



Published in final edited form as:

*Trends Pharmacol Sci.* 2015 February ; 36(2): 72–77. doi:10.1016/j.tips.2014.12.004.

## Supersaturation is a major driving force for protein aggregation in neurodegenerative diseases

Prajwal Ciryam<sup>1,2</sup>, Rishika Kundra<sup>1</sup>, Richard I. Morimoto<sup>2</sup>, Christopher M. Dobson<sup>1</sup>, and Michele Vendruscolo<sup>1</sup>

<sup>1</sup>Department of Chemistry, University of Cambridge, Cambridge CB2 1EW, UK

<sup>2</sup>Department of Molecular Biosciences, Rice Institute for Biomedical Research, Northwestern University, Evanston IL 60208, USA

### Abstract

The solubility of proteins is an essential requirement for their function. Yet these ubiquitous molecules can undergo aggregation when the protein homeostasis system becomes impaired. Here we ask which is the driving force for protein aggregation in the cellular environment. Emerging evidence suggests that this phenomenon arises because the native states of many proteins are inherently metastable as their cellular concentrations exceed their critical values. Such 'supersaturated' proteins are strongly driven towards aggregation, and are over-represented in specific biochemical pathways associated with neurodegenerative conditions. These observations suggest that effective therapeutic approaches to combat neurodegenerative diseases could be aimed at enhancing the ability of the cell to maintain protein solubility.

### Protein aggregation in neurodegenerative disease

Neurodegenerative disorders, including Alzheimer's, Parkinson's and Huntington's diseases, are increasingly prevalent with the ageing of the modern world, and present great challenges to our healthcare systems as their early diagnosis and effective treatment remain elusive [1–6]. This situation results, at least in part, from our limited understanding of the fundamental nature and origins of such diseases. Several contributing factors have been proposed to explain their onset and progression, including oxidative stress, mitochondrial dysfunction, the disruption of the endoplasmic reticulum and of membrane trafficking, the failure of protein folding and clearance mechanisms, and the activation of inflammatory responses [2, 3, 7, 8]. A common upstream feature of these disorders, however, is that specific peptides and proteins, including A $\beta$  and tau in Alzheimer's disease,  $\alpha$ -synuclein in Parkinson's disease and huntingtin in Huntington's disease, misfold and aggregate to form amyloid assemblies with a characteristic cross- $\beta$  structure [4, 9, 10]. The presence of such aberrant aggregates can generate a cascade of pathological events, leading to the failure of protein homeostasis and the loss of normal biological function [4, 9, 11–14]. While the causal role of aggregation in these neurodegenerative disorders has not been established conclusively in general, this phenomenon is strongly associated with pathogenesis in a wide range of cases.

In this view, one promising avenue for progress in the development of therapeutic approaches for neurodegenerative disorders is to improve our understanding of the mechanisms by which cellular dysfunction arises from the initial protein aggregation events

[4, 9, 11, 12, 15]. It is increasingly evident that, far from being an unusual property of a few specific proteins, protein aggregation is a common phenomenon and that most, if not all, proteins can form amyloid assemblies under appropriate conditions [1]. It has also been proposed that the intrinsic propensity of proteins to aggregate is encoded in their amino acid sequences [16], and a variety of complementary approaches have been developed to predict such aggregation propensities [16–19]. Moreover, misfolded species, particularly the oligomers associated with the formation of larger heterologous assemblies, possess a generic cellular toxicity even under physiological conditions, the degree of which can be rationalized in terms of the physico-chemical properties of the proteins involved [20, 21]. In this article we discuss the fundamental reasons that underlie protein aggregation, and identify supersaturation (Figure 1) as the driving force responsible for this phenomenon, suggesting that the enhancement of the protein homeostasis mechanisms responsible for the maintenance of proteins in their soluble state represents an effective therapeutic route to prevent and combat neurodegenerative disorders.

