Engineering Aspects of Protein Interactions and Self-assembly

Miriam Linsenmeier and Paolo Arosio*

Abstract: In the new Laboratory for Biochemical Engineering (LBCE) at ETH Zurich researchers combine principles of chemical engineering with microfluidic technology and biophysical methods to investigate the physical determinants of biomolecular self-assembly in living organisms. In this account, we show the impact of this activity on concrete applications in biomedical sciences and biotechnology. We focus in particular on the field of protein aggregation and phase separation, and we highlight examples in the context of diagnosis and treatment of Alzheimer's disease and neurodegenerative disorders, cell compartmentalization as well as manufacturing and delivery of therapeutic proteins.

Keywords: Alzheimer's disease · Amyloids · Microfluidics · Protein aggregation · Protein interactions · Protein phase separation

1. Protein Structure and Interactions

Self-organization is one of the fundamental processes underlying the origin of life. The series of cellular functions on which life depends originate at the molecular level from interactions between biomolecules, namely proteins and nucleic acids. Proteins represent a unique class of lyophilic colloids with heterogeneous and complex structures, which exhibit amphoteric polyelectrolyte nature and the simultaneous presence of hydrophobic and hydrophilic patches. Proteins are synthesized as linear chains of amino acids which can fold into suitable secondary and tertiary 3-dimensional structures. In some cases, the stability and the activity of proteins require the formation of quaternary structures, such as complexation into oligomers or assembly into large filaments.

Structure and intermolecular forces are two strongly interconnected aspects in proteins: evolution has optimized the formation of secondary, tertiary and quaternary structure by controlling a delicate balance of intramolecular and intermolecular forces, which include electrostatic, van der Waals, solvent and hydrophobic interactions. In particular, hydrophobic forces play a major role in protein folding, since globular proteins show stable native structures in which hydrophilic groups are solvent-exposed, while hydrophobic patches are buried intramolecularly. At the same time, the complex surface chemistry and structure of proteins allow these molecules to perform specific tasks by generating highly specific intermolecular interactions with other proteins and biomolecules.

This subtle balance of intermolecular forces is highly sensitive to environmental changes and even to single point mutations of the protein sequence, and is not free of mistakes. In some cases, aberrant interactions leading to misfolding and undesired aggregation can occur.[1] This is the case for instance of amyloids, which represent a particularly important class of aberrant aggregates of peptides and proteins which are involved in several human disorders including Alzheimer’s, Parkinson’s and systemic amyloidoses.[2] Amyloids can originate from a broad range of very different peptides and proteins and exhibit a remarkably consistent supramolecular structure, which is mostly independent of the precursor protein and consists of a diameter of few nanometers, a length of several microns, and a predominant cross β-sheet secondary structure.[2c] This common structure is primarily due to the high number of hydrogen bonds that proteins can form along the β-sheet structure. In addition to these specific interactions, the thermodynamic stability of the fibrils is further promoted by the exclusion of water molecules and specific side-chain interactions within the inner core of the fibrils.[2c]

Although amyloid structures were originally discovered in association with a variety of pathological conditions,[2a,c] increasing evidence demonstrates that amyloids underlie a series of physiological functions,[3] including the natural storage of peptide hormones,[4] the formation of adhesive microbe biofilms[5] and cell compartmentalization.[6]

As biochemical engineers we are interested in understanding the physical determinants of biomolecular assembly in living organisms and their connection with functional and aberrant behaviors. This activity requires addressing a series of theoretical and analytical challenges, since biological
systems consist of heterogeneous mixtures that often exhibit transient interactions, low concentration of metastable species and complex aggregation networks. In the following paragraphs we show how concepts of chemical engineering lead to theoretical and experimental advances which allow to achieve quantitative information about protein self-assembly processes at the molecular level. We highlight the direct implications of these findings in different areas of biological and biomed-ical sciences, including the diagnosis and therapy of Alzheimer’s disease and cell compartmentalization. We finally discuss how the developed approaches and the lessons learned from nature allow to design functional materials as well as to monitor and engineer the stability of proteins in biotechnology.

