA structures with no dihedral angle violation >4 degrees were

selected for analysis.

Author Contributions

Conceived and designed the experiments: JXS; Performed the experiments: LZL YML.

Analyzed the data: JXS LZL YML. Prepared figures and wrote the paper: JXS.

Acknowledgement

References

1. Barron LD: Chirality and life. *Space Sci Rev*. 2008; 135 (1–4): 187–201.

2. Anfinsen CB: Principles that govern the folding of protein chains. *Science*. 1973; 181 (4096): 223–230.

3. Kim PS, Baldwin RL: Specific intermediates in the folding reactions of small proteins and the mechanism of protein folding. *Annu Rev Biochem*. 1982; 51: 459–489.

4. Dill KA: Dominant forces in protein folding. *Biochemistry*. 1990; 29: 7133–7155.

5. van der Lee R, Buljan M, Lang B, Weatheritt RJ, Daughdrill GW, Dunker AK, Fuxreiter M, Gough J, Gsponer J, Jones DT, Kim PM, Kriwacki RW, Oldfield CJ, Pappu RV, Tompa P, Uversky VN, Wright PE, Babu MM. Classification of intrinsically disordered regions and proteins. *Chem Rev*. 2014;114(13):6589-631.

6. Ross CA, Poirier MA: Opinion: What is the role of protein aggregation in neurodegeneration? *Nat Rev Mol Cell Biol*. 2005; 6: 891–8.

7. Chiti F, Dobson CM: Protein misfolding, functional amyloid, and human disease. *Annu Rev Biochem*. 2006; 75: 333–366.

8. Kotler SA, Walsh P, Brender JR, Ramamoorthy A. Differences between amyloid-β aggregation in solution and on the membrane: insights into elucidation of the mechanistic details of Alzheimer's disease. *Chem Soc Rev*. 2014; 43: 6692–6700.

9. Ling SC, Polymenidou M, Cleveland DW. Converging mechanisms in ALS and FTD: disrupted RNA and protein homeostasis. *Neuron*. 2013; 79: 416–438.

10. Hock EM1, Polymenidou M1. Prion-like propagation as a pathogenic principle in frontotemporal dementia. *J Neurochem*. 2016;138 Suppl 1:163-83.

11. Lansbury PT, Lashuel HA: A century-old debate on protein aggregation and neurodegeneration enters the clinic. *Nature*. 2006; 443 (7113): 774–779.

12. Brender JR, Salamekh S, Ramamoorthy A. Membrane disruption and early events in the aggregation of the diabetes related peptide IAPP from a molecular perspective. *Acc Chem Res.* 2012; 45: 454– 462.

13. Hebert DN, Molinari M: In and out of the ER: Protein folding, quality control, degradation, and related human diseases. *Physiol Rev.* 2007; 87 (4): 1377–1408.

14. Willis MS, Patterson C. Proteotoxicity and cardiac dysfunction--Alzheimer's disease of the heart? *N Engl J Med.* 2013;368(5):455-64.

15. Lindner AB, Demarez A: Protein aggregation as a paradigm of aging. *Biochim Biophys Acta.* 2009; 1790 (10): 980–996.

16. Kourtis N, Tavernarakis N: Cellular stress response pathways and ageing: Intricate molecular relationships. *EMBO J.* 2011; 30 (13): 2520–2531.

17. Lindner AB, Madden R, Demarez A, Stewart EJ, Taddei F. Asymmetric segregation of protein aggregates is associated with cellular aging and rejuvenation. *Proc Natl Acad Sci U S A.* 2008; 105: 3076–3081.

18. Sun S, Sun Y, Ling SC, Ferraiuolo L, McAlonis-Downes M, Zou Y, Drenner K, Wang Y, Ditsworth D, Tokunaga S, Kopelevich A, Kaspar BK, Lagier-Tourenne C, Cleveland DW. Translational profiling identifies a cascade of damage initiated in motor neurons and spreading to glia in mutant SOD1-mediated ALS. *Proc Natl Acad Sci U S A.* 2015; 112(50):E6993-7002.

19. Arrasate M, Mitra S, Schweitzer ES, Segal MR, Finkbeiner S. Inclusion body formation reduces levels of mutant huntingtin and the risk of neuronal death. *Nature.* 2004; 431(7010): 805-10.