### **Native states of proteins can be metastable towards aggregation**

Increasing evidence indicates that many proteins are kinetically, but not thermodynamically, stable in their native states even under physiological conditions [22, 23]. Despite the high stability of the aggregated state, however, such proteins can remain soluble for long periods of time because of the presence of high kinetic barriers that separate the native and the aggregated states [23, 24]. In essence, proteins have evolved to remain soluble, but only to the levels required to meet their functions [25]. A series of recent studies has shown that the concentrations of proteins in the cell are close to their critical values [25] and a negative correlation between protein aggregation propensity and protein expression has been observed among groups of bacterial, yeast and human proteins [26, 27]. Exceptions to this rule are represented by functional amyloid proteins, such as Pmel17, which is associated with the production of melanin, whose expression levels are high relative to their aggregation propensities, thus facilitating aggregation under normal conditions *in vivo* [25]. It has also been suggested that the codon usage in highly expressed genes is under pressure to avoid mistranslation-induced protein misfolding [28], so that not only the amino acid sequences, but also the codon choice may be under selection to avoid aggregation-prone sequences.

Given the metastability of native proteins, it is not surprising that protein aggregation has emerged as a much more widespread phenomenon than previously supposed, with hundreds of different proteins shown to aggregate under stress, or as a consequence of ageing and disease [29–39]. These studies suggest the existence of a metastable sub-proteome that is inherently at danger of aggregation, and consequently raise the question of why some proteins are more at risk of aggregation than others. We believe that this question may be the key to determining how specific misfolding events in the cell evolve into a broad collapse of protein homeostasis [4].

## Supersaturated proteins are driven towards aggregation

To understand why some proteins aggregate under specific stress conditions while others remain soluble, we directed our attention towards proteins that are less soluble than one would expect given their expression levels [40]. We refer to such proteins as being supersaturated [40]. The concept of supersaturation has a long history in the physical and biological sciences, as it was for example used over 40 years ago to elucidate the interplay between kinetic and thermodynamic factors determining the aggregation of sickle hemoglobin [41, 42], and more recently to characterise protein aggregate formation [43–45]. While supersaturation and solubility are thermodynamic concepts, they also have implications for the kinetics of aggregation, as when the concentration above the critical value becomes greater, the driving force towards aggregation is likely to become stronger (Figure 1).

We recently tested the consequences of the metastability of at least a part of the proteome against aggregation and its implications for the collapse of protein homeostasis by analyzing the supersaturation levels of thousands of proteins. To obtain these results we estimated the cellular concentrations of proteins from transcriptomic and proteomic data and the critical concentrations from the prediction of the aggregation propensity from the physicochemical properties of the corresponding amino acid sequences [40]. We found that a substantial fraction of proteins in the proteome is intrinsically supersaturated, and thus at risk of aggregation even in the absence of destabilizing amino acid mutations or increased expression. We also found that a wide range of pathological consequences associated with neurodegenerative diseases can be rationalized to a substantial degree on the basis of the supersaturation levels of these metastable proteins [40]. The specific proteins that have been found to co-aggregate with A $\beta$  plaques and tau tangles in Alzheimer's diseases, with  $\alpha$ -synuclein Lewy bodies in Parkinson's disease, and with artificial amyloid-forming  $\beta$  constructs are substantially supersaturated (Figure 2). In addition, we observed that proteins which preferentially aggregate during ageing in the nematode *C. elegans* are also highly supersaturated [40]. A recent proteomic study in yeast similarly found that arsenite stress results in the aggregation of high concentration, poorly soluble proteins [46]. We thus propose that supersaturation is a major driving force of aggregation across the proteome, and a central aspect of understanding the widespread cellular dysfunction evident in many conformational disorders.

## Pathways affected in neurodegenerative diseases are enriched in supersaturated proteins

Neurodegenerative diseases are associated with dysfunction in one or more biochemical pathways [4, 9, 11, 12, 15]. In the case of Alzheimer's disease, according to the 'amyloid cascade hypothesis' [9], the aggregation of the A $\beta$  peptide triggers a series of downstream pathological events (Figure 3), ranging from the formation of neurofibrillary tangles to the disruption of lipid membranes and to oxidative stress. The list of these secondary consequences is large, as Alzheimer's disease is multifactorial both in terms of its genetic risk factors and its physiological effects in the cell [2, 3, 9]. This bewildering complexity may explain why so many distinct mechanisms have been proposed for disease pathology,

including mitochondrial dysfunction, trafficking failure, loss of protein homeostasis and inflammation [4, 9, 11, 12, 15].

