

PACS: 87.14.Cc, 87.16.Dg

**MODELIZATION OF AMYLOID FIBRIL SELF-ASSEMBLY****V. Trusova, G. Gorbenko***Department of Nuclear and Medical Physics, V.N. Karazin Kharkiv National University**4 Svobody Sq., Kharkiv, 61022, Ukraine**E-mail: valerija.trusova@karazin.ua**Received February 22, 2018*

Intermolecular noncovalent interactions between protein molecules result in the formation of a wide spectrum of supramolecular assemblies the structure of which varies from disordered amorphous aggregates to the crystals with strictly defined translational symmetry in three directions. One-dimensional protein aggregates (amyloid fibrils) represent highly ordered semiflexible polymers with unique mesoscopic properties which can be tuned by both intrinsic physicochemical characteristics of polypeptide chain and milieu conditions. In the present work the molecular mechanisms of amyloid formation are discussed and mathematical description of the existing models of protein fibrillization are given. For disease-related amyloids, deeper understanding of fibril growth process may shed light on the pathogenesis and molecular mechanisms of the disorders, as well as on the strategies of amyloidosis prevention at atomistic level. In the context of nanotechnology and functional material science, knowing the details of amyloid formation is crucially required for the design of novel nanomaterials with unprecedented qualities.

**KEYWORDS:** amyloid fibrils, misfolding, mathematical models, amyloid diseases, nanomaterials**МОДЕЛЮВАННЯ ПРОЦЕСУ САМОЗБИРАННЯ АМІЛОЇДНИХ ФІБРИЛ****В. Трусова, Г. Горбенко***Кафедра ядерної та медичної фізики, Харківський національний університет імені В.Н. Каразіна**пл. Свободи 4, Харків, 61022, Україна*

Міжмолекулярні нековалентні взаємодії між білковими молекулами призводять до формування широкого спектру надмолекулярних ансамблів, структура яких коливається від неупорядкованих аморфних агрегатів до кристалів із чітко визначеною трансляційною симетрією у трьох напрямках. Одномірні білкові агрегати (амілоїдні фібрили) представляють собою високопорядковані напівгнучкі полімери з унікальними мезоскопічними властивостями, що визначаються як внутрішніми фізико-хімічними характеристиками поліпептидного ланцюга, так і зовнішніми умовами. У даній роботі обговорюються молекулярні механізми утворення амілоїдів, а також наведено математичний опис існуючих на сьогодні моделей фібрилізації білків. Для амілоїдів, які пов'язані із різноманітними патологіями, розуміння процесу росту фібрил може пролити світло на патогенез та молекулярні механізми амілоїдних захворювань, а також на стратегії попередження амілоїдозу на атомістичному рівні. У контексті нанотехнологій та функціональних матеріалів, розуміння деталей утворення амілоїдів необхідне для дизайну нових наноматеріалів із перспективними властивостями.

**КЛЮЧОВІ СЛОВА:** амілоїдні фібрили, місфолдінг, математичні моделі, амілоїдні захворювання, наноматеріали**МОДЕЛИРОВАНИЕ ПРОЦЕССА САМОСБОРКИ АМИЛОИДНЫХ ФИБРИЛЛ****В. Трусова, Г. Горбенко***Кафедра ядерной и медицинской физики, Харьковский национальный университет имени В.Н. Каразина**пл. Свободы 4, Харьков, 61022, Украина*

Межмолекулярные нековалентные взаимодействия между белковыми молекулами приводят к формированию широкого спектра надмолекулярных ансамблей, структура которых колеблется от неупорядоченных аморфных агрегатов до кристаллов с четко определенной трансляционной симметрией в трех направлениях. Одномерные белковые агрегаты (амилоидные фибриллы) представляют собой высокоупорядоченные полугибкие полимеры, с уникальными мезоскопическими свойствами, которые определяются как внутренними физико-химическими характеристиками полипептидной цепи, так и внешними условиями. В данной работе обсуждаются молекулярные механизмы образования амилоидов, а также приведено математическое описание существующих на сегодняшний день моделей фибриллизации белков. Для амилоидов, связанных с различными патологиями, понимание процесса роста фибрилл может пролить свет на патогенез и молекулярные механизмы амилоидных заболеваний, а также на стратегии предупреждения амилоидоза на атомистическом уровне. В контексте нанотехнологий и функциональных материалов, понимание деталей образования амилоидов необходимо для дизайна новых наноматериалов с перспективными свойствами.

