Protein aggregation: A perspective from amyloid and inclusion-body formation

Susan Idicula-Thomas and Petety V. Balaji*

School of Biosciences and Bioengineering, Indian Institute of Technology Bombay, Powai, Mumbai 400 076, India

Over the past few decades, an overwhelmingly vast amount of research has been dedicated to understanding protein aggregation. This review summarizes the current understanding of protein aggregation from the viewpoint of its two important manifestations: formation of amyloid fibrils and inclusion bodies. The article summarizes the structure, mechanism of formation, predisposing factors and measures of overcoming inclusion body and amyloid formation. The protective role played by molecular evolution in curbing aggregation and results from recent studies on the prediction of aggregation rates based on primary structure have also been discussed.

Keywords: Folding intermediates, molecular evolution, natively unfolded, protein folding.

PROTEIN aggregation is a phenomenon wherein the protein loses its native structure and adopts a non-native conformation leading to aggregation. There are three manifestations of aggregation: fibrils, amorphous aggregates and soluble oligomers[1]. Aggregates are also commonly classified as ordered or amorphous, based on their macroscopic morphology[2]. Aggregation differs from precipitation (e.g. salting out, isoelectric precipitation, crystallization) in that the aggregates exhibit partial to total loss of native structure of the aggregating protein, whereas the native structure of the protein remains unaffected in the precipitates[3].

Aggregation is thought to be a reversible process of self-association of several identical protein molecules driven by stereospecific intermolecular contacts[4]. Aggregates are enriched with a specific protein. Nevertheless, they have a range of other components associated with them. Proteolytic fragments of the aggregating protein and/or other aggregation-prone proteins are generally the heterogeneous components associated with inclusion bodies. They may also include contaminants arising from the process of purification[3]. Glycosaminoglycans, proteoglycans, apolipoprotein E and serum amyloid P component are the non-fibrillar components that are generally associated with amyloid fibrils. It is believed that these components aid in the formation and stability of the fibrils[5].

The native conformation of the aggregating protein could belong to any SCOP (Structural Classification of Proteins) class[6]: all-α (e.g. myoglobin, cytochrome c552, methionine aminopeptidase), all-β (transthyretin, β2-microglobulin, immunoglobulin light chain variable domains), $\alpha+\beta$ (lysozyme, α-lactalbumin) or α/β (cystic fibrosis transmembrane regulator). In some cases, the protein does not have a well-defined conformation and is natively unfolded (Aβ fragments, IAPP (Islet Amyloid Precursor Polypeptide), calcitonin, α-synuclein, etc.; Table 1).

Studies on protein aggregation have gained significant momentum in the last decade due to the discovery of several debilitating human disorders associated with protein aggregation (Table 1). The aggregates in these 'protein aggregation disorders' or 'protein conformational diseases' may perturb the health of the individual by either loss of biological activity and/or gain of toxicity[7–9]. Based on the type and localization of aggregates, protein aggregation disorders are classified (Figure 1) as (a) extracellular aggregation disorders or amyloidoses caused by amyloid fibril formation, e.g. Alzheimer's, type-II diabetes mellitus, and the prion encephalopathies[10], and (b) intracellular aggregation disorders caused by inclusion bodies, e.g. cataract and Huntington's[7]. This classification is however not stringent since the neurodegenerative disorders could have both intracellular and extracellular aggregates associated with them[11].

This review aims to summarize the knowledge gained in the area of protein aggregation with special emphasis on inclusion body and amyloid formation. Comparisons of these two manifestations of protein aggregation with regard to their structure, formation, predisposing and control factors are discussed. The review also highlights the importance of the evolutionary adaptations in living systems towards restricting aggregation.

Inclusion bodies

Escherichia coli is the preferred host for production of recombinant proteins due to the ease of genetic manipulation and low cost of production. Overexpression of proteins in *E. coli* usually results in the accumulation of the overexpressed protein into inclusion bodies[12]. Proteins that (a) have exposed hydrophobic patches, (b) undergo post-translational modifications, or (c) have disulphide bonds are believed to be at an increased risk to form such amorphous aggregates[13,14]. A majority of overexpressed proteins,

*For correspondence. (e-mail: balaji@iitb.ac.in)

Table 1. Proteins involved in some of the protein aggregation disorders

Disease	Protein	Native structure*	Modification	Manifestation	Reference
Alzheimer's	Amyloid β-peptide Tau	NU NU	Aberrant processing Aberrant Tau phosphorylation	Extracellular amyloid plaques Intracellular neurofibrillary tangles and neurodegeneration	98
Parkinson's	α-Synuclein	NU	Aberrant ubiquination (mutations in Parkin gene)	Cytoplasmic Lewy bodies Neurodegeneration	99
Scrapie/Creutzfeld-Jakob's disease	Prion	NU/ α-helical	Conformational conversion of PrP (C) to infectious PrP (SC)	Amyloid fibrils in CNS Neurodegeneration	100
Cystic fibrosis	Cystic fibrosis *trans*-membrane regulator	α/β	Aberrant folding in the mutant (ΔF508)	Loss of activity	101
Phenylketonuria	Phenylalanine hydroxylase	$\alpha + \beta$	Mutants with defective folding followed by degradation/ aggregation	Loss of activity Metabolic disorder	102
Huntington's	Huntingtin	α-helical/ NU	Poly-glutamine-mediated protein aggregation	Intranuclear inclusions in neurons Neurodegeneration	91
Marfan syndrome	Fibrillin	Small proteins	Mutants have aberrant domain folding	Defective assembly into microfibrils Connective tissue disease	103
Osteogenesis imperfecta	Procollagen	β-sheet	Mutants have structural defects	Bone fragility. Connective tissue disease	104
Sickle cell anaemia	Haemoglobin	α-helical	Substitution of Glu to Val	Intracellular polymerization of mutant Hb on deoxygenation	105
Tay-Sachs disease	β-Hexosaminidase	$\alpha + \beta$	Misfolding of mutant (HexA)	Loss of activity. Neurodegeneration	106
Amyloidoses	Transthyretin Lysozyme Apolipoprotein A1 Gelsolin	β-sheet $\alpha + \beta$ α-helical $\alpha + \beta$	Mutations that reduce stability of native structure	Extracellular fibril deposits	5, 107
Diabetes (type II)	Amylin	NU	Structural instability	Pancreatic amyloid plaques	108
α_1-Antitrypsin deficiency	α_1-Antitrypsin	Multi-domain	Mutations cause polymerization	Inclusion bodies in the ER of hepatocytes	109
Hemodialysis related	β_2-microglobulin	β-sheet	Self-association at elevated concentration	Amyloid plaques in joints	110
Cerebral amyloid angiopathy	Cystatin C	$\alpha + \beta$	Mutations (L68Q) cause polymerization	Amyloid deposits in brain arteries	111
Retinitis pigmentosa	Rhodopsin	–	Mutants exhibit misfolding and aggregation	Intracellular inclusion bodies formed	112
Cataract	Crystallins	β-sheet	Mutants exhibit misfolding and aggregation	Intranuclear inclusions	113

*NU, Natively unfolded protein.

heterologous or native, are thought to form inclusion bodies if their concentration exceeds 2% of the total cellular protein concentration, irrespective of other factors[15].

