Protein Aggregation: A Review

Ganesh V. Devkate[1*], Sharwaree R. Hardikar[2] and Rajendra N. Patil[3]

[1]Department of Pharmaceutics, Shivnagar Vidya Prasarak Mandal's College of Pharmacy, Malegaon (BK), Baramati-413115, India.
[2]Department of Pharmaceutics, Seth Govind Raghunath Sable College of Pharmacy, Saswad, Pune 412301, India.
[3]Department of Pharmaceutical Chemistry, Shivnagar Vidya Prasarak Mandal's College of Pharmacy, Malegaon (BK), Baramati-413115, India.

Authors' contributions

This work was carried out in collaboration between all authors. Author RNP designed the study, wrote the protocol and supervised the work. Authors GVD and SRH carried out all laboratories work and performed the statistical analysis. Author SRH managed the analyses of the study. Author GVD wrote the first draft of the manuscript. Author RNP managed the literature searches and edited the manuscript. All authors read and approved the final manuscript.

Review Article

ABSTRACT

Protein aggregation is arguably the most common and troubling manifestation of protein instability, encountered in almost all stage of protein drug development. Protein generally will tend to aggregate under variety of conditions. The extent of aggregation depends on many factors that can be classified as intrinsic (primary, secondary, tertiary or quaternary structure) and extrinsic (environment in which protein present or processed). These protein aggregate may exhibit less desirable characteristics like reduce or no biological activity. Although a variety of method has been used to inhibit protein aggregation, it involved structural (internal) and environmental (external) modifications.

_____

Corresponding author: E-mail: gvdevkate@gmail.com;

1. INTRODUCTION

Protein is a molecule composed of polymers of amino acids joined together by peptide bonds. It can be distinguished from fats and carbohydrates by containing nitrogen. Other components include carbon, hydrogen, oxygen, sulphur, and sometimes phosphorus.

A protein is a linear polymer built from about 20 different amino acids. The type and the sequence of amino acids in a protein are specified by the DNA in the cell that produces them. This sequence of amino acids is essential since it determines the overall structure and function of a protein.

A protein has several functions. It may serve as a structural material (e.g. keratin, collagene), as enzymes, as transporters (e.g. hemoglobin), as antibodies, or as regulators of gene expression.

A protein may be classified based on its form and main functions: It can be a globular protein like most enzymes, fibrous protein which are for structural role; and membrane proteins that serve as receptors or channels for polar or charged molecule to pass through the cell membrane [1].

1.1 Functions of Protein in the Human Body

Protein has a large number of important functions in the human body and in fact, the human body is about 45% protein. It's an essential macromolecule without which our bodies would be unable to repair, regulate, or protect themselves.

Protein has a range of essential functions in the body, including the following:

- Required for building and repair of body tissues (including muscle).
- Enzymes, hormones, and many immune molecules are proteins.
- Essential body processes such as water balancing, nutrient transport, and muscle contractions require protein to function.
- Protein is a source of energy.
- Protein helps keep skin, hair, and nails healthy.
- Protein, like most other essential nutrients, is absolutely crucial for overall good health.

Proteins are the main actioners in cells and in an entire organism. Without proteins the most basic functions of life could not be carried out. Respiration, for example, requires muscle contractions, and muscle contractions require proteins.

In last three decades therapeutics peptide and protein have raised in prominence a potential drug of the future. Biotechnology plays a key role in development of peptide and protein drug. A new series of peptide and protein based pharmaceutical have discover method to clone and express the cDNA encoding. The resent advances in large scale fermentation and purification processes and analytical characterization has further widened. The protein is available at much purer form in significant quantities, at reasonable coast. The problem of immunogenicity and antigenicity has also been considerably reduced. The peptide and protein drug produced by recombinant DNA technology are the exact replicas of that obtained from natural sources. Endogenously the protein is synthesized in small amount, normally act in transient fashion and are stabilized by the cellular and extracellular milieu. In contrast, the protein drug must be highly purified and concentrated and have a shelf life at least two year. Oral administration is limited due to enzymatic degradation. Only after some viable novel delivery system develops to improve their systemic bioavailability, these peptide and protein drug would be of therapeutic importance [1,2].

1.2 Stability of Protein

protein stability primarily in terms of the thermodynamic stability of a protein that unfolds and refolds rapidly, reversibly, cooperatively, and with a simple, two-state mechanism:
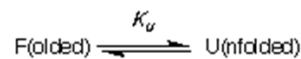
$$F\text{(olded)} \underset{}{\overset{K_u}{\rightleftharpoons}} U\text{(nfolded)}$$

Where K_u, is the equilibrium constant for unfolding.

The stability of the protein is simply the difference in Gibbs free energy, $\triangle G$, between the folded and the unfolded states. The only factors affecting stability are the relative free energies of the folded (G_f) and the unfolded (G_u)

states. The larger and more positive $\triangle G_u$, the more stable is the protein to denaturation.

$$\triangle G_u = G_u - G_f$$

The Gibbs free energy, G, is made up the two terms enthalpy (H) and entropy (S), related by the equation:

$$G = H - TS$$

Where T is the temperature in Kelvin [3-5].

Table 1. Some representative peptidal and proteinaceous drug with their potential function and biomedical application [1]

Peptide or protein drug (s)	Function or biomedical application
Cardiovascular active	
Angiotensin II antagonist	Lowering blood pressure
Bradykinin	Improving peripheral circulation
Captopril	Heart failure management
Tissue plasminogen activator	Dissolution of blood clot
CNS active	
Cholecystokinin	Suppressing appetite
B-endorphine	Relieving pain
Neuropeptide γ	Controlling feeding and drinking behavior
Nerve growth factore	Simulating nerve growth and repair
Gastrointestinal active	
Somatostatine	Reducing bleeding of gastric ulcer
Pancreatic enzyme	Digestive supplement
Immunomodulating	
Interferon	Enhancing activity of killer cell
Tumer necroising factor	Controlling polymorphonuclear function
Cyclosporine	Inhibiting function of T lymphocyte
Metabolic modulating	
Insulin	Treating diabetes mellitus
Vasopressine	Treating diabetes insipidus
Human growth hormone	Treating hypopituitary dwarfism