Two complementary possibilities could explain this collapse of cellular function. The first is that specific disease-related aggregating peptides and proteins (e.g. A $\beta$  and tau in Alzheimer's disease,  $\alpha$ -synuclein in Parkinson's, and huntingtin in Huntington's disease) specifically disrupt numerous cellular processes. The second possibility is that the aggregation of specific proteins saturates the capacity of the protein quality control system, making it difficult for the cell to maintain the solubility of the proteome as a whole. These observations raise the intriguing possibility that the amyloid cascade may originate from intrinsic vulnerabilities in the proteome and their consequences on the quality control machinery. Recent investigations of the events that trigger aggregation in supersaturated protein solutions showed that sonication, TFE exposure, and HFIP exposure are each able to perturb the metastable state and induce aggregation [47, 48]. These *in vitro* studies suggest that also in the disease state specific perturbations of protein homeostasis may potentiate broader disruptions in the maintenance of a fully functional proteome. As mechanical or chemical stress may reduce *in vitro* the kinetic barrier to aggregation, so can age-related stress *in vivo*. Indeed, it was recently shown that, during ageing, a specific subset of molecular chaperones able to prevent protein aggregation decrease in expression [49].

Our investigations of the properties of the pathways populated by proteins predicted to be supersaturated support such an interpretation. We carried out an unbiased search of the hundreds of biochemical pathways assembled in the Kyoto Encyclopedia of Genes and Genomes (KEGG) for those that are enriched in the most supersaturated proteins. We found that the 'KEGG pathways' (i.e. as defined in KEGG) for Alzheimer's, Parkinson's, and Huntington's diseases are highly and specifically populated with supersaturated proteins (Figure 3) [40]. As the proteins common to these pathways are particularly supersaturated, we referred to them as the core metastable sub-proteome [40]. We can also speculate that supersaturated proteins cluster in given pathways because they function together, and hence their expression levels are correlated [50, 51] as their expression are often controlled by the same transcription factors [52]. While these findings by no means reduce the importance of the toxic events associated with the interactions of specific aggregated species with cellular components, they also suggest that fundamental physicochemical principles contribute strongly to the widespread aggregation and cellular dysfunction that follows the onset of neurodegenerative diseases [40].

The presence of supersaturated proteins in the cell may follow from the requirement to carry out efficiently specific cellular functions. Protein complexes offer intriguing examples that illustrate this possibility. Many of the pathways that we found to be supersaturated involve large protein complexes, such as the ribosome and components of the electron transport chain. In order for a complex to form, its constituent proteins have been observed to possess surfaces that tend to be aggregation-prone [53]. If their stoichiometry and assembly process are disrupted, therefore, aggregation can occur, as seen for instance for components of the ribosome [33].

## Maintaining the homeostasis of the metastable sub-proteome

The metastability of a substantial part of the proteome with respect to aggregation suggests that diagnostic and therapeutic approaches aimed at limiting this phenomenon could look beyond individual disease-related proteins. Such a shift in perspective is timely, as efforts to target the aggregation of individual proteins have proven less fruitful than anticipated [54]. While an interpretation of these disappointments is to question the amyloid hypothesis entirely, it has also been pointed out that many of the trials of aggregation-reduction therapies had substantial limitations because of the difference between attempting to reduce pre-existing aggregate levels and preventing aggregation in the first place [55]. Timing is essential, as without an effective molecular marker for early diagnoses of sporadic forms of neurodegenerative disorders such as Alzheimer's and Parkinson's diseases, treatments so far could have been initiated too late.

It is therefore of great importance to understand the mechanisms that help maintain the metastable sub-proteome in healthy cells and organisms. Although altered transcriptional levels, inherited polymorphisms, and environmental disturbances may push proteins towards supersaturation, the protein homeostasis machinery can still be capable of restoring the balance (Figure 4). It has been for example proposed that the molecular chaperone Hsp90 can facilitate evolution by permitting a broader search of amino acid sequence space by reducing the cost of destabilizing mutations that might otherwise excessively increase protein supersaturation levels [56]. Indeed, numerous studies have demonstrated that molecular chaperones can suppress protein aggregation [57–59], essentially reducing the effective supersaturation level of proteins below that predicted from their physicochemical properties. Recently, it was shown that Hsp90 helps to regulate protein homeostasis in a cell non-autonomous fashion, revealing another means by which molecular chaperones can act at an organismal level to suppress physicochemical instabilities [60].