Structure of the inclusion bodies

Inclusion bodies are dense particles of aggregated protein. The diameter of the spherical, bacterial inclusion bodies can range from 0.5 to 1.3 µm. In spite of their high density (1.3 mg/ml), inclusion bodies are known to be porous and hydrated[15].

Studies on the structure and morphology of inclusion bodies formed from various proteins have revealed that these aggregates can exhibit amorphous as well as partly ordered structure[16]. The presence of 'native-like' structure and retention of functional activity in inclusion bodies of proteins such as interleukin-1β[17], endoglucanase D[18], β-galactosidase[19] and dihydrofolate reductase[20] has lent evidence to the possibility of inclusion bodies of certain proteins preserving a certain degree of ordered structure and not being completely amorphous. The nature of the aggregating protein, predisposing conditions for aggregate formation and the site of inclusion body formation contribute to the structural heterogeneity observed in the inclusion bodies. The inclusion bodies of β-lactamase formed in the periplasmic space were found to be amorphous compared to the paracrystalline structure of the same

enzyme formed in the cytoplasm. Inclusion bodies from the two sites were also different in terms of protein composition, degree of solubilization in the presence of denaturants and sensitivity to protease digestion. These differences have been attributed to difference in the conformation of the associated polypeptide chains[21].

An increase in the content of non-native β-strands has been observed in the inclusion body of β-lactamase[22]. Inter-molecular β-sheets, a hallmark of amyloid fibrils, has also been observed in the inclusion bodies of proteins such as dihydrofolate reductase[20] and a variant of β-galactosidase[23]. Some of the bacterial inclusion bodies are known to bind to amyloid diagnostic dyes, Congo red and thioflavin-T. These observations suggest that the inclusion bodies may possess a varied degree of order in their architecture.

Advantages and disadvantages of inclusion body formation

Considering that inclusion bodies are repositories of the overexpressed proteins, which, at times are greater than 50% of the total cellular proteins, overexpression of proteins into inclusion bodies makes their isolation and purification easy. Inclusion bodies also protect the proteins from proteolysis and/or other degradation pathways[13]. If the recombinant protein is toxic to the host cell, overexpression into inclusion bodies is desirable for large-scale production since the non-functional nature of the inclusion bodies would protect the host[24]. Despite these advantages, inclusion-body formation is often regarded as a bottleneck that has lowered the spectrum of recombinant proteins that could be conveniently overexpressed in prokaryotic expression systems such as *E. coli* and *Bacillus* sp.[25]. This is so because conditions for refolding the aggregated protein to its native, functional form has to be standardized for each protein separately by trial-and-error procedures; this is often both time-consuming and expensive[26].

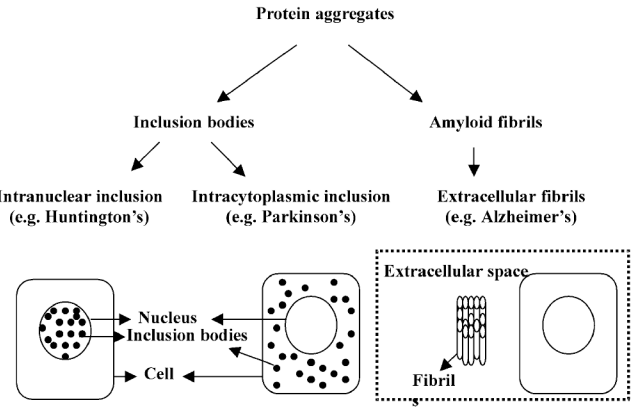

Figure 1. Cartoon representation of types of protein aggregates involved in various protein aggregation disorders. Both intranuclear and cytoplasmic inclusions are known to be associated with Huntington's disease[27]. Although amyloid fibrils are generally seen to occupy the extracellular space, there have been reports of intracellular amyloid fibrils as well[114,115].

Inclusion bodies NOT associated with overexpression of the recombinant protein

Inclusion bodies are often associated with the overexpression of a recombinant protein in a heterologous host such as *E. coli*. However, inclusion bodies also form naturally: because of mutations or inappropriate post-translational modifications which disturb the native conformation of the protein, intracellular inclusion bodies are formed leading to certain neurodegenerative disorders[27]. Such intracellular inclusion bodies may be present in the cytoplasm, nucleus, oligodendrocytes or neurocytes. Based on their location and the protein involved, they are classified as Lewy bodies, Lewy neurites, NFT (neurofibrillary tangles), Pick's bodies, Hirano bodies, Collins bodies, Glial inclusions, etc. Each of these intracellular aggregates has characteristic structure and morphology[27].

Amyloid formation

To date, only a few proteins (~20) are known to form amyloid fibrils under physiological conditions[28]. This is in contrast to inclusion bodies, which are formed under physiological conditions by a large number of proteins when expressed at high levels[4]. However, almost any protein could be induced to form amyloid fibrils *in vitro* under destabilizing environmental conditions. This suggests that amyloid formation is a generic property of all proteins[29].

Structure of the amyloid fibril

Knowledge of the structure of amyloid fibrils is important for understanding their mechanism of formation and stability, related pathogenesis and rational design of drugs to inhibit their formation. The structure of amyloid fibrils has been deciphered using chemical staining methods, hydrogen/deuterium exchange techniques, electron and atomic force microscopies, X-ray diffraction and spectroscopic methods such as solid-state nuclear magnetic resonance (NMR), circular dichroism (CD), Fourier transformed infrared (FTIR), and electron paramagnetic resonance (EPR). All amyloid fibrils, on staining with the diazo dye Congo red, exhibit apple-green birefringence when viewed under a microscope using polarized light. A shift in fluorescence after staining with thioflavin-T is also observed. The amyloid fibrils are viewed as straight, unbranched uniform fibrils of 7–12 nm diameter and of indeterminate length under electron microscopy. X-ray diffraction pattern suggests a cross-β structure wherein ordered β-pleated sheets (separated by ~ 0.98 nm) and the constituent β-strands (separated by

~ 0.47 nm) propagate parallel and perpendicular to the long axis of the fibre, respectively. Spectroscopic studies also have revealed a high β-sheet content for amyloid fibrils[30,31].

Several plausible extensions/modifications have been proposed for the cross-β spine, which is established as the common molecular core for amyloid fibrils[31]. The NMR-derived fibril structure of Alzheimer's β1–42 peptide revealed that the residues 18–42 of each peptide adopt a strand–turn–strand (β–t–β) motif, which assembles into two intermolecular parallel β-sheets that constitute the fibril[32]. Yet another modification of the cross-β structure is the β-helical structure, wherein one or more extended β-sheets wrap in a helical manner around a hollow tube[30]. Support for such β-helical structures comes from the observation that β-helices retain the cross-β orientation[33] and are found in the native structures of proteins such as pectate lyase C wherein multiple β-strands (with 4–6 residues per strand) separated by a short bend segment of 1–2 residues are present[34].