2. PROTEIN AGGREGATIONS

Protein aggregation means two or more protein molecule come together and form small mass like structure call aggregate. The definition of newly formed aggregate is still debatable. We propose that protein aggregates be defined as any protein species in non-native states and whose sizes are at least twice as that of the native protein. Proteins generally will tend to aggregate under a variety of environmental conditions in comparison with small drug molecules [2]. The extent of aggregation is dependent on many factors that can be broadly classified as intrinsic (primary, secondary, tertiary or quaternary structure) or extrinsic (environment in which protein is present, processing conditions, etc). These protein aggregates may exhibit less desirable characteristics like reduced or no biological activity, potential for immunogenicity or other side effects. The scientific and regulatory challenges in overcoming protein aggregation issues remain one of the major hurdles in commercialization of such drug products. Thus, it has been one of the most important areas of intensive research and development in both academia and biopharmaceutical companies to identify the cause(s) of protein aggregation and to control protein aggregation to an acceptable level. Among these physical and chemical deterioration pathways, protein aggregation is arguably the most common and troubling manifestation of protein instability, almost in all phases of protein drug development. Presence of any insoluble aggregates in a protein pharmaceutical is generally unacceptable for product release. Aggregation is a general term that encompasses several types of interactions or characteristics. Aggregates of proteins may arise from several mechanisms and may be classified in numerous ways, including soluble/insoluble, covalent/non covalent, reversible/irreversible, and native/ denatured. For protein therapeutics, the presence of aggregates of any type is typically considered to be undesirable because of the concern that the aggregates may lead to an immunogenic reaction (small aggregates) or may cause adverse events on administration [3-5].

2.1 Protein Aggregation Pathways and Characteristics

Proteins aggregate through different mechanisms/pathways. The major pathways are shown in Fig. 1 and can be roughly divided into: (1) aggregation through unfolding intermediates and unfolded states (pathway 1); (2) aggregation through protein self association (pathway 2a) or chemical linkages (pathway 2b); and (3) aggregation through chemical degradations (pathway 3).

2.1.1 Aggregation through unfolding intermediates and unfolded states

Under normal conditions, native proteins in solution are in equilibrium with a small amount of unfolding intermediates—I states, which are further in equilibrium with the completely unfolded/denatured states—D states in Fig. 1. Depending on the degree of protein unfolding, the unfolding intermediates may be further divided into more native like intermediates and more unfolded like intermediates. The latter are also referred to as molten globules for globular proteins, commonly used in the description of the protein folding/unfolding process.

In terms of the physical aggregation process, significant evidence suggests that the poorly populated protein folding/unfolding intermediates are precursors of the aggregation process. This is because these intermediates expose more hydrophobic patches and have higher flexibility relative to the folded state. Completely folded or unfolded proteins, in contrast, do not aggregate easilysince the hydrophobic side chains are either mostly buried out of contact with water or randomly scattered. Recent simulation studies indicate that aggregation of native proteins with accessible hydrophobic areas on its surface is intensified when the solution conditions favor the

partially unfolded conformation as opposed to either the native or fully unfolded conformations. Interaction (attraction) of these intermediates leads to formation of protein aggregates. Dimmers, trimmers, which maintain the native-like state, will not fall under the definition of aggregates. Another similar term is oligomers or critical oligomers. These terms are often used for the description of early protein or peptide fibrillation process, also referred to as fibrillogenesis or amyloidogenesis [6].

2.1.2 Aggregation through self association

Many proteins can directly associate into protein aggregates physically from the native (folded) state (pathway 2a in Fig. 1) without going through the intermediate state (I). Such associations can be simply electrostatic or both electrostatic and hydrophobic depending on the experimental conditions. Other weaker forces (Van der Waals forces) can also initiate the association process. Protein self-association may or may not be accompanied by subtle conformational changes. Such association often leads to formation of reversible aggregates which can be considered the precursors of irreversible aggregates. If protein self-association is the rate-limiting step, protein aggregation resulting from self association would be association-limited [7].
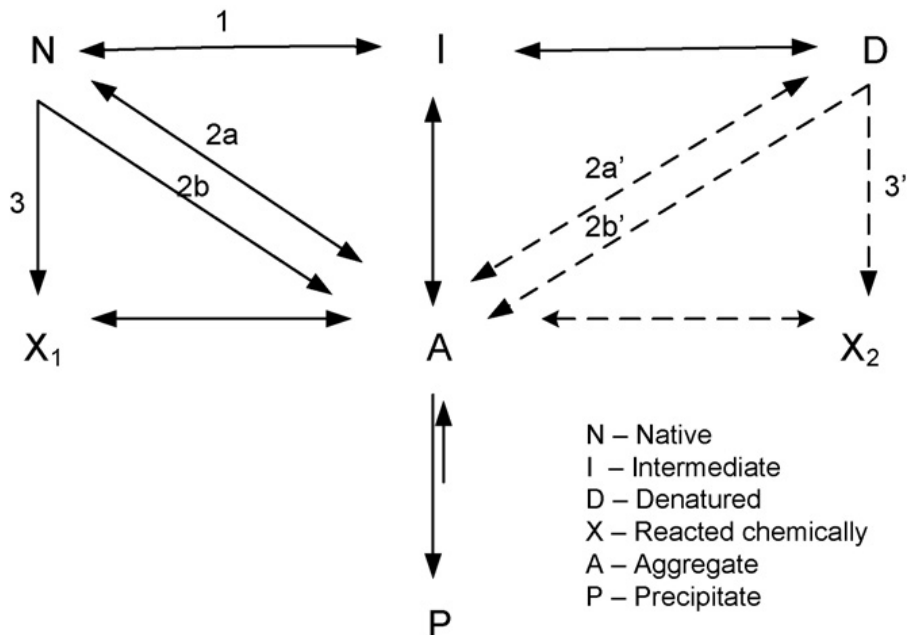

N – Native
I – Intermediate
D – Denatured
X – Reacted chemically
A – Aggregate
P – Precipitate

Fig. 1. Graphic illustration of protein aggregation

2.1.3 Aggregation through chemical linkage

Many chemical degradations directly crosslink protein chains and thereby lead to aggregation (pathway 2b in Fig. 1). The most common cross-linking reaction is the intermolecular disulfide bond formation/exchange. Formation of disulfide linkages can further promote physical aggregation of proteins [8].