2. Reaction Engineering in Amyloids and Alzheimer’s Disease

The formation of amyloid plaques of the peptide Abeta42 (Aβ42) or analogous peptides, which originate from the fragment-ation of the amyloid precursor protein (APP), is a hallmark of Alzheimer’s disease (AD), a devastating form of dementia that affects more than one thousand people in Switzerland and millions of people worldwide. An increasing amount of evidence indicates that the formation of amyloid fibrils is a critical upstream process in the series of events leading to the loss of neuronal function. Today no effec-tive pharmaceutical treatment exists for this disorder, a fact that reflects our current lack of understanding of the molecular mechanisms underpinning the disease. Despite the direct proof of the causality between the formation of aggregates and the disease is still lacking, the aggregation process represents an attractive target for therapeutic intervention.

In particular, it is emerging that soluble oligomers generated during the aggregation process could represent a source of toxicity via several possible mechanisms. The experimental characterization of these species is primarily impaired by their metastability and low concentrations. Moreover, the identification of the time evolution of their concentration is complicated by the variety of microscopic steps of nucleation and growth which compose the aggregation process. Indeed, in addition to primary nucleation events, amyloid fibrils grow via elongation and several secondary nucleation processes, including fibril fragment-ation and monomer-dependent secondary nucleation reactions (Fig. 1A). It is therefore crucial to identify the contribution of these individual reactions to the overall aggregation rate and in particular to the generation of the soluble intermediates. In analogy to traditional fields of chemical engineering, such as polymer reaction en-gineering, colloidal dispersions and combustion processes, also in this context concepts of reaction engineering and chemical kinetics allow to unravel complex kinetic schemes from a limited number of available macroscopic data (Fig. 1). In addition to the identification of the critical microscopic steps that are most responsible for the generation of toxic species, this information is fundamental to explain several aspects of the aggregation process. For instance, the application of chemical kinetic analysis has elucidated the nature and the molecular origin of the macroscopic lag-phase that is typically observed in the formation of amyloids from solutions of soluble peptides and proteins. This lag-phase could be erroneously interpreted as a waiting time required to reach a critical concentration of nuclei before fibril growth can occur. Chemical kinetic analysis has revealed that each individual microscopic reaction of growth and secondary nucleation is actually present from the very beginning of the aggregation process, and that the lag-phase represents the time required for fibrils to amplify and reach a critical concentration which is detectable by experimental analytical methods. This concept has been exploited to develop a highly sensitive assay for the quantification of low concentra-tions of amyloid fibrils. Measurements of aggregation rates can also be used to indirectly quantify specific interactions between molecules by analyzing changes in the aggregation profiles in the absence and presence of the binding partner. This approach is attractive in applications characterized by the presence of complex mixtures as well as transient and metastable interactions. One example is the analysis of the mechanisms of inhibi-

![Image](https://example.com/image.png)

Fig. 1. Reaction engineering in protein aggregation. (A) Chemical kinetics allows to extract information about the microscopic mechanisms of nucleation and growth underlying the formation of amyloids. (B) This analysis can lead to the quantification of specific interactions in complex mixtures as well as to the identification of perturbations of individual reactions in complex kinetic schemes. This strategy has allowed for instance to identify the variety of microscopic mechanisms of inhibition of molecular chaperones against the formation of amyloid fibrils. (C) The same approach shed light on the combined effect of interfaces and mechanical agitation on the formation of amyloid fibrils from soluble human insulin. Figures reproduced with permission of Springer Nature (ref. [18a]) and ACS Publications (ref. [41]).
bition of amyloid fibrils. Inhibitors can interact with several soluble and aggregated protein species in the reaction mixture, and therefore they can potentially perturb many different microscopic reactions. Depending on the specific reactions that are affected in the aggregation network, different modalities of inhibition can have dramatically different consequences on the generation of the toxic intermediates. In particular, some modalities could increase the toxicity even if the overall aggregation rate is retarded. It is clear that an effective intervention should aim at inhibiting the specific steps responsible for the generation of the toxic oligomers rather than at generically arresting the aggregation rate. In this context, chemical kinetics is a powerful tool to identify specific interactions between the inhibitor and the several protein species present in the reaction mixture, as well as to discriminate which specific microscopic reactions are affected by the presence of the inhibitor. For instance, the aggregation scheme of the Aβ42 peptide is characterized by the presence of a secondary nucleation reaction catalyzed by the surfaces of existing fibrils. Molecules capable of inhibiting this molecular process are highly efficient in removing the toxicity associated with the formation of amyloids without modifying the final fibril load. The application of this kinetic approach is leading to the identification of several small molecules and antibodies capable of targeting the nucleation processes in AD.