The fraction of total residues that gets incorporated in the core structure varies substantially for amyloidogenic proteins. The manner in which all the residues of a protein are accommodated in the amyloid fibril is not clear. It is possible that only a fragment of the amyloidogenic protein unfolds to contribute a strand to the growing fibril with the rest of the residues hanging out intact[35] or gets associated in some other manner with the fibrillar assembly[28]. One of the few models that attempts to accommodate the complete native structures in amyloid fibres is the zipper-spine model based on two distinct domain swapped structures of the RNase A molecule[36]. In the 'polar zippers', in addition to the intermolecular backbone hydrogen bonds, the cross-β structural motif is further stabilized by intermolecular hydrogen bonds between side chain amide and carbonyl groups. The polar zipper model is believed to be relevant in fibrils formed by proteins such as Huntingtin and yeast prion protein containing glutamine- or asparagine-rich segments[33].

Mature amyloid fibrils typically consist of 2–6 unbranched protofilaments. These protofilaments have a diameter of 2–5 nm and associate laterally or get twisted together to form fibrils of 4–13 nm in diameter[1]. However, the fibrils formed from the different proteins differ in number, packing arrangement and helical twists of the constituent protofilaments[37].

There have been conflicting reports on the role of sequence dependence on the assembly of the polypeptide chains in the protofilaments. Study of a dozen soluble oligomers from different amyloidogenic proteins/peptides, suggested that the structure of the soluble oligomers of amyloidogenic proteins/peptides is independent of the sequence[38]. Independent simulation studies on some peptides (whose amino acid sequences are AGAAAAGA, KFFE and NFGAIL) revealed that well-aligned, anti-parallel sheets oriented in parallel could possibly seed fibril formation[39]. In contrast, the amyloid-forming peptide (amino acid sequence: GNNQQNY) from the yeast prion Sup35 has been proposed to have precisely the opposite arrangement, viz. parallel β-sheets oriented antiparallel to each other[40]. It is envisaged that the side-chain chemical diversity induces formation of a complex network of interactions that ultimately dictate the microscopic arrangement of the strands at the protofilament level[41]. Solid-state NMR studies on Aβ peptides have suggested that optimization of hydrophobic contacts along with the electrostatic and other interactions decide the finer details of supramolecular organization of protofilaments[33]. Apart from the role of the sequence, differences in the conditions used for fibril growth could also affect its architecture with regard to side chain interactions and β registry[30].

Advantages and disadvantages of amyloid formation

Amyloid formation is associated with a number of fatal diseases. It is not clear if amyloid fibril formation is a blessing or a curse in amyloid-related pathology, since the soluble oligomeric species and not the mature fibrils are now believed to be associated with cytotoxicity[42]. If this is true, then amyloid fibril formation can be viewed as a desirable process. Nevertheless, the pathophysiological mechanisms could arise as a result of lack of biological function and/or gain of toxicity contributed by both the oligomers and the fibrils[43]. In a development that is unrelated to human pathology, amyloid fibrils are presently being exploited as useful scaffolds for biomimetic materials with significant optical, magnetic, electronic and/or mechanical properties. Amyloid fibrils thus have applications in materials science[33].

Mechanism of protein aggregation

Protein aggregation can be broadly viewed as occurring through three major steps. In the first step, the soluble native proteins get transformed into aggregation-prone precursor molecules, also referred to as 'off-pathway' partially folded intermediates[1]. The ease of attaining the partially unfolded/folded state is therefore considered an important parameter in governing the aggregation propensity of proteins. Natively unfolded or intrinsically unstructured proteins do not harbour significant tertiary structure and can therefore easily attain the partially folded state compared to proteins with compact tertiary structures[44]. As expected, a majority of the known amyloidogenic proteins are known to be natively unstructured[1]. Likewise, helices that have high propensity for attaining extended strand conformations are predicted to be aggregation-prone[45,46].

The folding intermediates are more prone to aggregation compared to the unfolded species of the same protein. Partial folding can bring together the distantly placed hydrophobic residues to create a contiguous 'aggregation-prone'

hydrophobic surface. Apart from the conformation, another determinant of aggregation is the lifetime of the partially folded intermediates of a protein. The presence of long-lived partially folded intermediates in the folding pathway of a protein predisposes the protein for aggregation: (a) by interacting with each other forming non-native intermolecular contacts, and (b) by exhausting the availability of cellular molecular chaperones for other nascent polypeptide chains[2]. Thus, protein aggregation is to a large extent dependent on the folding kinetics of a protein. All proteins in Nature do not have optimized folding rates and the crowded cellular environment further promotes the aggregation of such slow folders[47].

In the second step of aggregation, which can be referred to as the 'nucleation' phase, the intermediates assemble in a specific manner to form discrete-structured oligomers or 'nuclei'. Formation of nuclei is a kinetically disfavoured step and hence is the rate-limiting or the lag phase of the aggregation process[48]. The structures of the oligomers are believed to depend on the protein and the environment[30].

The third step is the 'polymerization' phase wherein the oligomers assemble to form amyloid fibrils or inclusion bodies. This phase is kinetically favoured and hence is much faster than the nucleation phase. This 'nucleation–polymerization' mechanism of aggregation can be accelerated by the addition of 'pre-formed nuclei', which reduces the lag phase of nucleation[2].

The funnel landscape of protein folding could help in understanding the homogeneous constitution of protein aggregates. Each protein sequence dictates a unique folding funnel under a given set of environmental conditions. Hence, there is greater opportunity for interaction between the folding intermediates that get trapped in local energy minima of a protein species as they traverse the funnel to reach their native folded state. Preventive measures of protein aggregation are thus broadly targetted at smoothening of the folding energy landscapes of the respective proteins[49].

Factors affecting protein aggregation

As with protein folding, aggregation caused by misfolding is also determined by the sequence and the environment. These factors may act in an independent or cooperative fashion to bring about aggregation. Generally, any sequence or environment characteristic that perturbs the stability of the native structures but is still favourable for non-covalent interactions can lead to aggregation[1].

The sequence-based characteristics that are associated with aggregation are hydrophobicity, net charge and secondary structural propensities[50]. In certain polypeptides the N- and C-terminal residue compositions are found to be more critical in determining the intrinsic propensity for inclusion body[51,52] and amyloid formation[53,54] compared to the rest of the protein. The nature of the partially folded intermediates has also been implicated in inclusion body and amyloid formation[2,55]. Protein concentration, pH, temperature and ionic strength are some of the environment-based determinants of aggregation[56].

Inclusion body formation

Some of the factors that have been associated with inclusion body formation in *E. coli* are: (i) high local concentration of the overexpressed protein[57]; (ii) reducing environment in the cytoplasm due to high levels of glutathione, preventing disulphide bond formation[13,14]; (iii) lack of post-translational modifications such as glycosylation which could improve solubility of the protein[58]; (iv) improper interactions with chaperones and other proteins participating in folding *in vivo*[59]; (v) intermolecular cross-linking via disulphides; however, proteins without cysteine residues also form aggregates, e.g. apomyoglobin, lactamase and interferon and (vi) kinetics of protein translation in the context of rare codons[60].

Statistical and mutation studies have been conducted to decipher the sequence determinants for inclusion-body formation in *E. coli*. These studies have associated hydrophobicity, charge, turn-forming residues, aliphatic index and instability index of the residues which measures the thermostability and *in vivo* half-life of the proteins respectively, to play a role in the propensity for inclusion body formation[61–65].