## Looking to the future

The presence of proteins near or above their solubility limits is inherently dangerous for living systems as the uncontrolled formation of aggregates can give rise to a variety of pathological events. Such events may then be aggravated by the co-regulation of aggregation-prone proteins, which tend to cluster into specific biochemical pathways. There have been recent advances in our understanding of the nature and properties of the conformational species populated during the aggregation process and of the roles in this context of molecular chaperones and of the protein expression and degradation pathways. We believe that the study of protein supersaturation as a major driving force of aggregation will help bring together these advances and clarify how the failure to maintain proteins in their functional soluble states can give rise to the multifactorial pathology of neurodegenerative conditions. As the failure to control protein solubility is likely to be a major contributor of these complex diseases, novel types of therapeutic strategies could be aimed not only at their treatment, but also at their prevention, by reducing the propensity of these proteins to aggregate and trigger loss of protein homeostasis, and also by enhancing through chemical intervention the quality control processes that regulate the response to protein aggregation.

## References

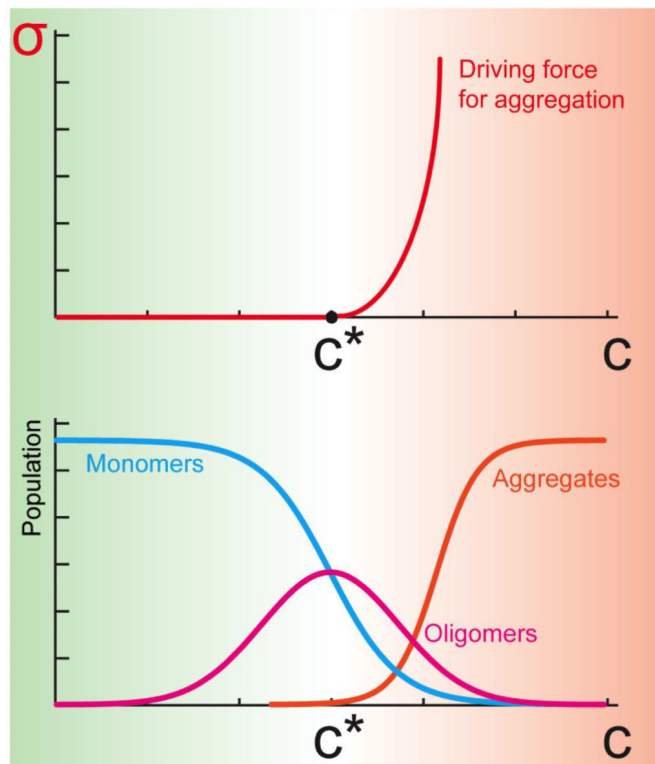
1. Dobson CM. Protein folding and misfolding. *Nature*. 2003; 426:884–890. [PubMed: 14685248]
2. Querfurth HW, LaFerla FM. Mechanisms of disease: Alzheimer's disease. *N. Engl. J. Med.* 2010; 362:329–344. [PubMed: 20107219]
3. Selkoe D, et al. Deciphering Alzheimer disease. *Cold Spring Harb. Perspect. Med.* 2012; 2:a011460. [PubMed: 22315723]
4. Knowles TP, et al. The amyloid state and its association with protein misfolding diseases. *Nat Rev. Mol. Cell Biol.* 2014; 15:384–396. [PubMed: 24854788]
5. Alzheimer's Research Trust. 2010. Dementia 2010: The prevalence, economic cost and research funding compared with other major diseases.
6. Alzheimer's Association. 2012. 2012 Alzheimer's disease: Facts and figures.
7. Rubinsztein DC. The roles of intracellular protein-degradation pathways in neurodegeneration. *Nature*. 2006; 443:780–786. [PubMed: 17051204]
8. Glass CK, et al. Mechanisms underlying inflammation in neurodegeneration. *Cell*. 2010; 140:918–934. [PubMed: 20303880]
9. Hardy J, Selkoe DJ. Medicine - The amyloid hypothesis of Alzheimer's disease: Progress and problems on the road to therapeutics. *Science*. 2002; 297:353–356. [PubMed: 12130773]
10. Eisenberg D, Jucker M. The amyloid state of proteins in human diseases. *Cell*. 2012; 148:1188–1203. [PubMed: 22424229]
11. Balch WE, et al. Adapting proteostasis for disease intervention. *Science*. 2008; 319:916–919. [PubMed: 18276881]
12. Sontag EM, et al. Sorting out the trash: The spatial nature of eukaryotic protein quality control. *Curr. Opin. Cell Biol.* 2014; 26:139–146. [PubMed: 24463332]
13. Hipp MS, et al. Proteostasis impairment in protein-misfolding and-aggregation diseases. *Trends Cell Biol.* 2014
14. Yu A, et al. Protein aggregation can inhibit clathrin-mediated endocytosis by chaperone competition. *Proc. Natl. Acad. Sci. USA*. 2014; 111:E1481–E1490. [PubMed: 24706768]
15. Haass C, Selkoe DJ. Soluble protein oligomers in neurodegeneration: Lessons from the Alzheimer's amyloid  $\beta$ -peptide. *Nat Rev. Mol. Cell Biol.* 2007; 8:101–112. [PubMed: 17245412]
16. Chiti F, et al. Rationalization of the effects of mutations on peptide and protein aggregation rates. *Nature*. 2003; 424:805–808. [PubMed: 12917692]
17. Tartaglia GG, et al. Prediction of aggregation-prone regions in structured proteins. *J. Mol. Biol.* 2008; 380:425–436. [PubMed: 18514226]
18. Fernandez-Escamilla A-M, et al. Prediction of sequence-dependent and mutational effects on the aggregation of peptides and proteins. *Nat. Biotechnol.* 2004; 22:1302–1306. [PubMed: 15361882]
19. Belli M, et al. Prediction of amyloid aggregation in vivo. *EMBO reports*. 2011; 12:657–663. [PubMed: 21681200]
20. Luheshi LM, et al. Systematic in vivo analysis of the intrinsic determinants of amyloid  $\beta$  pathogenicity. *PLoS Biol.* 2007; 5:e290. [PubMed: 17973577]
21. Brorsson A-C, et al. Intrinsic determinants of neurotoxic aggregate formation by the amyloid  $\beta$  peptide. *Biophys. J.* 2010; 98:1677–1684. [PubMed: 20409489]
22. Gazit E. The “correctly folded” state of proteins: Is it a metastable state? *Angew. Chem. Int. Ed.* 2002; 41:257–259.
23. Baldwin AJ, et al. Metastability of native proteins and the phenomenon of amyloid formation. *J. Am. Chem. Soc.* 2011; 133:14160–14163. [PubMed: 21650202]
24. Buell AK, et al. Detailed analysis of the energy barriers for amyloid fibril growth. *Angew. Chem. Int. Ed.* 2012; 51:5247–5251.
25. Tartaglia GG, et al. Life on the edge: A link between gene expression levels and aggregation rates of human proteins. *Trends Biochem. Sci.* 2007; 32:204–206. [PubMed: 17419062]
26. Tartaglia GG, Vendruscolo M. Correlation between mRNA expression levels and protein aggregation propensities in subcellular localisations. *Mol. BioSyst.* 2009; 5:1873–1876.