20. Bodner RA, Outeiro TF, Altmann S, Maxwell MM, Cho SH, Hyman BT, McLean PJ, Young AB, Housman DE, Kazantsev AG. Pharmacological promotion of inclusion formation:

a therapeutic approach for Huntington's and Parkinson's diseases. *Proc Natl Acad Sci U S A.* 2006; 103(11):4246-51.

21. Liu J, Lillo C, Jonsson PA, Vande Velde C, Ward CM, Miller TM, Subramaniam JR, Rothstein JD, Marklund S, Andersen PM, Brännström T, Gredal O, Wong PC, Williams DS, Cleveland DW. Toxicity of familial ALS-linked SOD1 mutants from selective recruitment to spinal mitochondria. *Neuron.* 2004; 43(1):5-17.

22. Vande Velde C, Miller TM, Cashman NR, Cleveland DW. Selective association of misfolded ALS-linked mutant SOD1 with the cytoplasmic face of mitochondria. *Proc Natl Acad Sci U S A.* 2008; 105(10):4022-7.

23. Auluck PK, Caraveo G, Lindquist S. α-Synuclein: membrane interactions and toxicity in Parkinson's disease. *Annu Rev Cell Dev Biol.* 2010; 26:211-33.

24. Elfrink K, Ollesch J, Stöhr J, Willbold D, Riesner D, Gerwert K. Structural changes of membrane-anchored native PrP(C), *Proc. Natl. Acad. Sci. U. S. A.* 2008; 105: 10815–10819.

25. Shao H, Jao S, Ma K, Zagorski MG. Solution structures of micelle-bound amyloid beta-(1–40) and beta-(1–42) peptides of Alzheimer's disease, *J. Mol. Biol.* 1999; 285: 755–773.

26. Künze G, Barré P, Scheidt HA, Thomas L, Eliezer D, Huster D. Binding of the three-repeat domain of tau to phospholipid membranes induces an aggregated-like state of the protein, *Biochim. Biophys. Acta* 2012; 1818 (9) 2302–2313.

27. Kegel KB, Sapp E, Yoder J, Cuiffo B, Sobin L, Kim YJ, Qin ZH, Hayden MR, Aronin N, Scott DL, Isenberg G, Goldmann WH, DiFiglia M. Huntingtin associates with acidic phospholipids at the plasma membrane, *J. Biol. Chem.* 2005; 280(43): 36464–36473.

28. Brender JR1, Salamekh S, Ramamoorthy A. Membrane disruption and early events in the aggregation of the diabetes related peptide IAPP from a molecular perspective, *Acc. Chem. Res.* 2012; 45(3): 454–462.

29. Kagan BL, Jang H, Capone R, Teran Arce F, Ramachandran S, Lal R, Nussinov R. Antimicrobial properties of amyloid peptides, *Mol. Pharm.* 2012; 9(4): 708–717.

30. Jiang Z, de Messieres M, Lee JC. Membrane remodeling by α-synuclein and effects on amyloid formation, *J. Am. Chem. Soc.* 2013; 135(43): 15970–15973.

31. Song J. Insight into "insoluble proteins" with pure water. *FEBS Lett.* 2009;583(6): 953-9.

32. Song J. Why do proteins aggregate? "Intrinsically insoluble proteins" and "dark mediators" revealed by studies on "insoluble proteins" solubilized in pure water. *F1000 Res* 2013; 2:94.

33. Li M, Liu J, Ran X, Fang M, Shi J, Qin H, Goh JM, Song J. Resurrecting abandoned proteins with pure water: CD and NMR studies of protein fragments solubilized in salt-free water. *Biophys J.* 2006; 91(11):4201-9.

34. Li M, Liu J, Song J. Nogo goes in the pure water: solution structure of Nogo-60 and design of the structured and buffer-soluble Nogo-54 for enhancing CNS regeneration. *Protein Sci*. 2006; 15(8):1835-41.

35. Li M, Song J. Nogo-B receptor possesses an intrinsically unstructured ectodomain and a partially folded cytoplasmic domain. *Biochem Biophys Res Commun.* 2007; 360(1):128-34.

36. Liu J, Song J. A novel nucleolar transcriptional activator ApLLP for long-term memory formation is intrinsically unstructured but functionally active. *Biochem Biophys Res Commun.* 2008; 366(2):585-91.

37. Li M, Li Y, Liao X, Liu J, Qin H, Xiao ZC, Song J. Rational design, solution conformation and identification of functional residues of the soluble and structured Nogo-54, which mimics Nogo-66 in inhibiting the CNS neurite outgrowth. *Biochem Biophys Res Commun.* 2008; 373(4):498-503.