In the case of mammalian intracellular aggregation, the key causative factors are (a) lack of ligands that stabilize the native conformations of the proteins, (b) inadequacy of chaperones that assist in proper folding of the nascent polypeptide chains, and (c) aberrant degradation of misfolded proteins by the lysosome or proteasome complex[7].

Amyloid formation

Elucidation of the factors leading to amyloid formation is an active area of research. Some of the elicitors of amyloid formation *in vivo* are high protein concentration, improper proteolysis, mutations in the polypeptide, local change in pH at membranes and oxidative or heat stress[66,67].

Sequence-based analyses conducted on amyloidogenic proteins have revealed that amyloidogenicity is positively correlated to the *in vivo* half-life and presence of order-promoting residues with high sheet propensity in a protein, and negatively correlated to the thermostability of the protein as decided by the content of aliphatic amino acids[46]. The other sequence-related parameters that have been implicated in deciding the amyloidogenicity of polypeptides are hydrophobic–hydrophilic patterning, net charge and placement of the charged residues, β-sheet and β-turn propensity, nature of the peripheral residues of amyloidogenic stretches

and the presence of prolines and aromatic amino acids[50,67,68]. The presence of numerous structural defects, such as backbone hydrogen bonds insufficiently shielded from water, in a protein have also been suggested to play a role in dictating amyloidogenic propensity[69].

Amyloid formation is also believed to be critically dependent on the presence of other molecules, such as metal ions, glycosaminoglycans, glycoproteins (e.g. serum amyloid P and apolipoprotein E), and constituents of basement membranes such as perlecan, laminin and agrin. The significant role played by these 'pathological chaperones' in the formation and stability of amyloid fibrils has led to the view that amyloid formation could be the consequence of improper heterologous interactions rather than improper protein folding[70].

However, the contribution/importance of these factors to fibril formation varies among proteins. For example, sequence-dependent factors such as secondary structure propensity, peptide length, isoelectric point (pI) and hydrophobicity were not found to affect the amyloidogenicity of β2-microglobulin peptides; the high content of aromatic side chains was found to be the only major determinant[54].

Overcoming protein aggregation

Efforts to overcome protein aggregation can be broadly divided into environment-based and protein-based modifications. In spite of the plethora of studies devoted to exploring the methods of overcoming protein aggregation, trial-and-error-methods are still largely in vogue to achieve respite from this problem. A universal protocol for overcoming aggregation in all proteins is yet to be identified and probably such a target is unrealistic since the mechanism and the predisposing factors for aggregation vary with proteins[70,71].

Inclusion body formation

Environment-based approaches for prevention of formation of inclusion bodies in *E. coli* include modifications in the host bacterial strain, expression vector and growth conditions[72–74]. Inducing an increase in chaperone activity in the overexpressing cells either by subjecting the host cells to heat shock prior to induction[75] and/or co-expressing with molecular chaperones[14] has been found to reduce inclusion-body formation. Use of certain additives in the growth medium like L-arginine[76], ethanol[77], detergents[78], monoclonal antibodies[79] and reducing the growth temperature[80] also aids in decreasing inclusion-body formation.

Site-directed mutagenesis, based on existing knowledge of sequence determinants for inclusion-body formation or random mutagenesis and the use of fusion tags are the main sequence-based procedures that have been explored for reducing or preventing bacterial inclusion-body formation[81,82].

High-throughput screening methods are being developed which allow simultaneous evaluation of different fusion constructs, mutations, refolding conditions, etc.[12,83–87]. A few other studies have used directed evolution methods, in which protein diversity libraries are screened for soluble variants; these methods do not require structural or functional information of the target protein[88].

Amyloid formation

Amyloidogenicity of a protein can be reduced by stabilizing the native structure by ligand binding, addition of monoclonal antibody fragments, disaccharides (e.g. trehalose), etc.[89–91]. Addition of compounds that inhibit interaction of the molecules with amyloidogenic proteins such as glycosaminoglycans and serum amyloid P has also shown to reduce fibril formation[5]. Other examples of environment-based factors that have been tried to reduce amyloid formation are the addition of 'mini-chaperones', β-breakers or removal of pathological chaperones such as metal ions and apolipoprotein E that stabilize the β-structures[66,70]. Sensitivity of amyloid proteins to mutations has led to exploiting protein engineering as a tool to prevent amyloid formation[92]. In this approach proteins are designed with markedly reduced propensities to aggregate, based on the current knowledge of relation between protein structure and fibril formation.

Similarities between inclusion body and amyloid formation

Inclusion body and amyloid formation being different manifestations of protein aggregation, do share certain similarities in their mechanism of formation, structure and function. It was earlier believed that aggregation into inclusion bodies does not involve any specific interaction between the aggregating proteins[93]. However, it is now established that, similar to amyloid fibrils, inclusion bodies also arise from specific interactions between the partially folded proteins[94] and that the inclusion bodies do have certain degree of order in their architecture[23].

Amyloid and inclusion-body formation are exhibited by a wide range of unrelated proteins[4,29]. Proteins that exhibit amyloid formation *in vitro* have been shown to form inclusion bodies in *E. coli* and correspondingly, mutants with reduced amyloidogenicity are found to have increased solubility levels when overexpressed in *E. coli*.[23]. Hence, the characteristics of primary structure that influence aggregation propensity are similar in both cases. Environmental triggers for both forms of aggregation are also similar and can be broadly termed as 'denaturing' or 'destabilizing'. These similarities in the factors that affect their aggregation potential are also suggestive of a common mechanism of formation.

Another evidence for similarity in the mechanism of formation and structure of the two forms of aggregates comes from the observation that pre-formed inclusion bodies and amyloid fibrils can accelerate the aggregation of soluble homologous polypeptides in a selective and dose-dependent manner[23]. This seeding effect of the aggregates strengthens the view that both inclusion body and amyloid formation occur via a nucleation–polymerization pathway[2].

Role of evolution in controlling protein aggregation

The ability of a nascent polypeptide chain to fold into a unique native conformation is in itself a remarkable achievement of evolution[95]. Evolution has helped in minimizing protein aggregation by natural selection of both the non-aggregating sequences and a favourable environment in which these proteins are synthesized and function[28].

The cellular environment has evolved to assist the nascent chains from aggregation by hosting an efficient quality control system. This system encompasses a group of molecular chaperones, folding catalysts and proteasomes that function in an orchestrated manner to assist a wide repertoire of nascent polypeptides in their folding pathway and prevent accumulation of misfolded proteins in the cell. A cellular environment conducive for protein folding is maintained by tight regulation of the cellular temperature and pH. Cellular homeostasis is further ensured in events of abrupt changes in the cellular temperature or redox status by the induction of a plethora of heat shock proteins, which guard the proteins from aggregation[96].

In eukaryotic cells, ER (endoplasmic reticulum) is the site for folding of non-cytosolic proteins. As can be expected, ER has evolved with stringent quality control mechanisms to ensure that only the correctly folded proteins are exported. Likewise, its efficient signalling pathways, called the unfolded protein response, have helped in coping with the proteins that have failed to achieve the correct native structure[97].