2.1.4 Aggregation through chemical degradation

The last major aggregation pathway (pathway 3 in Fig. 1) is the chemical degradation induced protein aggregation. Many chemical degradation pathways have been shown to increase the aggregation tendency of proteins. Chemical degradation often changes the physical properties of a protein, such as protein hydrophobicity or association tendency, secondary/tertiary structures, and the thermodynamic/kinetic barrier to protein unfolding [9].

Chemical degradation may occur in a number of different ways. This type of degradation affects the primary sequence and may also lead to significant changes in the higher order structure. Examples of chemical degradation include deamidation, oxidation, isomerization, clipping/fragmentation, and cross-linking [10].

Deamidation is probably the most common type of chemical degradation encountered in mAb-based biotherapeutics. The reaction is favored at neutral and basic pH. There are examples of deamidation occurring at lower pH, however that has been reported to occur primarily through a mechanism independent of succinimide formation; for example, deamidation of Asn in the A chain of insulin is favored at pH < 5, which is mediated via the formation of cyclic anhydride intermediate. There are factors other than pH that influences the rate of deamidation, e.g., sequence and local structure (steric effect). Amino acids present at the carboxyl end of Asn have been reported to influence the rate of deamidation; the decrease in rate has been correlated with increase in size and branching of the side chain [11,12].

Isomerization is another common method of chemical degradation that shares the common outcome with deamidation. It can occur directly from Asp residue or from the succinimide intermediate. Similarly to deamidation,

isomerization is favored at neutral and basic pH, and the rate has been reported to be influenced by the steric effect [13].

Hydrolysis of a peptide bond leads to fragmentation of the protein. Some peptide bonds, such as Asp-Gly and Asp-Pro, are sensitive to hydrolysis. Asp-Y sequence has been reported to be 100 times more sensitive to hydrolysis than any other peptide bonds. At times, hydrolysis is the natural, subsequent reaction following deamidation of Asn residue, as is observed with insulin [14,15].

Chemical degradation may also result from cross-linking, which may or may not be mediated by the formation of disulfide bond. When cross-linking is disulfide bond-mediated, the reaction occurs either by the formation of a new disulfide bond or by disulfide bond exchange. Intra-molecular disulfide linkage may lead to change in tertiary structure, whereas inter-molecular (or inter-domain) disulfide linkage may lead to change in quaternary structure or formation of covalent aggregates. At higher pH, the formation of reactive thiolate ion (S−), from the thiol group (–SH) of cys residue, is favored, which may increase the probability of disulfide linkage formation. There are instances of covalent cross-linking that may occur between different amino acid residues besides Cys, which, unlike disulfide linkage, are non-reducible [16].

3. TYPE OF PROTEIN AGGREGATION

Protein aggregation can occur via physical and chemical mechanisms.

3.1 Physical Aggregation

In protein aggregations protein may aggregate simply by physical association with one another without any change in primary structure or by formation of a new covalent bond.

3.2 Chemical Aggregation

In protein aggregation formation of bond by directly crosslink protein or indirectly alter aggregation tendency of the original protein.

Both above mechanism can occur simultaneously to a protein and may lead to formation of either soluble or insoluble aggregates, depending on the protein, environmental condition, and stage of the aggregation process. For example, insulin can

undergo both physical aggregation process, leading to formation of either soluble hexamers or insoluble fibrils and chemical aggregation process, leading to formation of either soluble dimmer via cyclic anhydride intermediate or insoluble disulfide-bonded aggregates [17-19].

4. MECHANISM OF PHYSICAL PROTEIN AGGREGATION

4.1 Folding/Unfolding Intermediates and Protein Aggregation

Traditional view of protein aggregation is the association of the unfolded state (s) of proteins [20,21]. However, there is over whelming evidence that protein folding/unfolding intermediates are precursors in protein aggregation [22], even though the intermediates are usually not stable and poorly populate. In contrast, completely folded or unfolded proteins do not aggregate easily as the hydrophobic side chains are either mostly buried out of contact with water, or randomly scattered. It is the patches of continuous hydrophobic groups in the folding/unfolding intermediates that initiate the aggregation process [23].

The aggregation process can be described as scheme (1), where proteins form reversible unfolded intermediates, which then form reversible unfolded proteins or irreversible/reversible aggregates.
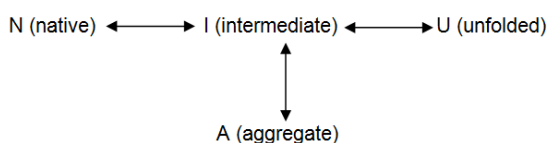
The intermediate state (I) is equivalent to the aggregation-competent state (A) or transition state [24].

4.2 Mechanism of Chemical Protein Aggregation

There are many chemical reactions that can directly crosslink protein chains or change the hydrophobicity of a protein, indirectly changing its aggregation behavior. Chemical reactions that cause aggregation involve disulfide bond formation/exchange, non-disulfide cross linking pathways, oxidation, maillard reaction [25].

5. CHARACTERIZATION OF PROTEIN AGGREGATION

5.1 Morphological Method

There is no consistent definition of what is meant by a "soluble" aggregate, so working definitions are often employed. Soluble aggregates refer to those that are not visible as discrete particles and that may not be removed by a filter with a pore size of 0.22 µm. Conversely, insoluble aggregates may be removed by filtration and are often visible to the unaided eye. Both types of aggregates may be problematic for the development of a therapeutic protein. There are clear guidelines and limitations on the number of particles ≥10 µm and ≥25 µm in size that may be present in pharmaceutical preparations. However, the levels of soluble aggregates such as dimers and trimers that are acceptable are not well defined.

Protein aggregates come in different shapes and sizes, even for aggregates of a single protein. In most cases, protein aggregates are amorphous. For many small proteins or peptides, fibrils are a common form of aggregates such as insulin [26]. β-amyloid peptide and its fragments [27]. Many factors may influence the morphology of protein aggregates. Major ones include the primary sequence of the protein. A single mutation in a protein or peptide can change the morphology of aggregates.