**КЛЮЧЕВЫЕ СЛОВА:** амилоидные фибриллы, мисфолдинг, математические модели, амилоидные заболевания, наноматериалы

The aggregation of specific peptides and proteins into highly ordered  $\beta$ -sheet fibrillar structures, termed amyloid fibrils, plays a crucial role in the onset and progression of a number of amyloid-related pathologies including Alzheimer's, Parkinson's and Huntington's diseases, type II diabetes, rheumatoid arthritis, spongiform encephalopathies, etc. [1,2]. The impairment of cell functions by fibrillized proteins and their intermediates is causatively linked to the complex conformational behavior of these aggregates, which are distinguished by the presence of a core cross- $\beta$ -sheet structure with  $\beta$ -strands orienting perpendicularly to the fibril long axis and  $\beta$ -sheets propagating in its direction. Such molecular architecture is stabilized by the main-chain hydrogen bonding, ionic pairing, van der

Waals, aromatic  $\pi$ - $\pi$  interactions and hydrogen bonds between amino-acid side chains, with a central role in self-assembly propensity being played by protein hydrophobicity. Along with pathogenic activities, amyloid fibrils display superior physical and structural properties that render them the ideal candidates for the design of novel functional materials for various applications. Specifically, protein fibers were successfully used for constructing the nanowires, as scaffolds for tissue engineering, depot for drug transport, just to name a few [3-5].

In the light of versatility of amyloid phenomena, the understanding of the pathways of fibril formation is of utmost importance both from the viewpoint of combating the amyloid disorders and smart design of novel functional materials. The aim of the present contribution is to analyze the molecular basis for protein fibrillization, and overview some mathematical models proposed to describe the process of amyloid formation.

### MOLECULAR BASIS FOR FIRILLIZATION

For a long time the ability to form amyloid fibrils was attributed only to several proteins and peptides involved in the development of neurodegenerative diseases. However, continuously growing evidence supports the idea that a propensity to convert into fibrillar aggregates represents a generic property of polypeptide chain. The molecular basis for this process lies in the incorrect protein folding or misfolding.

In general, under physiological conditions, polypeptide chain reaches its native functional state through the folding pathways. The process of protein folding is commonly described by the principle of minimal frustration [6]. According to this principle, interdependent interactions between the side chains of amino acid residues cooperatively lead to the most stable low-energy structure which thus “minimally frustrated” [7]. This results in the funnel-shaped energy landscape of a protein molecule, described by an energy function  $F(x) = F(x_1, x_2, \dots, x_n)$ , where  $n$  denotes protein degrees of freedom,  $x_1, x_2, \dots, x_n$  specify microscopic state of the molecule (for instance, dihedral angles of polypeptide chain),  $F(x)$  is defined as protein free energy in a given conformation. The most thermodynamically stable protein conformation corresponds to the set of  $x_1, x_2, \dots, x_n$  values that yields the minimum value of  $F(x)$ . The funnel shape of the energy landscape predicts that the free energy decreases while approaching the native state (Figure). The wide top of the funnel represents the denatured, or unfolded states of a protein, while narrow bottom represents the global minimum of  $F(x)$ .

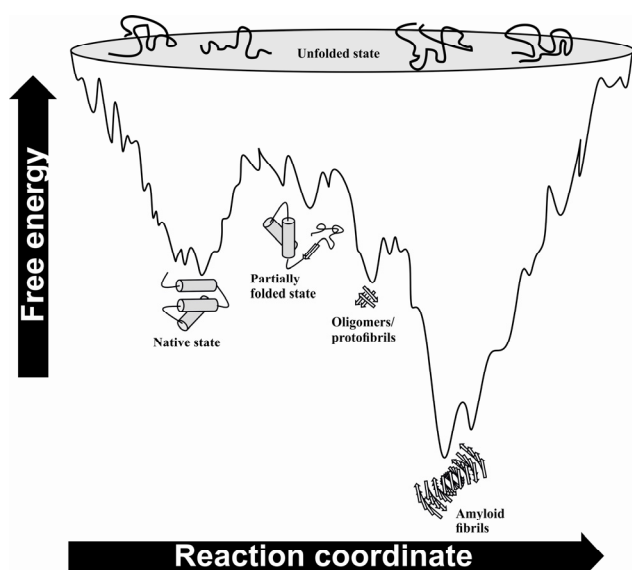


Figure. Energy funnel landscape of protein molecules.