Although cells are equipped to assist the folding of nascent polypeptides with an efficient molecular chaperone system, the sequences have evolved to fold independent of them such that less than a third of the proteins needs the assistance of folder chaperones in attaining their native conformation[56]. Some of the evolutionary strategies developed by the proteins to counter aggregation are:

(a) Hydrophobic stretches of the proteins are usually buried in the native conformation.
(b) Aggregation-prone stretches of the proteins are flanked by 'gate-keeper' residues such as glycine, proline, arginine and lysine. These residues prevent aggregate formation either by their structure-breaking properties (Gly and Pro) or by creating a high repulsive force on self-assembly (Lys and Arg). The long side chains of Lys and Arg have large conformational entropy, which further restricts their immobilization in an aggregate. These positive residues are also exploited by molecular chaperones in selectively identifying the aggregation-prone hydrophobic stretches of the nascent polypeptides[4].
(c) Statistical analyses have revealed that sequences with alternating polar and nonpolar amino acids favour amyloid formation and interestingly, sequences with such binary patterns are seen to be rare in the database of natural proteins[68].
(d) The free-edge strands of proteins are protected from non-native intermolecular β-sheet interactions by the strategic placement of prolines, charged residues, very short edge strands, β-bulges and long loops[43].

Even though strategies have evolved to counter aggregation, there is an increasing incidence of amyloid-related diseases, most of which occur in the later stages of life. In case of such age-related amyloidoses, weakening of the quality control mechanism, evolved for ensuring cellular homeostasis, has been implicated. It has been proposed that amyloidogenic proteins might have evaded evolutionary selection – one way this might happen is if the extent to which/conditions under which these proteins can form amyloid fibrils does not affect the reproductive ability of the organism[7]. Another reason is that some of these diseases are an outcome of recently introduced practices and thus, there has not been any time to evolve mechanisms that can effectively protect against such diseases[28]. Examples of such diseases include BSE (bovine spongiform encephalopathy; because of new agricultural practices) and Type II diabetes (because of changing diet).

While taking stock of the role of evolution in curbing protein aggregation, the fact that proteins are evolved not only to fold to a compact stable state but also to function effectively, must not be ignored. Thus the structures of the proteins need to be metastable, giving room for 'conformational breathing'. A possible offshoot of this structural flexibility is the ease of partial unfolding of these polypeptides and thus succumbing to aggregation in an unfavourable environment. In this regard, aggregation of proteins might be the evolutionary cost one has to pay for the benefits associated with their functions[7,49].

Concluding remarks

The main manifestations of protein aggregation are amyloid and inclusion body formation. The number of individuals reported to be suffering from protein conformational disorders has increased substantially over the past decade. Large-scale production of soluble proteins in *E. coli* has been hindered due to the formation of inclusion bodies. Several experimental and computational studies have been devoted to comprehend the underlying factors and mechanism

of protein aggregation. However, studying the effects of the sequence and environment in all the possible permutations and combinations which exist in nature, on protein aggregation is a rather daunting task. As of now, a unifying mechanism of protein aggregation is yet to be deciphered and such a generalized mechanism may in fact be elusive, since the predisposing factors for aggregation vary with the proteins and their environment. However, in the absence of such a unique mechanism of aggregation for proteins, the process of identifying measures for overcoming aggregation would also have to be specific for the protein and the environment. This perhaps explains the slow progress that has been made in the field of understanding and combating protein aggregation.