5.2 Analytical Method

The amount, type, and size of the aggregates in biopharmaceuticals can have important consequences for the safety and efficacy of biopharmaceuticals

The primary tools we use for aggregation analysis are:

- analytical ultracentrifugation (AUC), including both sedimentation velocity and sedimentation equilibrium
- both static light scattering (SEC-MALS) and dynamic light scattering (DLS)
- native gel electrophoresis
- Proteostat protein aggregation analysis

5.2.1 ProteoStat™ protein aggregation assay [28]

ProteoStatTM Protein aggregation assay provides a simple, homogenous assay format for

monitoring peptide and protein aggregation. The assay can be employed to streamline protein processing and optimize formulation procedures during from protein expression through manufacturing. Relative to conventional protein aggregation dyes, such as Thioflavin T.

ProteoStat™ detection reagent can detect a broader range of different protein. The assay yields a much brighter, provides at least 2 orders of magnitude linear dynamic, and offers superior performance across a broad range of values (4~10) and buffer compositions. Sensitivity of this assay is in the submicro molar so that less than 1% protein aggregate is detectable in a protein solution. The assay Z'factor scores greater than 0.5 providing quantitative analysis of protein aggregation in a robust and high-throughput fashion. Lyophilized native and aggregated proteins are included in the kit as negative and positive controls for monitoring changes in protein aggregation status.

5.2.1.1 Assay mechanism

Proteins misfold and aggregate under stress from pH levels, temperature, buffers, light, oxygen, mechanical agitation and freeze/thaw cycles. Applications performed with the ProteoStat™ Protein aggregation dye can quantify aggregates in protein samples. ProteoStat™ is a molecular rotor dye that rotates like a propeller in the absence of protein aggregates and does not fluoresce. When the dye binds to the aggregate, it is immobilized slowing down the rotational movement and causing the dye to fluoresce. Unlike other environmentally sensitive dyes that measure unfolding and exposure of hydrophobic regions of proteins, there is little interference from hydrophobic compounds or detergents present at normal concentrations.
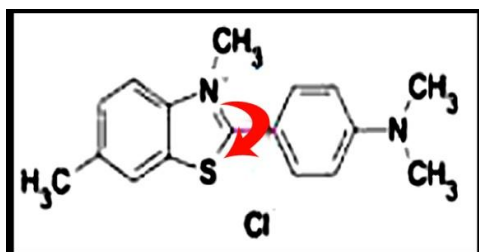

Fig. 2. Dye Thioflavin T rotate along with protein (red arrow) in absence of aggregation
Thioflavin T: early prototype dye in the design of ProteoStat™ assay which also rotates around a single bond (red arrow) in the absence of protein aggregates

Fig. 3. Dye Thioflavin T immobilized when bound to the aggregate and begins to fluoresce

5.2.2 Rayleigh scattering measurements

The optical properties of a solution are a result of the ability of molecules or particles present therein to absorb and scatter light. Rayleigh's theory of light scattering states that particles with diameter lesser than the wavelength of incident light are able to scatter light. The upper limit of the size of particles is about one-tenth of the wavelength, which implies that scattering does not take the shape of the particle into account. Light scattering is essentially used in the investigation of protein aggregation, and it can be monitor educing a fluorescence spectrophotometer by setting identical excitation and emission wavelengths, usually at 350 nm. This method has been considerably employed to detect the formation of protein aggregates [29,30].

5.2.3 Circular dichroism spectroscopy (CD)

Circular dichroism (CD) spectroscopy is a very important tool in structural biology and protein chemistry [31]. For determination of the secondary and tertiary structure of proteins and peptides, far-ultraviolet (UV) CD (190–250 nm) and near-UV CD (250–320 nm) are used, respectively. Aggregation of amyloid proteins into protofibrils and fibrils is accompanied by the formation of β sheet conformation, which gives the characteristic minima at 215–218 nm in the far-UV CD region. Qualitative information on amyloid assemblies and kinetics of conformational transition associated with aggregation can be obtained by CD [32,33].

5.2.4 Intrinsic fluorescence measurements

Intrinsic fluorescence parameters such as fluorescence intensity (FI) and emission maxima are very sensitive to the changes in the structural dynamics and polarity of protein chromophores such as Trp and Tyr; therefore, they can be used to study protein folding and assembly. The intrinsic fluorescence of aromatic residues has been used to probe conformational dynamic

sand the self-assembly of several amyloidogenic proteins such as prion, α -synuclein and immunoglobulin G (IgG) light chains [34-36].

5.2.5 Transmission electron microscopy (TEM)

The direct visualization of the protein self-assembly into oligomers, protofibrils and mature fibrils has been facilitated by high-resolution microscopic techniques such as transmission electron microscopy (TEM). It provides only qualitative information about the aggregates. Protein aggregates such as insulin, lysozyme and α-synuclein have been characterised by TEM [37,38].

5.2.6 Atomic force microscopy (AFM)

Biological sciences have exploited the ability of AFM to give 3D structural details of numerous protein based samples or fibrils with very high resolution. AFM does not require special sample preparation that would otherwise damage the sample. Its principle involves scanning of the tip over the samples, and the laser beam detects the deflection. It can be used in the examination of supramolecular structures of living cells and the growth of amyloid fibrils. AFM provides both qualitative and quantitative information at the nanometer level including length, width and surface characteristic of protein aggregates [39, 40].

5.2.7 High-resolution transmission electron microscopy (HRTEM)

High-resolution transmission electron microscopy (HRTEM) is a useful technique for monitoring the formation of amyloid fibrils in vitro and providing images of prefibrillar aggregates, circular species

and mature fibrils at high resolution. Moreover, it also provides information on protein aggregates formed by other misfolding pathways such as amorphous aggregate formation. HRTEM has the advantage of quick performance, which allows researchers to assess the fibril formation rapidly. HRTEM confirms the fibrillar morphology of protein aggregates, but it does not confirm the presence of the characteristic cross- β-sheet structure. Thus, ancillary techniques are required in addition to HRTEM for proving the fibrillar morphology or any other configuration of amyloid. It provides greater resolution than other techniques such as confocal and scanning electron microscopy (SEM) [41].

5.2.8 Field-emission scanning electron microscopy (FESEM)

Amyloid fibrils have a characteristic cross-β-sheet structure that may be formed in both *in vitro* and *in vivo* environments. An SEM is used to obtain the 3D image of an amyloid, but it is not as prevalent as other techniques although it possesses relatively higher resolution up to the nanometer level. In SEM, secondary electrons emitted from the specimen are used to construct the final image. In field-emission scanning electron microscopy (FESEM), a field-emission cathode provides narrower probing beams with high as well as low electron energy, which is responsible for the enhanced resolution and minimised sample damage of images [42,43].