However, due to high conformational freedom and fluctuations, the funnel energy landscape of every protein molecule is rugged, displaying several additional local minima on the energy surface. This implies that at each step of the folding process, protein molecule is in an equilibrium with partially folded intermediates, reducing thereby the probability of a polypeptide chain to reach its native state. The number and activity of such intermediates are controlled by polypeptide sequence and environmental conditions. These intermediate states generally have the exposed patches of hydrophobic amino acid residues which are prone to self-association. Hydrophobic interactions between partially folded states force the polypeptide chain to leave the native folding pathway and enter the off-folding route. Contrary to spontaneous and rapid folding process, the off-folding route is characterized by relatively slow kinetics. One of the possible off-folding pathways involves the transition of partially unfolded intermediates into the highly

ordered fibrillar structures, amyloid fibers [8]. These misfolded protein states are characterized predominantly by  $\beta$ -sheet secondary structure. The factors, which shift the balance between the intramolecular forces accounting for correct protein packing, and trigger the off-folding transformation into the ordered amyloid self-assemblies, involve genetic mutations, oxidative stress, physicochemical properties of the solvent (pH, temperature, ionic strength, pressure), etc.

### INTERMOLECULAR FORCES STABILIZING THE AMYLOID FIBRIL STRUCTURE

Modern models of amyloid fibers strongly highlight the fact that it is the hierarchical architecture that governs the physical properties of these aggregates. Extraordinary stability of the amyloids and their resistance to high hydrostatic pressures and temperatures suggest that fibrillar structure represents a free energy state lower than that of a native protein. Amyloid aggregates are stabilized by a set of interactions including hydrogen bonding, hydrophobic,

electrostatic, and aromatic  $\pi$ - $\pi$  interactions. The central role in self-assembly propensity is played by protein hydrophobicity [9,10].

A network of hydrogen-bonds acts like a molecular skeleton which holds the overall fibrillar architecture. Analogously to the folding of globular proteins, self-assembly of protein monomers into amyloid structures is essentially driven by hydrophobic effect. Hydrophobic collapse of polypeptide chain into a molten globule state, provoked the factors such as elevated protein concentration, temperature, pH and ionic strength of the solution, initiates the nucleation process [11,12]. The non-polar residues exposed during the protein unfolding attract to each other in order to minimize the unfavorable contacts with solvent. The mature fibrils, as well as the transient oligomeric intermediates, are stabilized by a network of hydrogen bonds along the amide backbone. A wealth of experimental results supports the idea that increase in sequence hydrophobicity accelerates the rate of fibrillization and enhances the thermodynamic stability of the resulting cluster [9,10]. A specific subclass of hydrophobic interactions, namely aromatic contacts was also proposed to contribute to the thermodynamic stabilization of the fibrils, but to a lesser extent. While Marek et al. concluded that  $\pi$ - $\pi$  interactions are kinetically important only during the lag phase of fibril formation [13], Gazit postulated that the role of aromatic interactions is two-fold: i) they provide positive energetic contribution arising from the  $\pi$ - $\pi$  stacking itself, and ii) their specific directional effect orients properly the  $\beta$ -strands to stack into a highly ordered structure [14].

Due to polyelectrolytic nature of the polypeptide chain, in the following it is tempting to briefly overview the role of electrostatics in the amyloid stability. The overall protein electrostatic potential, leading to the intra- and interchain repulsion, in couple with the unbalanced charges on N- and C-termini of the chain, result in the fact that under physiological conditions proteins are unable to self-assembly. However, once the protein is transferred into the conditions of Debye screening of the unfavorable repulsions between the residues (low pH and high ionic strength), the process of fibrillization is initiated. In addition, using the approach of capped peptide ends, the important role of electrostatic attraction between the charged termini in the lateral end-to-end association of the protofilaments has been revealed [15]. Likewise, several lines of evidence indicate that the role of electrostatics in amyloid formation lies in stabilization of the amyloid charged and polar segments as well as fibril turns by the formation the salt bridges between the side chain amino acid residues. Such salt bridges are typically of low energy and do not completely suppress the negative contribution of the net protein charge in the aggregation process. The formation of salt bridges in a fibrillar state was observed, for example, for apolipoprotein A-I fragment [16], prion protein [17], and A $\beta$  peptide [18].