27. Gsponer J, Babu MM. Cellular strategies for regulating functional and nonfunctional protein aggregation. *Cell Rep.* 2012; 2:1425–1437. [PubMed: 23168257]
28. Drummond DA, Wilke CO. Mistranslation-induced protein misfolding as a dominant constraint on coding-sequence evolution. *Cell.* 2008; 134:341–352. [PubMed: 18662548]
29. Chapman E, et al. Global aggregation of newly translated proteins in an escherichia coli strain deficient of the chaperonin groel. *Proc. Natl. Acad. Sci. USA.* 2006; 103:15800–15805. [PubMed: 17043235]
30. David DC, et al. Widespread protein aggregation as an inherent part of aging in *C. elegans*. *PLoS Biol.* 2010; 8:e1000450. [PubMed: 20711477]
31. Gidalevitz T, et al. Progressive disruption of cellular protein folding in models of polyglutamine diseases. *Science.* 2006; 311:1471–1474. [PubMed: 16469881]
32. Koga H, et al. Protein homeostasis and aging: The importance of exquisite quality control. *Ageing Res. Rev.* 2011; 10:205–215. [PubMed: 20152936]
33. Koplín A, et al. A dual function for chaperones ssb-rac and the nac nascent polypeptide-associated complex on ribosomes. *J. Cell Biol.* 2010; 189:57–68. [PubMed: 20368618]
34. Luján L, et al. Proteomic characterization of postmortem amyloid plaques isolated by laser capture microdissection. *J. Biol. Chem.* 2004; 279:37061–37068. [PubMed: 15220353]
35. Narayanaswamy R, et al. Widespread reorganization of metabolic enzymes into reversible assemblies upon nutrient starvation. *Proc. Natl. Acad. Sci. USA.* 2009; 106:10147–10152. [PubMed: 19502427]
36. Olzscha H, et al. Amyloid-like aggregates sequester numerous metastable proteins with essential cellular functions. *Cell.* 2011; 144:67–78. [PubMed: 21215370]
37. Reis-Rodrigues P, et al. Proteomic analysis of age-dependent changes in protein solubility identifies genes that modulate lifespan. *Ageing Cell.* 2012; 11:120–127. [PubMed: 22103665]
38. Wang Q, et al. Proteomic analysis of neurofibrillary tangles in Alzheimer disease identifies gapdh as a detergent-insoluble paired helical filament tau binding protein. *FASEB J.* 2005; 19:869–871. [PubMed: 15746184]
39. Xia Q, et al. Proteomic identification of novel proteins associated with Lewy bodies. *Front. Biosci.* 2008; 13:3850–3856. [PubMed: 18508479]
40. Ciryam P, et al. Widespread aggregation and neurodegenerative diseases are associated with supersaturated proteins. *Cell Rep.* 2013; 5:781–790. [PubMed: 24183671]
41. Eaton WA, Hofrichter J. Hemoglobin-S gelation and sickle-cell disease. *Blood.* 1987; 70:1245–1266. [PubMed: 3311198]
42. Hofrichter J, et al. Supersaturation in sickle cell hemoglobin solutions. *Proc. Natl. Acad. Sci. USA.* 1976; 73:3035–3039. [PubMed: 9640]
43. Harper JD, Lansbury PT Jr. Models of amyloid seeding in Alzheimer's disease and scrapie: Mechanistic truths and physiological consequences of the time-dependent solubility of amyloid proteins. *Annu. Rev. Biochem.* 1997; 66:385–407. [PubMed: 9242912]
44. Ikenoue T, et al. Heat of supersaturation-limited amyloid burst directly monitored by isothermal titration calorimetry. *Proc. Natl. Acad. Sci. USA.* 2014; 111:6654–6659. [PubMed: 24753579]
45. Crick SL, et al. Unmasking the roles of N- and C-terminal flanking sequences from exon 1 of huntingtin as modulators of polyglutamine aggregation. *Proc. Natl. Acad. Sci. USA.* 2013; 110:20075–20080. [PubMed: 24282292]
46. Ibstedt, S., et al. *Biology Open.* 2014. Global analysis of protein aggregation in yeast during physiological conditions and arsenite stress. BIO20148938
47. Muta H, et al. Supersaturation-limited amyloid fibrillation of insulin revealed by ultrasonication. *J. Biol. Chem.* 2014; 289:18228–18238. [PubMed: 24847058]
48. Umemoto A, et al. High-throughput analysis of the ultrasonication-forced amyloid fibrillation reveals the mechanism underlying the large fluctuation in the lag time. *J. Biol. Chem.* 2014; M114:569814. jbc.
49. Brehme M. *Cell Rep.* 2014 in press.
50. Eisen MB, et al. Cluster analysis and display of genome-wide expression patterns. *Proc. Natl. Acad. Sci. USA.* 1998; 95:14863–14868. [PubMed: 9843981]