Moreover, by applying the same platform we have shown that molecular chaperones, which are key components of the protein homeostasis network, can inhibit the formation of amyloids via a variety of molecular mechanisms, thereby proving that nature has evolved multiple strategies to avoid aberrant protein aggregation (Fig. 1B).

3. Protein Liquid–Liquid Phase Separation

Attractive protein–protein interactions do not result only in the formation of aggregates. Protein solutions exhibit a rich phase diagram that includes crystals, nematic phases and liquid–liquid phase separation (LLPS). The latter process is characterized by the formation of a highly concentrated disperse phase within a dilute continuous phase (Fig. 2). Recent findings revealed that LLPS plays an important role in cell compartmentalization. In addition to vesicle-like organelles, cells can form dynamic membraneless organelles with defined composition and widespread functions, including ribosome biogenesis, RNA processing and storage, as well as stress-induced adjustment of the cellular metabolism. These membraneless bodies have been observed in the nucleus, in the cytoplasm and in membrane-associated parts of the cell, and are rich in nucleic acids and RNA-binding proteins. Despite the widespread appearance in many different organisms, the mechanisms that control their reversible formation and disassembly as well as their biophysical properties remain largely elusive.

Most of these proteins possess N- and/or C-terminal low-complexity domains (LCDs), which are intrinsically disordered and dynamic structures that appear to play a key role in the liquid–liquid phase transition underlying the formation of the bodies. The sequences of the LCDs contain a limited set of amino acids which are frequently arranged in repetitive sequence motifs and encode specific intermolecular interactions: for instance, glutamine and asparagine amino acids induce polar forces, interspersed aromatic amino acids enable π-π stacking, and repetitive sequence motifs, such as polyQ and RG/RGG repeats, promote multivalent interactions. The specificity of these interactions is important for the cell to carefully control the balance of forces required for the assembly and disassembly of the bodies. The presence of attractive interactions is required to induce LLPS but at the same time this attractive energy should not lead to irreversible structures that would prevent the recovery of the soluble components. A picture is emerging in which cells have developed a series of control mechanisms to tune the biophysical properties and the dynamics of the protein-rich droplets, including for instance interactions between proteins and RNA molecules, post-translational modifications and binding to ATP.

The high specificity of the biomolecular interactions underpinning liquid phase transition is very different from the situation observed with amyloids, in which the dominant energetic driving force for aggregation is represented by aspecific hydrogen bonds along the cross-β-sheet structures. This aspecificity largely explains the reason why many different soluble peptides and proteins can convert into a similar, generic amyloid structure. In contrast, different membraneless compartments require specific compositions and tailored biophysical properties to perform specific tasks, which are encoded by specific interactions.

In analogy to protein folding, misregulation of the delicate balance of interactions in LLPS can promote the transition from a liquid-like state to solid, irreversible aggregates. This behavior has been observed for instance with the RNA-binding protein Fused in Sarcoma (FUS), which is associated with Amyotrophic Lateral Sclerosis (ALS), a neurodegenerative disease which is characterized by the degeneration of motor neurons in brain and spinal cord. Wild-type FUS possesses an N-terminal low-complexity domain and several RGG-rich regions which promote the formation of stable liquid–liquid phase separation. The substitution of a single amino acid (G156E) associated with ALS pathology results in the aging of the protein-rich droplets into fibrillar, solid amyloids.