5.2.9 Fluorescence microscopy

Fluorescence microscopy is an invaluable tool for detection and characterization of protein aggregates. Dyes such as thioflavin T, Congo red and Nile Red are used as extrinsic

Table 2. Analytical techniques for determination of protein aggregation

Categories	Individual techniques	Applications	Protein examples
Calorimetry	Differential scanning calorimetry (DSC)	Thermal protein unfolding/aggregation	FVIII SQ
Centrifugation	Analytical centrifugation	Size and shape estimation	Insulin
Chromatography	SEC-HPLC	Size estimation and Quantitation	Factor IX
Electrophoresis	SDS-PAGE	Size estimation and mechanistic probing	IL-2
	Native PAGE	Aggregation process and mechanistic probing	IL-2
Light scattering	Static light scattering	Size and shape estimation	Amyloid peptide
	Dynamic light scattering	Size distribution of soluble Aggregates	Deoxy hemoglobin
	Light scattering/obscuration	Size estimation and relative Distribution	Deoxy hemoglobin

fluorophores to monitor the protein aggregates; Congo red shows an apple green birefringence, and fluorescence of the thioflavin T increases in the presence of amyloid deposits. Fluorescence microscopy has been efficiently utilised even at high protein concentrations, minimising changes in the protein local environment for the detection and characterisation of aggregates. Nile Red is a low molecular weight phenoxazone dye, binds to hydrophobic surfaces of proteins and exhibits a strong fluorescence [44]. Recent approaches include confocal and two or multiphoton fluorescence microscopy and total internal reflection. Confocal fluorescence microscopy is well suited for 3D resolution investigation. Two- or multiphoton fluorescence microscopy has been used for the analysis of UV-excitable fluorescence, which is not possible with confocal fluorescence microscopy. Total internal reflection fluorescence microscopy (TIRFM) has been used to study the plasma membrane and single-molecule methods such as fibril progress [45].

6. INHIBITION OF PROTEIN AGGREGA-TION

Protein aggregation can be inhibited either by modifying protein's structure (internally) or by changing protein's environmental properties (externally). Modification can be made by site-directed mutagenesis or by chemical reactions. The key issue in structural modification for stabilization is to preserve the protein activity, as the modified protein can easily lose activity [46]. Many structurally modified proteins have been shown to have reduced tendency to aggregate, including single amino acid mutants of human macrophage inflammatory protein (hMIP)-1, thymidylate synthase, monosubstituted, disubstituted, and trisubstituted insulins with *p*-succinamidophenylglucopyranoside (SAPG) [47].

Changing protein's environmental properties often lead to inhibition of protein aggregation. A common method is to add excipients/additives in the protein preparation. A variety of excipients/additives are available for use, such as sugars and polyols, amino acids, amines, salts, polymers, and surfactants. These additives stabilize proteins by preferential interactions, a widely accepted concept of protein stabilization. Other mechanisms are also operable, such as increased rate of protein folding, reduction of solvent accessibility and conformational mobility, and increase in solvent viscosities [47].

6.1 Denature Concentration

In close relation to protein concentration, the denaturant concentration strongly affects both the rate and extent of on-pathway folding and off-pathway aggregation during refolding. This is because denaturants at different concentrations affect protein solubility to different degrees. Low concentrations of denaturants may decrease the solubility of a protein as the denaturants convert native to denatured and aggregation-prone species, while high denaturant concentrations should solubilize the protein as it weakens the protein–protein attractions in water. It seems that protein aggregation under such conditions results from the favorable formation of misfolded intermediates [48].

6.2 Use of Additives

The well-known function of molecular chaperones in bacteria led to the use of such molecular systems (such as DnaK-DnaJ-GrpE and GroEL-GroES) to inhibit protein aggregation during protein refolding [49]. They suppress protein aggregation by alternate binding and releasing of the folding intermediates [50]. However, two major limitations seem to be associated with a routine use of these folding modulators production cost and the extra purification burden. Therefore, other additives have been explored to achieve the same goal during refolding.

The term "artificial chaperone" has been used to describe such non-protein additives, which have chaperone-like effect. Such additives include detergents and cyclodextrins. Detergents were found to suppress formation of protein aggregates (rather than dissolve aggregates) during protein refolding [51]. Similarly, cyclodextrins were shown to prevent formation of protein aggregates during renaturation of carbonic anhydrase. To increase the efficiency, a combination of such additives was used for the refolding of lysozyme [52,53].

6.3 Processing of Proteins

Many routine processing steps can lead to protein aggregation to various degrees, such as heat-treatment, filtration, shaking/shearing, atomization, freezing, drying, and reconstitution. Proper care in that step can help to control protein aggregation.

6.4 Heat Treatment

One of the viral inactivation methods in protein purification is heat treatment, which is usually performed at 60°C. Unfortunately; many proteins aggregate at this temperature and thus, need proper protection. To protect a protein, the first step is to choose a stable solution pH, as proteins are usually stable in a narrow pH range [54]. Although many buffering agents are available for pH adjustment, the aggregation behavior of proteins can be significantly different in different buffer systems and at different buffer concentrations [55].

Another two critical parameters in controlling protein aggregation during thermal treatment are protein concentration and ionic strength. Generally, increasing protein concentration increases the rate of protein aggregation due to the increased chances of intermolecular interactions [56]. The effect of ionic strength is very much protein dependent. The salt type can also make a significant difference in protein aggregation. In addition, a stabilizer(s) is often added into a protein preparation to inhibit protein aggregation under thermal treatment. Polymers may inhibit protein aggregation through one or more of these properties: surface activity, preferential exclusion, steric hindrance of protein–protein interactions, and increased viscosity limiting protein structural movement. PEGs of different molecular weights were found to inhibit aggregation of recombinant human keratinocyte growth factor (rhKGF) at 45°C and to inhibit thermally induced aggregation [57].

6.5 Shaking and Shearing

Shaking creates air/water interface. The hydrophobic property of air relative to water induces protein alignment at the interface, maximizing exposure of the hydrophobic residues to the air and initiating aggregation [58]. Shearing also exposes the hydrophobic areas of proteins, causing aggregation. Many proteins easily aggregate during shaking or shearing, such as rFXIII [59,60], hGH [61], and insulin [62].