### MOLECULAR MODELS FOR AMYLOID FORMATION

Accumulating experimental and theoretical evidence indicates that protein fibrillization is an extremely complex process characterized by several more or less clearly distinguished stages. Furthermore, the existence of multiple amyloid polymorphs, along with a variety of structural intermediates, strongly suggests that assembly of highly ordered protein aggregates may occur via different mechanisms. A great deal of models have been proposed to describe the amyloid formation [19-23].

The pioneering work concerning the protein polymerization belongs to Oosawa et al., who in 1959 investigated the aggregation of G-actin to F-actin [20]. In 1967 Griffith proposed the mechanistic template-assembly (TA) model to describe self-replication of the prion disease agent scrapie [24]. TA model predicts the existence of two protein states, a stable soluble conformation (S), and assembly-competent state (A). The elongation of the amyloid fiber occurs via attachment of S monomers to the A-state nucleus. Next significant contribution to the field of protein aggregation was made by Ferrone and co-workers who in 1980 proposed the model for polymerization of sickle hemoglobin considering the so-called "homogeneous" and "heterogeneous" nucleation, where "heterogeneous" nucleation means that new polymers nucleate on the surface of existing ones [25].

Monomer-directed conversion (MDC) model was formulated for prion aggregation by Prusiner et al. in 1982 and further extended in 1991 [26]. This model suggests the pre-existence of S monomers in the A-state conformation, which subsequently associate with the soluble S-state monomers and eventually initiate their rate-determining conversion into reactive A-state form.

Following the work of Prusiner, in 1993 Lansbury et al. developed the nucleation-dependent polymerization (NDP) model [21]. Initially, it was proposed to explain the mechanisms of prion transmission and propagation, but nowadays this model represents one of the basic kinetic models of protein fibrillation. Within the framework of this model protein conversion into the amyloid state is described by biphasic kinetic profile with the lag and growth phases. The lag (or nucleation) phase involves the monomer self-assembly into the critical nucleus having the highest energy and thus being the most thermodynamically unstable state. This process is followed by the elongation of the nucleus via

attachment of the additional monomers eventually resulting in thermodynamically favorable process of exponential fibril growth. Notably, the nucleus formation represents the rate-limiting step for the entire fibrillization process. The kinetics of the lag phase can be controlled by the environmental conditions (pH, ionic strength, temperature), mutations and/or addition of the pre-aggregated species (template) to the solution (known as seeding effect).

Assuming that the nucleus is represented by 10-mer, and aggregates larger than 10-mer are treated as fibrils, the kinetic equations of NDP model may be written in a general form as [27]:

$$\frac{d[M_1]}{dt} = -k_1[M_1] \left( 2[M_1] + \sum_{j=2}^9 [M_j] \right) + k_2 \left( 2[M_2] + \sum_{j=3}^{10} [M_j] \right) - k_p[M_1][M_{10}] - k_p[M_1][P], \quad (1)$$

$$\frac{d[M_j]}{dt} k_1[M_1]([M_{j-1}] - [M_j]) - k_2([M_j] - [M_{j+1}]) \quad (j = 2 - 9), \quad (2)$$

$$\frac{d[M_{10}]}{dt} = k_1[M_1][M_9] - k_2[M_{10}] - k_p[M_1][M_{10}], \quad (3)$$

$$\frac{d[M_j]}{dt} = k_p[M_1]([M_{j-1}] - [M_j]) \quad (j = 11 \text{ to } N - 1), \quad (4)$$