1. Uversky, V. N. and Fink, A. L., Conformational constraints for amyloid fibrillation: The importance of being unfolded. *Biochim. Biophys. Acta*, 2004, **1698**, 131–153.
2. Fink, A. L., Protein aggregation: Folding aggregates, inclusion bodies and amyloid. *Fold Des.*, 1998, **3**, R9–R23.
3. Ventura, S., Sequence determinants of protein aggregation: Tools to increase protein solubility. *Microb. Cell Fact.*, 2005, **4**, 11.
4. Rousseau, F., Schymkowitz, J. and Serrano, L., Protein aggregation and amyloidosis: Confusion of the kinds?. *Curr. Opin. Struct. Biol.*, 2006, **16**, 118–126.
5. Hirschfield, G. M., Amyloidosis: A clinico-pathophysiological synopsis. *Semin. Cell Dev. Biol.*, 2004, **15**, 39–44.
6. Murzin, A. G., Brenner, S. E., Hubbard, T. and Chothia, C., SCOP: A structural classification of proteins database for the investigation of sequences and structures. *J. Mol. Biol.*, 1995, **247**, 536–540.
7. Buxbaum, J. N., Diseases of protein conformation: What do *in vitro* experiments tell us about *in vivo* diseases?. *Trends Biochem. Sci.*, 2003, **28**, 585–592.
8. Pieri, L., Bucciantini, M., Nosi, D., Formigli, L., Savistchenko, J., Melki, R. and Stefani, M., The yeast prion Ure2p native-like assemblies are toxic to mammalian cells regardless to their aggregation state. *J. Biol. Chem.*, 2006, **281**, 15337–15344.
9. Scheibel, T. and Buchner, J., Protein aggregation as a cause for disease. *Handb. Exp. Pharmacol.*, 2006, **172**, 199–219.
10. Kirschner, D. A., Elliott-Bryant, R., Szumowski, K. E., Gonnerman, W. A., Kindy, M. S., Sipe, J. D. and Cathcart, E. S., *In vitro* amyloid fibril formation by synthetic peptides corresponding to the amino terminus of apoSAA isoforms from amyloid-susceptible and amyloid-resistant mice. *J. Struct. Biol.*, 1998, **124**, 88–98.
11. Tabner, B. J., El-Agnaf, O. M., German, M. J., Fullwood, N. J. and Allsop, D., Protein aggregation, metals and oxidative stress in neurodegenerative diseases. *Biochem. Soc. Trans.*, 2005, **33**, 1082–1086.
12. Pedelacq, J. D. *et al.*, Engineering soluble proteins for structural genomics. *Nature Biotechnol.*, 2002, **20**, 927–932.
13. Lilie, H., Schwarz, E. and Rudolph, R., Advances in refolding of proteins produced in *E. coli*. *Curr. Opin. Biotechnol.*, 1998, **9**, 497–501.
14. Makrides, S. C., Strategies for achieving high-level expression of genes in *Escherichia coli*. *Microbiol. Rev.*, 1996, **60**, 512–538.
15. Singh, S. M. and Panda, A. K., Solubilization and refolding of bacterial inclusion body proteins. *J. Biosci. Bioeng.*, 2005, **99**, 303–310.
16. Schrodel, A. and de Marco, A., Characterization of the aggregates formed during recombinant protein expression in bacteria. *BMC Biochem.*, 2005, **6**, 10.
17. Oberg, K., Chrunyk, B. A., Wetzel, R. and Fink, A. L., Native-like secondary structure in interleukin-1 beta inclusion bodies by attenuated total reflectance FTIR. *Biochemistry*, 2005, **33**, 2628–2634.
18. Tokatlidis, K., Dhurjati, P., Millet, J., Beguin, P. and Aubert, J. P., High activity of inclusion bodies formed in *Escherichia coli* overproducing *Clostridium thermocellum* endoglucanase D. *FEBS Lett.*, 1991, **282**, 205–208.
19. Worrall, D. M. and Goss, N. H., The formation of biologically active beta-galactosidase inclusion bodies in *Escherichia coli*. *Aust. J. Biotechnol.*, 1989, **3**, 28–32.
20. Garcia-Fruitos, E. *et al.*, Aggregation as bacterial inclusion bodies does not imply inactivation of enzymes and fluorescent proteins. *Microb. Cell Fact.*, 2005, **4**, 27.
21. Bowden, G. A., Paredes, A. M. and Georgiou, G., Structure and morphology of protein inclusion bodies in *Escherichia coli*. *Biotechnology*, 1991, **9**, 725–730.
22. Przybycien, T. M., Dunn, J. P., Valax, P. and Georgiou, G., Secondary structure characterization of beta-lactamase inclusion bodies. *Protein Eng.*, 1994, **7**, 131–136.
23. Carrio, M., Gonzalez-Montalban, N., Vera, A., Villaverde, A. and Ventura, S., Amyloid-like properties of bacterial inclusion bodies. *J. Mol. Biol.*, 2005, **347**, 1025–1037.
24. Clark, E. D. B., Protein refolding for industrial purposes. *Curr. Opin. Biotechnol.*, 2001, **12**, 202–207.
25. Jurgen, B., Hanschke, R., Sarvas, M., Hecker, M. and Schweder, T., Proteome and transcriptome based analysis of *Bacillus subtilis* cells overproducing an insoluble heterologous protein. *Appl. Microbiol. Biotechnol.*, 2001, **55**, 326–332.
26. Lindwall, G., Chau, M., Gardner, S. R. and Kohlstaedt, L. A., A sparse matrix approach to the solubilization of overexpressed proteins. *Protein Eng.*, 2000, **13**, 67–71.
27. Dziewulska, D. and Rafalowska, J., Proteinaceous intracellular inclusions in neurodegenerative disorders. *Folia Neuropathol.*, 2005, **43**, 51–63.
28. Dobson, C. M., Principles of protein folding, misfolding and aggregation. *Semin. Cell Dev. Biol.*, 2004, **15**, 3–16.
29. Dobson, C. M., Protein misfolding, evolution and disease. *Trends Biochem. Sci.*, 1999, **24**, 329–332.
30. Makin, O. S. and Serpell, L. C., Structures for amyloid fibrils. *FEBS J.*, 2005, **272**, 5950–5961.
31. Nelson, R., Sawaya, M. R., Balbirnie, M., Madsen, A. O., Riekel, C., Grothe, R. and Eisenberg, D., Structure of the cross-beta spine of amyloid-like fibrils. *Nature*, 2005, **435**, 773–778.
32. Luehrs, T. *et al.*, 3D structure of Alzheimer's amyloid-beta (1–42) fibrils. *Proc. Natl. Acad. Sci. USA*, 2005, **102**, 17342–17347.
33. Tycko, R., Progress towards a molecular-level structural understanding of amyloid fibrils. *Curr. Opin. Struct. Biol.*, 2004, **14**, 96–103.
34. Pickersgill, R., Jenkins, J., Harris, G., Nasser, W. and Robert-Baudouy, J., The structure of Bacillus subtilis pectate lyase in complex with calcium. *Nature Struct. Mol. Biol.*, 1994, **1**, 717–723.
35. Ma, B. and Nussinov, R., Stabilities and conformations of Alzheimer's beta-amyloid peptide oligomers (Abeta 16-22, Abeta 16-35, and Abeta 10-35): Sequence effects. *Proc. Natl. Acad. Sci. USA*, 2002, **99**, 14126–14131.
36. Liu, Y., Gotte, G., Libonati, M. and Eisenberg, D., Structures of the two 3D domain-swapped RNase A trimers. *Protein Sci.*, 2002, **11**, 371–380.
37. Tattum, M. H. *et al.*, Elongated oligomers assemble into mammalian PrP amyloid fibrils. *J. Mol. Biol.*, 2006, **357**, 975–985.
38. Kayed, R., Head, E., Thompson, J. L., McIntire, T. M., Milton, S. C., Cotman, C. W. and Glabe, C. G., Common structure of soluble amyloid oligomers implies common mechanism of pathogenesis. *Science*, 2003, **300**, 486–489.