To protect proteins from shaking/shearing-induced aggregation, surfactants are most commonly used. They inhibit protein aggregation by accumulating, competitively with proteins, at hydrophobic surfaces/ interfaces and/or by binding directly to proteins. Certain surfactants such as poloxamers may also increase the viscosity of a protein solution, restraining the motion of protein backbone to inhibit protein aggregation [63].

6.6 Care in Freezing Step

Freezing of a protein solution may cause protein aggregation due to one or more of the following freezing induced stresses: low temperature, solute concentration, formation of ice-water interfaces, and potential pH changes or phase separation [64]. Understanding the cause(s) of protein aggregation is the first step in aggregation prevention/inhibition during freezing. For proteins that denature during freezing due to their interaction with the ice-water interface, a slower rate of freezing should inhibit aggregation of these proteins as a smaller ice-water interface will be generated [65-67]. On the other hand, a lower freezing rate may facilitate crystallization of other solution components, promoting potential Crystallization-induced protein aggregation [68]. A frequent cause of protein aggregation during freezing is the buffer-induced pH change during the freezing process. In this case, selection of a suitable buffer would prevent/inhibit protein aggregation. Many additives can be used to inhibit freezing induced protein aggregation. They may inhibit protein aggregation by preferential interaction and/or other mechanisms such as increased solution viscosity, steric hindrance of protein–protein interactions or suppression of pH changes during freezing [69].

6.7 Proper Care in Drying Step

The amount of water covering the surface of a protein in a fully hydrated state is around 0.3 g/g protein [70], while the water content of a dried protein product is usually less than 10%. Therefore, the drying process removes part of the hydration layer, which may disrupt the native state of a protein and cause protein aggregation. To protect the protein from dehydration induced aggregation, excipients, capable of forming hydrogen bonds with the protein, are added to replace water so that the hydrogen bonding requirements of polar groups on the protein surface can be satisfied [71]. These excipients serve as water substitutes. The formation of hydrogen bonding has been demonstrated by IR spectroscopy between carbohydrates and many freeze-dried proteins. These excipients are preferably amorphous. In fact, the formation of an amorphous glass during dehydration has been considered as a prerequisite for protein stabilization [72].

The commonly used stabilizing agents during drying are sugars. Both trehalose and sucrose inhibit aggregation of many proteins during lyophilization such as IL-6 and human factor XIII. Some polymers are also effective, as they can form a glass with a high transition temperature. Examples include hydroxyethyl cellulose at 1% in the complete inhibition of a FGF aggregation and PEG In the inhibition of human FXIII aggregation during lyophilization. Probably due to polymer-associated properties, many proteins protect themselves from aggregation during drying [73, 74,58].

6.8 Other Techniques

To prevent protein aggregation during refolding, a variety of novel techniques or methods have been tried. These include refolding of proteins on a size-exclusion column during elution, use of high hydrostatic pressure [6,57].

7. CONCLUSIONS

A common phenomenon of protein instability is the formation of aggregates, which can be soluble or insoluble, non-covalent or covalent, and reversible or irreversible. Although protein sequence determines the behavior of protein aggregation, many external factors play a critical role in controlling/affecting protein aggregation, including temperature, pH, ionic strength, surface adsorption, shearing, shaking, presence of metal ions, organic solvents and additives, protein concentration, purity and morphism, pressure, freezing and drying. Chemical transformations can lead to direct or indirect protein aggregation, such as disulfide bond formation/exchange,non-disulfide cross linking, transamidation, and oxidation. Although many analytical techniques are available in monitoring protein aggregation, their differences in quantization, accuracy, sensitivity, and ease of operation require careful selection of some of these methods in monitoring all aspects of protein aggregation.

Protein aggregation occurs readily in almost all the biopharmaceutical processes, especially during fermentation, refolding, purification, formulation, and storage. Recognition of different causes of protein aggregation in these processes is the basis for selection of different techniques/methods for aggregation inhibition. Although protein aggregation can be inhibited effectively by modifying the primary sequence of a protein, this method is seldom used mainly because of the unpredictable

consequences of structural modifications. Therefore, inhibition of protein aggregation is commonly achieved by changing the immediate environment of the protein. One particularly useful and simple method is the use of compatible excipients(s). The often used protein stabilizing excipients include sugars, polyols, surfactants, salts, PEGs, polymers, metal ions, and amino acids.

Although there are a variety of methods and a large pool of excipients for use to inhibit protein aggregation, a satisfactory level of aggregation inhibition has not been achieved for many proteins. Trial and error has been the main stay in the prevention or inhibition of protein aggregation in most biopharmaceutical processes. This is partly due to our lack of a clear understanding of the protein aggregation process. Therefore, intensive and rigorous investigation of mechanisms of protein aggregation is urgently needed for an ultimate solution of protein aggregation in biopharmaceutics.

COMPETING INTERESTS

Authors have declared that no competing interests exist.

REFERENCES

1. Vyas SP, Khar RK. Controlled drug delivery controlled drug delivery concept and advances published by Jain M.K. Vallabh Prakahan first print. (Reprint). 2002;2008:504-510.
2. Modler AJ, Gast K, Lutsch G, Damaschun G. Assembly of amyloid protofibrils via critical oligomers—a novel pathway of amyloid formation. J. Mol. Biol. 2003;325: 135–148.
3. Chi EY, Krishnan S, et al. Physical stability of proteins in aqueous solution: Mechanism and driving forces in nonnative protein aggregation. Pharm. Res. 2003; 20:1325–1336.
4. Arakawa JS, Philo, Kita Y. Kinetic and thermodynamic analysis of thermal unfolding of recombinant erythropoietin. Biosci. Biotechnol. Biochem. 2001;65:121-1327.
5. Morris AM, Watzky MA, Finke RG. Protein aggregation kinetics, mechanism and curve fitting: A review of the literature. Biochim. Biophys. Acta. 2009;1794:375–397.

6. Zhang L, Lu D, Liu Z. How native proteins aggregate in solution: A dynamic Monte Carlo simulation. Biophys. Chem. 2008; 133:71–80.

7. Weiss WF, Young TM, Roberts CJ. Principles, approaches and challenges for predicting protein aggregation rates and shelf life. J. Pharm. Sci. 2009;98:1246–1277.