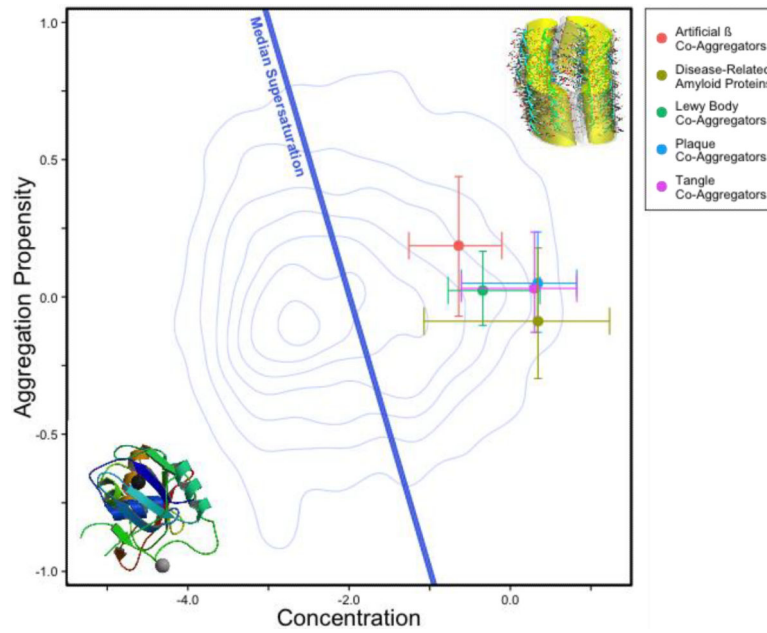
51. Wilhelm BG, et al. Composition of isolated synaptic boutons reveals the amounts of vesicle trafficking proteins. *Science*. 2014; 344:1023–1028. [PubMed: 24876496]
52. Tavazoie S, et al. Systematic determination of genetic network architecture. *Nat. Genet.* 1999; 22:281–285. [PubMed: 10391217]
53. Pechmann S, et al. Physicochemical principles that regulate the competition between functional and dysfunctional association of proteins. *Proc. Natl. Acad. Sci. USA*. 2009; 106:10159–10164. [PubMed: 19502422]
54. Doody RS, et al. Phase 3 trials of solanezumab for mild-to-moderate Alzheimer's disease. *N. Engl. J. Med.* 2014; 370:311–321. [PubMed: 24450890]
55. Karran E, Hardy J. A critique of the drug discovery and phase 3 clinical programmes targeting the amyloid hypothesis for Alzheimer's disease. *Ann. Neurol.* 2014
56. Rutherford SL, Lindquist S. Hsp90 as a capacitor for morphological evolution. *Nature*. 1998; 396:336–342. [PubMed: 9845070]
57. Gragerov A, et al. Cooperation of GroEL/GroES and DnaK/DnaJ heat shock proteins in preventing protein misfolding in *Escherichia coli*. *Proc. Natl. Acad. Sci. USA*. 1992; 89:10341–10344. [PubMed: 1359538]
58. Wiech H, et al. Hsp90 chaperones protein folding in vitro. *Nature*. 1992; 358:169–170. [PubMed: 1614549]
59. Tam S, et al. The chaperonin tric controls polyglutamine aggregation and toxicity through subunit-specific interactions. *Nat. Cell Biol.* 2006; 8:1155–1162. [PubMed: 16980959]
60. van Oosten-Hawle P, et al. Regulation of organismal proteostasis by transcellular chaperone signaling. *Cell*. 2013; 153:1366–1378. [PubMed: 23746847]