39. Ma, B. and Nussinov, R., Molecular dynamics simulations of alanine rich beta-sheet oligomers: Insight into amyloid formation. *Protein Sci.*, 2002, **11**, 2335–2350.
40. Balbirnie, M., Grothe, R. and Eisenberg, D. S., An amyloid-forming peptide from the yeast prion Sup35 reveals a dehydrated beta-sheet structure for amyloid. *Proc. Natl. Acad. Sci. USA*, 2001, **98**, 2375–2380.
41. Zanuy, D., Haspel, N., Tsai, H. H., Ma, B., Gunasekaran, K., Wolfson, H. J. and Nussinov, R., Side chain interactions determine the amyloid organization: A single layer beta-sheet molecular structure of the calcitonin peptide segment 15–19. *Phys. Biol.*, 2004, **1**, 89–99.
42. Bucciantini, M. *et al.*, Inherent toxicity of aggregates implies a common mechanism for protein misfolding diseases. *Nature*, 2002, **416**, 507–511.
43. Johansson, J., Molecular determinants for amyloid fibril formation: Lessons from lung surfactant protein C. *Swiss Med. Wkly.*, 2003, **133**, 275–282.
44. Uversky, V. N., Protein folding revisited. A polypeptide chain at the folding–misfolding–nonfolding cross-roads: Which way to go?. *Cell. Mol. Life Sci.*, 2003, **60**, 1852–1871.
45. Kallberg, Y., Gustafsson, M., Persson, B., Thyberg, J. and Johansson, J., Prediction of amyloid fibril-forming proteins. *J. Biol. Chem.*, 2001, **276**, 12945–12950.
46. Idicula-Thomas, S. and Balaji, P. V., Understanding the relationship between the primary structure of proteins and their amyloidogenic propensity: Clues from inclusion body formation. *Prot. Eng. Des. Sel.*, 2005, **18**, 175–180.
47. Viguera, A. R., Villegas, V., Aviles, F. X. and Serrano, L., Favourable native-like helical local interactions can accelerate protein folding. *Fold. Des.*, 1997, **2**, 23–33.
48. Ohnishi, S. and Takano, K., Amyloid fibrils from the viewpoint of protein folding. *Cell Mol. Life Sci.*, 2004, **61**, 511–524.
49. Jahn, T. R. and Radford, S. E., The Yin and Yang of protein folding. *FEBS J.*, 2005, **272**, 5962–5970.
50. Chiti, F., Stefani, M., Taddei, N., Ramponi, G. and Dobson, C. M., Rationalization of the effects of mutations on peptide and protein aggregation rates. *Nature*, 2003, **424**, 805–808.
51. Boeggeman, E. E., Ramakrishnan, B. and Qasba, P. K., The N-terminal stem region of bovine and human beta 1,4-galactosyltransferase I increases the *in vitro* folding efficiency of their catalytic domain from inclusion bodies. *Protein Exp. Purif.*, 2003, **30**, 219–229.
52. Sati, S. P., Singh, S. K., Kumar, N. and Sharma, A., Extra terminal residues have a profound effect on the folding and solubility of a *Plasmodium falciparum* sexual stage-specific protein overexpressed in *Escherichia coli*. *Eur. J. Biochem.*, 2002, **269**, 5259–5263.
53. Goldsbury, C. *et al.*, Amyloid fibril formation from full-length and fragments of amylin. *J. Struct. Biol.*, 2000, **130**, 352–362.
54. Jones, S., Manning, J., Kad, N. M. and Radford, S. E., Amyloid-forming peptides from beta 2-microglobulin – Insights into the mechanism of fibril formation *in vitro*. *J. Mol. Biol.*, 2003, **325**, 249–257.
55. King, J., Haase-Pettingell, C., Robinson, A. S., Speed, M. and Mitraki, A., Thermolabile folding intermediates: Inclusion body precursors and chaperonin substrates. *FASEB J.*, 1996, **10**, 57–66.
56. Yon, J. M., Protein folding in the post-genomic era. *J. Cell Mol. Med.*, 2002, **6**, 307–327.
57. Horwich, A., Protein aggregation in disease: A role for folding intermediates forming specific multimeric interactions. *J. Clin. Invest.*, 2002, **110**, 1221–1232.
58. Zhang, Y., Olsen, D. R., Nguyen, K. B., Olson, P. S., Rhodes, E. T. and Mascarenhas, D., Expression of eukaryotic proteins in soluble form in *Escherichia coli*. *Protein Exp. Purif.*, 1998, **12**, 159–165.
59. Mogk, A., Mayer, M. P. and Deuerling, E., Mechanisms of protein folding: Molecular chaperones and their application in biotechnology. *Chembiochem.*, 2002, **3**, 807–814.
60. Komar, A. A., Lesnik, T. and Reiss, C., Synonymous codon substitutions affect ribosome traffic and protein folding during *in vitro* translation. *FEBS Lett.*, 1999, **462**, 387–391.
61. Davis, G. D., Elisee, C., Newham, D. M. and Harrison, R. G., New fusion protein systems designed to give soluble expression in *Escherichia coli*. *Biotechnol. Bioeng.*, 1999, **65**, 382–388.
62. Bertone, P. *et al.*, SPINE: An integrated tracking database and data mining approach for identifying feasible targets in high-throughput structural proteomics. *Nucleic Acids Res.*, 2001, **29**, 2884–2898.
63. Luan, C. H. *et al.*, 2004. High-throughput expression of *C. elegans* proteins. *Genome Res.*, **14**, 2102–2110.
64. Idicula-Thomas, S. and Balaji, P. V., Understanding the relationship between the primary structure of proteins and their propensity to be soluble on overexpression in *Escherichia coli*. *Protein Sci.*, 2005, **14**, 582–592.
65. Idicula-Thomas, S., Kulkarni, A. J., Kulkarni, B. D., Jayaraman, V. K. and Balaji, P. V., A support vector machine-based method for predicting the propensity of a protein to be soluble or to form inclusion body on over-expression in *Escherichia coli*. *Bioinformatics*, 2006, **22**, 278–284.
66. Zerovnik, E., Amyloid-fibril formation. Proposed mechanisms and relevance to conformational disease. *Eur. J. Biochem.*, 2002, **269**, 3362–3371.
67. DuBay, K. F., Pawar, A. P., Chiti, F., Zurdo, J., Dobson, C. M. and Vendruscolo, M., Prediction of the absolute aggregation rates of amyloidogenic polypeptide chains. *J. Mol. Biol.*, 2004, **341**, 1317–1326.
68. Wurth, C., Kim, W. and Hecht, M. H., Combinatorial approaches to probe the sequence determinants of protein aggregation and amyloidogenicity. *Protein Pept. Lett.*, 2006, **13**, 279–286.
69. Fernandez, A., Kardos, J., Scott, L. R., Goto, Y. and Berry, R. S., Structural defects and the diagnosis of amyloidogenic propensity. *Proc. Natl. Acad. Sci. USA*, 2003, **100**, 6446–6451.
70. Alexandrescu, A. T., Amyloid accomplices and enforcers. *Protein Sci.*, 2005, **14**, 1–12.
71. Wetzel, R., Perry, L. J. and Veilleux, C., Mutations in human interferon gamma affecting inclusion body formation identified by a general immunochemical screen. *Biotechnology*, 1991, **9**, 731–737.
72. Clark, E. D. B., Refolding of recombinant proteins. *Curr. Opin. Biotechnol.*, 1998, **9**, 157–163.
73. Georgiou, G. and Valax, P., Isolating inclusion bodies from bacteria. *Methods Enzymol.*, 1999, **309**, 48–58.
74. Rudolph, R., Renaturation of recombinant, disulfide-bonded proteins from inclusion bodies. In *Modern Methods in Protein and Nucleic Acid Research* (ed. Tschesche, H.), Walter de Gruyter, Berlin, 1990, pp. 149–172.
75. Chen, J., Acton, T. B., Basu, S. K., Montelione, G. T. and Inouye, M., Enhancement of the solubility of proteins overexpressed in *Escherichia coli* by heat shock. *J. Mol. Microbiol. Biotechnol.*, 2002, **4**, 519–524.
76. Winter, J., Neubauer, P., Glockshuber, R. and Rudolph, R., Increased production of human proinsulin in the periplasmic space of *Escherichia coli* by fusion to DsbA. *J. Biotechnol.*, 2001, **84**, 175–185.
77. Nygaard, F. B. and Harlow, K. W., Heterologous expression of soluble, active proteins in *Escherichia coli*: The human estrogen receptor hormone-binding domain as paradigm. *Protein Exp. Purif.*, 2001, **21**, 500–509.
78. Tandon, S. and Horowitz, P., The effects of lauryl maltoside on the reactivation of several enzymes after treatment with guanidinium chloride. *Biochim. Biophys. Acta*, 1988, **955**, 19–25.
79. Katzav-Gozansky, T., Hanan, E. and Solomon, B., Effect of monoclonal antibodies in preventing carboxypeptidase A aggregation. *Biotechnol. Appl. Biochem.*, 1996, **23**, 227–230.