38. Liu J, Song J. Insights into protein aggregation by NMR characterization of insoluble SH3 mutants solubilized in salt-free water. *PLoS One.* 2009; 4(11):e7805.

28

39. Shaveta G, Shi J, Chow VT, Song J. Structural characterization reveals that viperin is a radical S-adenosyl-L-methionine (SAM) enzyme. *Biochem Biophys Res Commun.* 2010; 391(3):1390-5.

40. Shi J, Lua S, Tong JS, Song J. Elimination of the native structure and solubility of the hVAPB MSP domain by the Pro56Ser mutation that causes amyotrophic lateral sclerosis. *Biochemistry*. 2010; 49(18):3887-97.

41. Miao L, Qin H, Koehl P, Song J. Selective and specific ion binding on proteins at physiologically-relevant concentrations. *FEBS Lett*. 2011; 585(19):3126-32.

42. Lua S, Qin H, Lim L, Shi J, Gupta G, Song J. Structural, stability, dynamic and binding properties of the ALS-causing T46I mutant of the hVAPB MSP domain as revealed by NMR and MD simulations. *PLoS One.* 2011; 6(11):e27072.

43. Qin H, Lim L, Wei Y, Gupta G, Song J. Resolving the paradox for protein aggregation diseases: NMR structure and dynamics of the membrane-embedded P56S-MSP causing ALS imply a common mechanism for aggregation-prone proteins to attack membranes. Version 2. *F1000Res*. 2013; 2:221.

44. Qin H, Wang W, Song J. ALS-causing P56S mutation and splicing variation on the hVAPB MSP domain transform its β-sandwich fold into lipid-interacting helical conformations. *Biochem Biophys Res Commun.* 2013; 431(3):398-403.

45. Lim L, Lee X, Song J. Mechanism for transforming cytosolic SOD1 into integral membrane proteins of organelles by ALS-causing mutations. *Biochim Biophys Acta.* 2015; 1848(1 Pt A):1-7.

46. Theng SS, Wang W, Mah WC, Chan C, Zhuo J, Gao Y, Qin H, Lim L, Chong SS, Song J, Lee CG. Disruption of FAT10-MAD2 binding inhibits tumor progression. *Proc Natl Acad Sci U S A*. 2014;111(49):E5282-91.

47. Goyal S, Qin H, Lim L, Song J. Insoluble protein characterization by circular dichroism (CD) spectroscopy and nuclear magnetic resonance (NMR). *Methods Mol Biol.* 2015; 1258:371-85.

48. Qin H, Lim LZ, Wei Y, Song J. TDP-43 N terminus encodes a novel ubiquitin-like fold and its unfolded form in equilibrium that can be shifted by binding to ssDNA. *Proc Natl Acad Sci U S A.* 2014; 111(52):18619-24.

49. Gupta G, Lim L, Song J. NMR and MD Studies Reveal That the Isolated Dengue NS3 Protease Is an Intrinsically Disordered Chymotrypsin Fold Which Absolutely Requests NS2B for Correct Folding and Functional Dynamics. *PLoS One.* 2015;10(8):e0134823.

50. Lim L, Wei Y, Lu Y, Song J. ALS-Causing Mutations Significantly Perturb the Self-Assembly and Interaction with Nucleic Acid of the Intrinsically Disordered Prion-Like Domain of TDP-43. *PLoS Biol.* 2016; 14(1):e1002338.

51. Wei Y, Lim L, Wang L, Song J. Inter-domain interactions of TDP-43 as decoded by NMR. *Biochem Biophys Res Commun.* 2016; 473(2):614-9.

52. Lim L, Song J. SALS-linked WT-SOD1 adopts a highly similar helical conformation as FALS-causing L126Z-SOD1 in a membrane environment. *Biochim Biophys Acta.* 2016; 1858(9):2223-30.

53. Kthiri F1, Le HT, Gautier V, Caldas T, Malki A, Landoulsi A, Bohn C, Bouloc P, Richarme G. Protein aggregation in a mutant deficient in YajL, the bacterial homolog of the Parkinsonism-associated protein DJ-1. *J Biol Chem.* 2010; 285(14):10328-36.