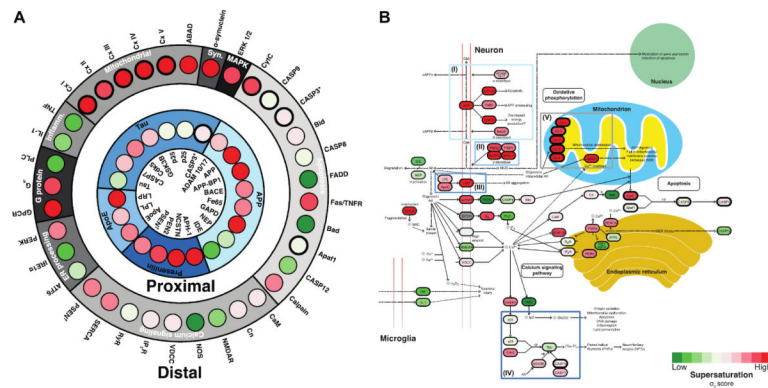
**Figure 1. Supersaturation is a major driving force for protein aggregation**

The supersaturation,  $\sigma$ , which can be defined as the ratio between the cellular concentration ( $C$ ) and the critical concentration ( $C^*$ ) of a protein, represents a driving force for aggregation. Below the critical concentration most proteins are in the monomeric state, with a small fraction in oligomeric states. Above the critical concentration, proteins begin to form insoluble aggregates, including amyloid fibrils, with some remaining fractions in soluble monomeric and oligomeric forms. Thermodynamically, the supersaturation represents the driving force for aggregation associated with the imbalance in the number of molecules in the aggregated and soluble states, and it grows as  $\sigma = \exp(\beta \mu)$ , where  $\mu$  is the difference of the chemical potentials of the aggregated and soluble states, and  $\beta = 1/k_B T$ .