These findings indicate that attractive intermolecular interactions in protein solutions can promote the formation of a continuum of protein-rich phases from liquid to solid, and that this process is highly sensitive to small perturbations, including for instance minor changes in the amino acid sequence induced by mutations or post-translational modifications. Given the fatal consequences that aberrant protein aggregation can have in vivo, it is of fundamental importance to identify the molecular determinants of control functional LLPS to identify targets for intervention when misregulated behavior is observed.

![Fig. 2. Thermodynamics of liquid mixtures. Protein solutions exhibit complex phase diagrams, including liquid-liquid phase separation (LLPS) induced by attractive intermolecular forces (θ). LLPS generates a dispersed phase at high protein concentration (θD) into a continuous phase at diluted protein concentration (θc). Recent findings indicate that this process underlies the formation of membraneless bodies in cells. Several mechanisms can trigger the assembly and disassembly of these bodies, including interactions between proteins and RNA molecules. The images show the phase separation of a solution of the model DEAD-box protein Dhh1 associated with the formation of processing bodies in the absence (i) and presence (ii) of RNA. Scale bar is 25 μm.](image-url)
In analogy with the field of amyloids, the application of concepts of polymer physics,[21] chemical kinetics and thermodynamics of complex mixtures[29] can have a significant impact on activity.

4. From Nature to Biotechnology

In the paragraphs above we have discussed some examples of the rich panorama of protein interactions and self-assembly in living organisms. The lessons learned from biology open the possibility to develop advanced protein materials that exhibit attractive features in terms of function, specificity and biocompatibility.[30] Protein cages,[31] functionalized amyloids[32] and self-assembling peptides[33] are only a few examples of the variety of materials that have been recently developed in laboratories in Switzerland and in the world. These materials find applications in a broad range of fields including water treatment,[33] food science,[34] vaccine delivery,[35] drug encapsulation[36] and enzyme technology.[37]

In other applications of material sciences protein aggregation represents an undesired process that must be carefully avoided. This is for instance the case of therapeutic proteins, a class of drugs that have been recently developed against diseases such as multiple sclerosis, rheumatoid arthritis, tumors and hemophilia.[38] The industrial production of these drugs involves a series of steps (expression, purification, shipping and storage) which can expose the proteins to different types of stresses, such as acidic, thermal and mechanical treatment.[39] Such stresses can compromise the chemical and physical stability of proteins and, in some cases, induce aggregation. Aggregation must be strictly avoided, not only because it leads to loss of valuable product, but, most importantly, because aggregates are potentially immunogenic.[40] The stability during manufacturing and administration is therefore crucial for the successful application of these drugs.

Also in this context, the application of concepts of reaction engineering and chemical kinetics can reveal the complex molecular mechanisms underlying aggregation.[40] This analysis has implications in the identification of lead candidates during the early-stages of the process as well as in the optimization of processing and formulation conditions. A recent example is represented by our study on the analysis of the synergistic effect of interfaces and mechanical forces on protein aggregation[41] (Fig. 1C). We have shown that under physiological conditions hydrophobic/hydrophilic interfaces can dramatically accelerate the formation of fibrils of human insulin, one of the most common therapeutic peptides. The application of chemical kinetic analysis enabled us to go beyond the empirical observation of this effect and demonstrate that interfaces specifically promote heterogeneous nucleation events, while mechanical agitation amplifies the aggregation rate by promoting the turnover of species on the surface and the fragmentation of fibrils in bulk.[42] These findings can explain some contradictory results reported in the literature about the effect of shear on protein aggregation and have practical implications in the bioprocessing of proteins.