54. Le HT, Gautier V, Kthiri F, Malki A, Messaoudi N, Mihoub M, Landoulsi A, An YJ, Cha SS, Richarme G. YajL, prokaryotic homolog of parkinsonism-associated protein DJ-1, functions as a covalent chaperone for thiol proteome. *J Biol Chem.* 2012; 287(8):5861-70.

55. Abdallah J, Mihoub M, Gautier V, Richarme G. The DJ-1 superfamily members YhbO and YajL from Escherichia coli repair proteins from glycation by methylglyoxal and glyoxal. *Biochem Biophys Res Commun*. 2016; 470(2):282-6.

56. Messaoudi N, Gautier V, Kthiri F, Lelandais G, Mihoub M, Joseleau-Petit D, Caldas T, Bohn C, Tolosa L, Rao G, Tao K, Landoulsi A, Bouloc P, Richarme G. Global stress response in a prokaryotic model of DJ-1-associated Parkinsonism. *J Bacteriol.* 2013; 195(6):1167-78.

57. Maisonneuve E, Fraysse L, Moinier D, Dukan S. Existence of abnormal protein aggregates in healthy Escherichia coli cells. *J Bacteriol*. 2008; 190(3):887-93.

58. Aliprandi P, Sizun C, Perez J, Mareuil F, Caputo S, Leroy JL, Odaert B, Laalami S, Uzan M, Bontems F. S1 ribosomal protein functions in translation initiation and ribonuclease RegB activation are mediated by similar RNA-protein interactions: an NMR and SAXS analysis. *J Biol Chem.* 2008; 283(19):13289-301.

59. Byrgazov K, Grishkovskaya I, Arenz S, Coudevylle N, Temmel H, Wilson DN, Djinovic-Carugo K, Moll I. Structural basis for the interaction of protein S1 with the Escherichia coli ribosome. *Nucleic Acids Res.* 2015; 43(1):661-73.

60. Stadtman ER. Protein oxidation and aging. *Science*. 1992; 257(5074):1220-4.

61. Davies MJ. Protein oxidation and peroxidation. *Biochem J*. 2016; 473(7):805-25.

62. Dean RT, Fu S, Stocker R, Davies MJ. Biochemistry and pathology of radical-mediated protein oxidation. *Biochem J*. 1997; 324 (Pt 1):1-18.

63. Aguilaniu H, Gustafsson L, Rigoulet M, Nystrom T. Asymmetric inheritance of oxidatively damaged proteins during cytokinesis. *Science* 2003; 299: 1751–1753.

64. Song J, Jamin N, Gilquin B, Vita C, Ménez A. A gradual disruption of tight side-chain packing: 2D 1H-NMR characterization of acid-induced unfolding of CHABII. *Nat Struct Biol.* 1999; 6(2):129-34.

65. Wei Z, Song J. Molecular mechanism underlying the thermal stability and pH-induced unfolding of CHABII. *J Mol Biol.* 2005; 348(1):205-18.

66. Dyson HJ, Wright PE. Unfolded proteins and protein folding studied by NMR. *Chem. Rev.* 2004; 104: 3607–3622.

67. Marsh JA, Singh VK, Jia Z, Forman-Kay JD. Sensitivity of secondary structure propensities to sequence differences between alpha- and gamma-synuclein: implications for fibrillation. *Protein Sci.* 2006; 15: 2795–2804.

68. Perutz MF, Kendrew, JC, Watson, HC. Structure and function of haemoglobin II. Some relations between polypeptide chain configuration and amino acid sequence. *J Mol Biol* 1965; 13:669–678.

69. Segrest JP, Jackson RL, Morrisett JD, Gotto AM Jr. A molecular theory of lipid–protein interactions in the plasma lipoproteins. *FEBS Lett* 1974; 38:247– 258.

70. Eisenberg D, Weiss RM, Terwilliger TC: The hydrophobic moment detects periodicity in protein hydrophobicity. *Proc Natl Acad Sci U S A.* 1984; 81(1):140–4.

71. Wimley WC, White SH. Experimentally determined hydrophobicity scale for proteins at membrane interfaces. *Nat Struct & Molec Bio.* 1996; 3: 842–848.

72. Fernández-Vidal M, Jayasinghe S, Ladokhin AS, et al.: Folding amphipathic helices into membranes: amphiphilicity trumps hydrophobicity. *J Mol Biol.* 2007; 370(3): 459–70.