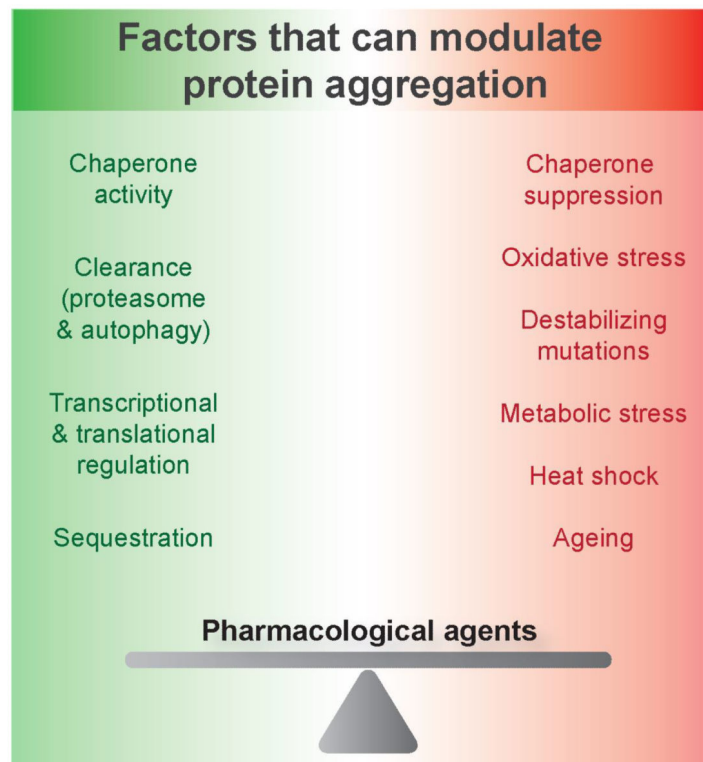
**Figure 2. Protein supersaturation underlies widespread protein aggregation**

The supersaturation scores for disease-related amyloid proteins and proteins that co-aggregate with several types of protein deposits (coloured circles) are above the median supersaturation line (blue line), which divides proteins whose supersaturation is high relative to that of the proteome from those whose supersaturation is low [40]. Error bars represent the 25<sup>th</sup> and 75<sup>th</sup> percentiles of concentration and aggregation propensity. Details of the calculation of the supersaturation score and the list of supersaturated proteins have been reported previously [40].



**Figure 3. Supersaturated proteins in Alzheimer's disease**

(A) Organisation of the proteins involved in the primary (proximal) and secondary (distal) events in the amyloid cascade hypothesis [9]. According to this hypothesis the initial aggregation of A $\beta$  triggers a series of proximal events (inner ring, blue scale), which is then followed by another series of distal events involving proteins in the same networks (outer ring, grey scale). Supersaturation scores are scaled from green (low scores) to red (high scores). (B) Supersaturated proteins (in red) in the KEGG 'Alzheimer's disease pathway'. Some of these pathways (blue boxes), such as the processing of APP (I), the presenilins (II), apolipoprotein E (III) and tau (IV) are known to be linked to genetic perturbations that drive the disease (inner ring in panel A). In addition to these factors, which are proximal to the primary cause of cellular toxicity, other major pathways, such as oxidative phosphorylation (V, grey box) and endoplasmic reticulum function are distal to the causes. Supersaturated proteins populate both those processes that are tied to the primary aggregation events (I–IV) and those that are likely the result of it (V).



**Figure 4. List of representative factors that can modulate protein aggregation**

Since the concentrations of proteins in the cellular environment should be maintained below the critical levels, a wide range of quality control mechanisms is in place for this purpose (green side). A variety of events and processes, however, could shift the balance towards more aggregation-prone conditions (red side). A general goal of pharmacological interventions against neurodegenerative diseases should be a reduction of the driving force for protein aggregation by suppressing protein supersaturation.