80. Schein, C. H. and Noteborn, M. H. M., Formation of soluble recombinant proteins in *Escherichia coli* is favored by lower growth temperature. *Biotechnology*, 1988, **6**, 291–294.
81. Jenkins, T. M., Hickman, A. B., Dyda, F., Ghirlando, R., Davies, D. R. and Craigie, R., Catalytic domain of human immunodeficiency virus type 1 integrase: Identification of a soluble mutant by systematic replacement of hydrophobic residues. *Proc. Natl. Acad. Sci. USA*, 1995, **92**, 6057–6061.
82. Sachdev, D. and Chirgwin, J. M., Solubility of proteins isolated from inclusion bodies is enhanced by fusion to maltose-binding protein or thioredoxin. *Protein Exp. Purif.*, 1998, **12**, 122–132.
83. Hammarstrom, M., Hellgren, N., van Den Berg, S., Berglund, H. and Hard, T., Rapid screening for improved solubility of small human proteins produced as fusion proteins in *Escherichia coli*. *Protein Sci.*, 2002, **11**, 313–321.
84. Knaust, R. K. and Nordlund, P., Screening for soluble expression of recombinant proteins in a 96-well format. *Anal. Biochem.*, 2001, **297**, 79–85.
85. Lesley, S. A., Graziano, J., Cho, C. Y., Knuth, M. W. and Klock, H. E., Gene expression response to misfolded protein as a screen for soluble recombinant protein. *Protein Eng.*, 2002, **15**, 153–160.
86. Shih, Y. P., Kung, W. M., Chen, J. C., Yeh, C. H., Wang, A. H. and Wang, T. F., High-throughput screening of soluble recombinant proteins. *Protein Sci.*, 2002, **11**, 1714–1719.
87. Wigley, W. C., Stidham, R. D., Smith, N. M., Hunt, J. F. and Thomas, P. J., Protein solubility and folding monitored *in vivo* by structural complementation of a genetic marker protein. *Nature Biotechnol.*, 2001, **19**, 131–136.
88. Waldo, G. S., Genetic screens and directed evolution for protein solubility. *Curr. Opin. Chem. Biol.*, 2003, **7**, 33–38.
89. Tjernberg, L. O. *et al.*, Arrest of beta-amyloid fibril formation by a pentapeptide ligand. *J. Biol. Chem.*, 1996, **271**, 8545–8548.
90. Solomon, B., Anti-aggregating antibodies, a new approach towards treatment of conformational diseases. *Curr. Med. Chem.*, 2002, **9**, 1737–1749.
91. Tanaka, M., Machida, Y. and Nukina, N., A novel therapeutic strategy for polyglutamine diseases by stabilizing aggregation-prone proteins with small molecules. *J. Mol. Med.*, 2005, **83**, 343–352.
92. Villegas, V., Zurdo, J., Filimonov, V. V., Aviles, F. X., Dobson, C. M. and Serrano, L., Protein engineering as a strategy to avoid formation of amyloid fibrils. *Protein Sci.*, 2000, **9**, 1700–1708.
93. Clark, E. D. B., Schwarz, E. and Rudolph, R., Inhibition of aggregation side reactions during *in vitro* protein folding. *Methods Enzymol.*, 1999, **309**, 217–236.
94. Ventura, S. and Villaverde, A., Protein quality in bacterial inclusion bodies. *Trends Biotechnol.*, 2006, **24**, 179–185.
95. Dinner, A. R., Sali, A., Smith, L. J., Dobson, C. M. and Karplus, M., Understanding protein folding via free-energy surfaces from theory and experiment. *Trends Biochem. Sci.*, 2000, **25**, 331–339.
96. Papp, E., Nardai, G., Soti, C. and Csermely, P., Molecular chaperones, stress proteins and redox homeostasis. *Biofactors*, 2003, **17**, 249–257.
97. Zhang, K. and Kaufman, R. J., The unfolded protein response: A stress signaling pathway critical for health and disease. *Neurology*, 2006, **66**, S102–S109.
98. Jacobsen, J. S., Reinhart, P. and Pangalos, M. N., Current concepts in therapeutic strategies targeting cognitive decline and disease modification in Alzheimer's disease. *NeuroRx.*, 2005, **2**, 612–626.
99. Jain, S., Wood, N. W. and Healy, D. G., Molecular genetic pathways in Parkinson's disease: A review. *Clin. Sci.*, 2005, **109**, 355–364.
100. Bratosiewicz-Wasik, J., Wasik, T. J. and Liberski, P. P., Molecular approaches to mechanisms of prion diseases. *Folia Neuropathol.*, 2004, **42** Suppl. A, 33–46.
101. Welch, W. J. Role of quality control pathways in human diseases involving protein misfolding. *Semin. Cell Dev. Biol.*, 2004, **15**, 31–38.
102. Pey, A. L., Desviat, L. R., Gamez, A., Ugarte, M. and Perez, B., Phenylketonuria: Genotype-phenotype correlations based on expression analysis of structural and functional mutations in PAH. *Hum. Mutat.*, 2003, **21**, 370–378.
103. Whiteman, P., Hutchinson, S. and Handford, P. A., Fibrillin-1 misfolding and disease. *Antioxid. Redox Signal.*, 2006, **8**, 338–346.
104. Gajko-Galicka, A., Mutations in type I collagen genes resulting in osteogenesis imperfecta in humans. *Acta Biochim. Pol.*, 2002, **49**, 433–441.
105. Schnog, J. B., Duits, A. J., Muskiet, F. A., ten Cate, H., Rojer, R. A. and Brandjes, D. P., Sickle cell disease: A general overview. *Neth. J. Med.*, 2004, **62**, 364–374.
106. Montalvo, A. L. *et al.*, Molecular analysis of the *HEXA* gene in Italian patients with infantile and late onset Tay-Sachs disease: detection of fourteen novel alleles. *Hum. Mutat.*, 2005, **26**, 282.
107. Merlini, G. and Bellotti, V., Lysozyme: A paradigmatic molecule for the investigation of protein structure, function and misfolding. *Clin. Chim. Acta*, 2005, **357**, 168–172.
108. Larson, J. L. and Miranker, A. D., The mechanism of insulin action on islet amyloid polypeptide fiber formation. *J. Mol. Biol.*, 2004, **335**, 221–231.
109. Carrell, R. W. and Lomas, D. A., Alpha1-antitrypsin deficiency – a model for conformational diseases. *N. Engl. J. Med.*, 2002, **346**, 45–53.
110. Radford, S. E., Gosal, W. S. and Platt, G. W., Towards an understanding of the structural molecular mechanism of $\beta(2)$-microglobulin amyloid formation *in vitro*. *Biochim. Biophys. Acta*, 2005, **1753**, 51–63.
111. Mussap, M. and Plebani, M., Biochemistry and clinical role of human cystatin C. *Crit. Rev. Clin. Lab. Sci.*, 2004, **41**, 467–550.
112. Mendes, H. F., van der Spuy, J., Chapple, J. P. and Cheetham, M. E., Mechanisms of cell death in rhodopsin retinitis pigmentosa: Implications for therapy. *Trends Mol. Med.*, 2005, **11**, 177–185.
113. Sandilands, A. *et al.*, Altered aggregation properties of mutant gamma-crystallins cause inherited cataract. *EMBO J.*, 2002, **21**, 6005–6014.
114. Otto, M., Hansen, T. and Kriegsmann, J., Images in cardiology: Intracellular cardiac amyloidosis in familiar amyloid polyneuropathy of the Portuguese type. *Heart*, 2002, **88**, 90.
115. Eriksson, L. and Westermark, P., Characterization of intracellular amyloid fibrils in the human choroid plexus epithelial cells. *Acta Neuropathol.*, 1990, **80**, 597–603.

ACKNOWLEDGEMENT. S.I.-T. acknowledges CSIR, New Delhi for research fellowship.

Received 14 July 2006; revised accepted 3 January 2007