73. Drin G, Antonny B. Amphipathic helices and membrane curvature. *FEBS Lett.* 2010; 584(9):1840-7.

74. Sharadadevi A1, Sivakamasundari C, Nagaraj R. Amphipathic alpha-helices in proteins: results from analysis of protein structures. *Proteins.* 2005; 59(4):791-801.

75. Parker W, Sood A, Song A. Organization of regions with amphiphilic alphahelical potential within the three-dimensional structure of beta-sheet proteins. *Protein Eng.* 2001; 14(5): 315–9.

76. Lee SY, Parker W: Amphiphilic α-helical potential: a putative folding motif adding few constraints to protein evolution. *J Mol Evol.* 2011; 73(3–4): 166–80.

77. Kyte J, Doolittle RF: A simple method for displaying the hydropathic character of a protein. *J Mol Biol.* 1982; 157(1): 105–32.

78. Hartl FU. Cellular Homeostasis and Aging. *Annu Rev Biochem.* 2016; 85:1-4.

79. Rujano MA, Bosveld F, Salomons FA, Dijk F, van Waarde MA, et al. Polarised asymmetric inheritance of accumulated protein damage in higher eukaryotes. *PLoS Biol* 2006; 4: e417.

80. Ben-Zvi A, Miller EA, Morimoto RI. Collapse of proteostasis represents an early molecular event in Caenorhabditis elegans aging. *Proc Natl Acad Sci U S A.* 2009; 106: 14914–14919.

81. David DC, Ollikainen N, Trinidad JC, Cary MP, Burlingame AL, Kenyon C. Widespread protein aggregation as an inherent part of aging in C. elegans. *PLoS Biol.* 2010; 8(8):e1000450.

82. Walther DM, Kasturi P, Zheng M, Pinkert S, Vecchi G, Ciryam P, Morimoto RI, Dobson CM, Vendruscolo M, Mann M, Hartl FU. Widespread Proteome Remodeling and Aggregation in Aging C. elegans. *Cell.* 2015; 161(4):919-32.

83. Liu J, Song J. NMR evidence for forming highly populated helical conformations in the partially folded hNck2 SH3 domain. *Biophys J.* 2008; 95(10):4803-12.

84. Tomoyasu T, Mogk A, Langen H, Goloubinoff P, Bukau B. Genetic dissection of the roles of chaperones and proteases in protein folding and degradation in the *Escherichia coli* cytosol. *Mol Microbiol.* 2001; 40(2):397-413.

85. Kerner MJ, Naylor DJ, Ishihama Y, Maier T, Chang HC, Stines AP, Georgopoulos C, Frishman D, Hayer-Hartl M, Mann M, Hartl FU. Proteome-wide analysis of chaperonin-dependent protein folding in *Escherichia coli. Cell.* 2005; 122(2):209-20.

86. Schubert U, Anton LC, Gibbs J, et al: Rapid degradation of a large fraction of newly synthesized proteins by proteasomes. *Nature.* 2000; 404 (6779): 770–774.

87. Duttler S, Pechmann S, Frydman J: Principles of cotranslational ubiquitination and quality control at the ribosome. *Mol Cell.* 2013; 50(3): 379–93.

88. Pace CN, Vajdos F, Fee L, Grimsley G, Gray T. How to measure and predict the molar absorption coefficient of a protein. *Protein Sci.* 1995; 4: 2411–2423.

89. Gupta G, Qin H, Song J. Intrinsically unstructured domain 3 of hepatitis C Virus NS5A forms a "fuzzy complex" with VAPB-MSP domain which carries ALS-causing mutations. *PLoS One.* 2012; 7(6):e39261.

90. Qin H, Lim LZ, Song J. Dynamic principle for designing antagonistic/agonistic molecules for EphA4 receptor, the only known ALS modifier. *ACS Chem Biol.* 2015; 10(2):372-8.

91. Delaglio F, Grzesiek S, Vuister GW, Zhu G, Pfeifer J, Bax A. NMRPipe: A multidimensional spectral processing system based on UNIX pipes. *J Biomol NMR.* 1995; 6: 277–293.

92. Johnson BA, Blevins RA. NMR View: A computer program for the visualization and analysis of NMR data. *J Biomol NMR.* 1994; 4: 603–614.

93. Cornilescu G, Delaglio F, Bax A. Protein backbone angle restraints from searching a database for chemical shift and sequence homology. *J Biomol NMR.* 1999; 13: 289–302.

94. Güntert P. Automated NMR structure calculation with CYANA. *Methods Mol Biol.* 2004; 278: 353–378.