8. Cabra V, Vazquez-Contreras E, Moreno A, Arreguin-Espinosa R. The effect of sulfhydryl groups and disulphide linkage in the thermal aggregation of Z19 alpha zein. Biochim. Biophys. Acta. 2008;1784:1028–1036.

9. Rosenfeld MA, Leonova VB, Konstantinova ML, Razumovskii SD. Self-assembly of fibrin monomers and fibrinogen aggregation during ozone oxidation. Biochemistry (Mosc.). 2009;74: 41–46.

10. Wang W, Singh S, Zeng DL, King K, Nema S. Antibody structure, instability and formulation. J. Pharm. Sci. 2007;96:1–26.

11. Tyler-Cross R, Schirch V. Effects of amino acid sequence, buffers and ionic strength on the rate and mechanism of deamidation of asparagine residues in small peptides. J. Biol. Chem. 1991;266:22549–22556.

12. Son K, Kwon C. Stabilization of human epidermal growth factor (hEGF) in aqueous formulation. Pharm. Res. 1995; 12:451–454.

13. Cacia J, Keck R, Presta LG, Frenz J. Isomerization of an aspartic acid residue in the complementarily determining regions of a recombinant antibody to human IgE: Identification and effect on binding affinity. Biochemistry. 1996;35:1897–1903.

14. Powell MF. A compendium and hydropathy/flexibility analysis of common reactive sites in proteins reactivity at Asn, Asp, Gln, and Met motifs in neutral pH solution. In Formulation, Characterization, and Stability of Protein Drugs; Pearlman, R., Wang, Y.J., Eds.; Plenum Press: New York, NY, USA. 1996;1–140.

15. Brange J, Langkj L, Havelund S, Volund A. Chemical stability of insulin. 1. Hydrolytic degradation during storage of pharmaceutical preparations. Pharm. Res. 1992;9:715–726.

16. Tous GI, Wei Z, Feng J, Bilbulian S, Bowen S, Smith J, Strouse R, McGeehan P, Casas-Finet J, Schenerman MA. Characterization of a novel modification to monoclonal antibodies: Thioether cross-link of heavy and light chains. Anal. Chem. 2005;77:2675–2682.

17. Sluzky V, Klibanov AM, et al. Mechanism of insulin aggregation and stabilization in agitated aqueous solutions. Biotechnol. Bioeng. 1992;40:895–903.

18. Costantino HR, Firouzabadian L, et al. Protein spray-freeze drying. Effect of atomization conditions on particle size and stability. Pharm. Res. 2000;17:1374–1383.

19. Darrington RT, Anderson BD. Evidence for a common intermediate in insulin deamidation and covalent dimer formation-effects of pH and aniline trapping in dilute acidic solutions. J. Pharm. Sci. 1995;84: 275–282.

20. De Young LR, Dill KA, et al. Aggregation and denaturation of a pomyoglobin in aqueous urea solutions. Biochemistry. 1993;32:3877–3886.

21. Stigter D, Dill KD. Theory for protein solubilities. Fluid Phase Equilib. 1993; 82:237–249.

22. Fields G, Alonso D, et al. Theory for the aggregation of proteins and copolymers. J. Phys. Chem. 1992;96:3974–3981.

23. Uversky VN, Karnoup AS, et al. Association of partially folded intermediates of staphylococcal nuclease induces structure and stability. Protein Sci. 1999;8:161–173.

24. Krishnamurthy R, Manning MC. The stability factor: Importance in formulation development. Curr. Pharm. Biotechnol 2002;3:361–371.

25. Apenten RK. Protein stability function relations: Beta- lacto globulin a sulphydryl group reactivity and its relationship to protein unfolding stability. Int. J. Biol. Macromol. 1998;23:19–25.

26. Schuler B, Rachel R, et al. Formation of fibrous aggregates from a non-native intermediate: The isolated P22 tailspike beta-helix domain. J. Biol. Chem. 1999; 274:18589–18596.

27. Tomski SJ, Murphy RM. Kinetics of aggregation of synthetic b-amyloid peptide. Arch. Biochem. Biophys. 1992;294:630–638.

28. ProteoStat™ protein aggregation assay Pdf. Enzolifescience. International Edition 1-Aug-2010.

29. Xu R. Particle characterization: Light scattering methods. Springer Science & Business Media. 2001;13:1–399.

30. Khurana R, Gillespie JR, Talapatra A, Minert LJ, Ionescu-Zanetti I, Millett AL.

Fink, Partially folded intermediates as critical precursors of light chainamyloid fibrils and amorphous aggregates. Biochemistry. 2001;40:3525–3535.

31. Tang Y, Cohen AE. Enhanced enantioselectivity in excitation of chiral molecules by superchiral light. Science. 2011;332:333–336.

32. Alam P, Rabbani G, Badr G, Badr BM, Khan RH. The surfactant-induced conformational and activity alterations in rhizopus niveus lipase. Cell Biochem. Biophys. 2015;71:1199–1206.

33. Alam P, Chaturvedi SK, Anwar T, Siddiqi MK, Ajmal MR, Badr G, Mahmoud MH, Hasan Khan R. Biophysical and molecular docking insight into the interaction of cytosine D-arabinofuranoside with human serum albumin. J. Lumin. 2015;164:123–130.

34. Chaturvedi SK, Ahmad E, Khan J, Alam P, Ishtikhar M, Khan R. Elucidating the interaction of limonene with bovine serum albumin: A multi technique approach. Mol. Biosyst. 2015;11:307–316.

35. Dusa A, Kaylor J, Edridge S, Bodner N, Hong DP, Fink AL. Characterization of oligomers during α-synuclein aggregation using intrinsic tryptophan fluorescence. Biochemistry. 2006;45:2752–2760.

36. Serno T, Carpenter JF, Randolph TW, Winter G. Inhibition of agitation induced aggregation of an IgG-antibody by hydroxypropyl betacyclodextrin. J. Pharm. Sci. 2010;99:1193–1206.

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

38. Bozzola JJ, Russell LD. Electron microscopy: Principles and techniques for biologists. Jones & Bartlett Learning; 1999.

39. Gregoire S, Irwin J, Kwon I. Techniques for monitoring protein misfolding and aggregation *in vitro* and in living cells. Korean J. Chem. Eng. 2012;29:693–702.