The analytical characterization of protein aggregates and, more generally, of biocolloids is another area where biochemical engineers can largely contribute. In this context novel approaches based on microfluidic technology[42] are emerging as powerful tools to improve the experimental characterization of proteins and biocolloids in both fundamental studies and practical applications,[43] including for instance the industrial bioprocessing of therapeutic proteins.[44]

The manipulation of matter at the micron scale allows to accurately control heat and mass transport phenomena, which have different features with respect to the bulk and open the possibility to perform operations which are not achievable with traditional methods.[43] Moreover, the reduced amount of material and time required by the analysis, inherent to miniaturization, largely increase the throughput of the measurements. A recent example is represented by a microfluidic diffusion platform in which the dominance of Brownian motion under laminar flow regime is exploited to accurately measure the time and spatial evolution of diffusion profiles[45] (Fig. 3A). This technique allows to evaluate the size distribution of complex mixtures in the submicron range, as well as to quantify specific interactions between biomolecules in solution.[45] The same approach can also be applied to measure the viscosity of complex solutions at zero shear by probing the Brownian motion of standard tracers[46] (Fig. 3B).

Another strength of microfluidics in protein science is the possibility to manipulate multiphase flows and compartmentalize solutions in small reactors of picoliter and nano-liter by generating water-in-oil emulsions[42,47] (Fig. 3C). Droplet microfluidic technology is particularly suitable to create physiological-like environments and investigate phase transition, nucleation events, reaction-diffusion phenomena and confinement effects in the absence of air-water interfaces.[48] Moreover the possibility to generate thousands of identical droplets in short time dramatically improves the throughput of the measurements.

5. Conclusions and Outlook

In summary, we have shown that key concepts of chemical engineering, such as reaction engineering, thermodynamics of complex mixtures and mass transport phenomena at the micron scale, can contribute to theoretical and practical problems of biomolecular self-assembly. Implications of these studies range from biomedical engineering to biotechnology. Concrete examples include the diagnosis and treatment of Alzheimer’s disease, the issue of cell compartmentalization and the stability of therapeutic proteins during industrial manufacturing.

We finally note that the applications of these approaches in biological sciences are not limited to the field of protein aggregation. Theoretical and experimental advancements based on microfluidic technology have the power to address a variety of other problems including for instance interactions between proteins and nucleic acids, cellular chemotaxis, separation and characterization of exosomes and vesicles. We have started to apply our methods to these systems, which have practical implications in other areas of biomedical sciences and biotechnology such as immunology, drug delivery, and cosmetics.

Acknowledgements

Work in our laboratory is funded by a start-up grant from ETH Zurich, the Swiss National Science Foundation (205321_179055), the ETH Zurich Research Grant ETH-33 17-2, the Synapsis Foundation and the Novartis Foundation for medical-biological Research. We thank Marie Kopp (ETH Zurich) for acquiring the images shown in Figure 3C.

Received: February 28, 2018

Fig. 3. Separation processes and mass transport at the micron scale. A) Microfluidic diffusion platform for the analysis of interactions and size distributions in the submicron range. The same device can be applied to measure the viscosity of protein solutions and complex mixtures. B) Images of a co-flow and a T-junction 2-D device for the generation of water-in-oil emulsions. These images have been acquired by brightfield and fluorescence microscopy, respectively. Scale bar is 50 µm. These new methods find applications in many fields of protein science, including the analysis of protein stability during bioprocessing and formulation of therapeutic proteins.\(^\text{[46]}\) Figures reproduced with permission of ACS Publications, refs \([45, 46]\).

---

\[ \frac{\partial r(t)}{\partial t} = [\nabla \cdot v + D \nabla^2 \theta] r(t) \]

\[ \text{Deconvolution} \]

\[ \text{Size distribution} \]

\[ \text{Free Ligand} \]

\[ \text{Complex Ligand - Target} \]

---

\[ s(t) = \text{BSA concentration (mg/mL)} \]

\[ \text{BSAc concentration (mg/mL)} \]

---

\[ \text{Size distribution} \]

\[ \text{Free Ligand} \]

\[ \text{Complex Ligand - Target} \]

---

\[ \text{Microfluidics} \]

\[ \text{DLS} \]

---

\[ \text{Microfluidics} \]

\[ \text{DLS} \]

---

\[ \text{Microfluidics} \]

\[ \text{DLS} \]

---

\[ \text{Microfluidics} \]

\[ \text{DLS} \]

---

\[ \text{Microfluidics} \]

\[ \text{DLS} \]