40. Muller DJ, Dufrene YF. Atomic force microscopy: A nanoscopic window on the cell surface. Trends Cell Biol. 2011; 21:461–469.

41. Hill AF, Barnham KJ, Bottomley SP, Cappai R, Gras S, Waddington L, Goldie K. Transmission electron microscopy of amyloid fibrils protein folding. Misfolding, and Disease. Humana Press. 2011;752: 197–214.

42. Seiler H. Secondary electron emission in the scanning electron microscope. J. Appl. Phys. 1983;54:R1–R18.

43. Hughes L, Hawes C, Monteith S, Vaughan S. Serial block face scanning electron microscopy the future of cell ultra structure imaging. Protoplasma. 2014;251:395–401.

44. Demeule R, Gurny BI, Arvinte T. Detection and characterization of protein aggregates by fluorescence microscopy. Int. J. Pharm. 2007;329:37–45.

45. Patil SM, Mehta A, Jha S, Alexandrescu, AT. Heterogeneous a α-mylin fibril growth mechanisms imaged by total internal reflection fluorescence microscopy. Biochemistry. 2011;50:2808–2819.

46. Knappik A, Plhckthun A. Engineered turns of a recombinant antibody improve its *in vivo* folding. Protein Eng. 1995;8:81–89.

47. Baudys M, Uchio T, et al. Physical stabilization of insulin by glycosylation. J. Pharm. Sci. 1995;84:28–33.

48. Hevehan D, De Bernardez-Clark E. Oxidative renaturation of lysozyme at high concentrations. Biotechnol. Bioeng. 1997; 54:221–230.

49. King J. Refolding with a piece of the ring. Nat. Biotechnol. 1997;15:514–515.

50. Thomas JG, Ayling A, et al. Molecular chaperones, folding catalysts, and the recovery of active recombinant proteins from *E. coli.* To fold or to refold. Appl. Biochem. Biotechnol. 1997;66:197–238.

51. Wetlaufer DB, Xie Y. Control of aggregation in protein refolding: A variety of surfactants promote renaturation of carbonic anhydrase II. Protein Sci. 1995;4:1535–1543.

52. Rozema D, Gellman SH. Artificial chaperone assisted refolding of carbonic anhydrase B. J. Biol. Chem. 1996a;271: 3478–3487.

53. Rozema D, Gellman SH. Artificial chaperone assisted refolding of denatured-reduced lysozyme: Modulation of the competition between renaturation and aggregation. Biochemistry. 1996b;35: 15760–15771.

54. Vrkljan M, Foster TM, et al. Thermal stability of low molecular weight urokinase during heat treatment. II. Effect of polymeric additives. Pharm. Res. 1994;11: 1004–1008.

55. Won CM, Molnar TE, et al. Stabilizers against heat Induced aggregation of RPR, 114849, an acidic fibroblast growth factor (AFGF). Int. J. Pharm. 1998;167:25–36.

56. Narhi LO, Philo JS, et al. Reversibility of heat induced denaturation of the recombinant human megakaryocyte growth and development factor. Pharm. Res. 1999;16:799–807.

57. Zhang MZ, Wen J, et al. A new strategy for enhancing the stability of lyophilized protein: The effect of the reconstitution medium on keratinocyte growth factor. Pharm. Res. 1995b;12:1447–1452.

58. Volkin DB, Middaugh CR. The characterization, stabilization, and formulation of acidic fibroblast growth factor. Pharm. Biotechnol. 1996;9:181–217.

59. Kreilgaard L, Frokjaer S, et al. Effects of additives on the stability of recombinant human factor XIII during freeze-drying and storage in the dried solid. Arch. Biochem. Biophys. 1998a;360:121–134.

60. Kreilgaard L, Jones LS, et al. Effect of Tween 20 on freeze-thawing-and agitation-induced aggregation of recombinant human factor XIII. J. Pharm. Sci. 1998b;87:1597–1603.

61. Katakam M, Banga AK. Use of poloxamer polymers to stabilize recombinant human growth hormone against various processing stresses. Pharm. Dev. Technol. 1997;2:143–149.

62. Brange J, Andersen L, et al. Toward understanding insulin fibrillation. J. Pharm. Sci. 1997;86:517–525.

63. Bam NB, Randolph TW, et al. Stability of protein formulations: Investigation of surfactant effects by a novel EPR spectroscopic technique. Pharm. Res. 1995;12:2–11.

64. Wang YJ, Shahrokh Z, et al. Characterization, stability and formulations of basic fibroblast growth factor. Pharm. Biotech-nol. 1996;9:141–180.

65. Chang BS, Beauvais RM, et al. Physical factors affecting the storage stability of freeze dried interleukin-1 receptor antagonist: Glass transition and protein conformation. Arch. Biochem. Biophys. 1996a;331:249–258.

66. Chang BS, Kendrick BS, et al. Surface induced denaturation of proteins during freezing and its inhibition by surfactants. J. Pharm. Sci. 1996b;85:1325–1330.

67. Chang BS, Reeder G, et al. Development of a stable freeze dried formulation of recombinant human interleukin-1 receptor antagonist. Pharm. Res. 1996b;13:243–249.

68. Eckhardt BM, Oeswein JQ, et al. Effect of freezing on aggregation of human growth hormone. Pharm. Res. 1991;8:1360–1364.

69. Anchordoquy TJ, Carpenter JF. Polymers protect lactate dehydrogenase during freeze drying by inhibiting dissociation in the frozen state. Arch. Biochem. Biophys. 1996;332:231–238.

70. Rupley JA, Careri G. Protein hydration and function. Adv. Protein Chem. 1991;41:37–172.

71. Crowe JH, Crowe LM, et al. Preserving dry biomaterials: The water replacement hypothesis, part 2. Biopharm. 1993;6:40–43.

72. Franks F. Long-term stabilization of biologicals. Biotechnology (NY). 1994;12:253–256.

73. Sarciaux JM, Mansour S, et al. Effect of species, processing conditions and phosphate buffer composition on IgG aggregation during lyophilization. Pharm. Sci. (Suppl.). 1998;1:545.

74. Lueckel B, Helk B, et al. Effects of formulation and process variables on the aggregation of freeze-dried interleukin-6 (IL-6) after lyophilization and on storage. Pharm. Dev. Technol. 1998;3:337–346.