$$\frac{d[M_N]}{dt} k_p[M_1][M_{N-1}], \quad (5)$$

where  $k_1$  and  $k_2$  are the rate constants for monomer attachment (nucleus growth) and dissociation;  $k_p$  is the rate constant for nucleus elongation;  $[M_j]$  is the concentration of  $n$ -mer;  $[P] = \sum_{j=11}^{N-1} [M_j]$  is the concentration of fibrils. The total concentration of fibril-forming monomers is given by:

$$[F] = \sum_{j=11}^N j[M_j]. \quad (6)$$

The rates of increasing the monomer and fibril concentrations over time can be expressed as:

$$\begin{aligned} \frac{d[F]}{dt} &= k_p[M_1][P] + 11k_p[M_1][M_{10}] \\ \frac{d[P]}{dt} &= k_p[M_1][M_{10}] \end{aligned} \quad (7)$$

Numerical solution of the above equations yields the rate constants and the degree of oligomerization ( $N$ ). Analysis of protein fibrillization in terms of NDP model allowed explaining the well-known sigmoidal shape of the time-course of protein aggregation.

Further development of fibrillization theory comes from Flyvbjerg et al. who in 1996 developed a model considering the subsequent monomer addition mechanism [22]. Based on the assumptions that i) there is only one route for protein aggregation, and ii) every particular step is connected with the next one only by addition of monomers, the authors derived the kinetic equations for analysis of protein fibrillization.

At the same time Lomakin et al. developed the ideas of Ferrone's [25] and proposed the nucleated-polymerization model (NP) for A $\beta$  fibrillization [23]. This model stipulates that above a certain critical micelle concentration (CMC)  $c^*$ , there exists a rapid thermodynamic equilibrium between the protein monomers and micelles. Formation of the nucleus occurs within the unordered micellar oligomers and is followed by fibril growth via irreversible binding of the monomers to the fibril ends:

$$\dot{N}_p = ck_e N_{p-1} - ck_e N_p + k_n M \delta_{p n_0}, \quad (8)$$

$$nL_f = \sqrt{\frac{m_0 k_e c^*}{k_n}}, \quad N_f = c_0 \sqrt{\frac{k_n}{m_0 k_e c^*}}, \quad (9)$$

where  $M$ ,  $N_p$  and  $c$  represent, respectively, the number concentration of micelles,  $p$ -mer fibers and free protein monomers;  $\dot{N}_p$  is the time-derivative;  $k_n$  and  $k_e$  are nucleation and elongation rate constants;  $\delta_{pm_0}$  represents the correction factor, or scaling term related to the formation of critical nucleus;  $m_0$  and  $n$  are the structural parameters, which mean the number of monomers in a micelle and in a fibril, respectively;  $L_f$  and  $N_f$  being the length and number of fibrils, respectively;  $c_0$  is the initial protein concentration.

Later on, in 1999, Ferrone derived the kinetic equations for the rates of monomer addition and nucleus elongation using the perturbation theory and the assumptions that a) all species present in the solution are classified either as monomers or polymers, b) the rate of polymer formation is determined by the concentration of the nuclei and the rate of their elongation, c) polymer growth occurs via accretion to their ends with the same size-independent rate constant [28]:

$$\frac{dc_p}{dt} = J^* c^*, \quad \frac{d\Delta}{dt} = J c_p, \quad c^* = K_n^* c^{n^*}, \quad \Delta = \frac{1}{2} J J^* c^* t^2, \quad (10)$$

where  $c_p$ ,  $c^*$  and  $\Delta$  standing for the concentrations of polymers, nuclei and polymerized monomers, respectively;  $J$  and  $J^*$  representing the rate of polymer addition and the rate of nucleus elongation;  $K_n^*$  and  $c^{n^*}$  denoting the equilibrium constant for nucleation and critical nucleus size, respectively.

In 2000 Serio et al. have complemented the basic statements of TA and NP models with some additional features to formulate the nucleated-conformational conversion (NCC) model [29]. This model insinuates the transition from the stable monomeric state S to the assembly-competent conformation A via structurally molten oligomeric intermediates. Once oligomeric nucleus is formed, its elongation rapidly proceeds at the fibril ends through templating or induced-fit mechanism. These unstructured intermediates are supposed to be implicated in the modification of the folding environment of the protein allowing it i) to rearrange slowly to the structured nucleus, ii) to undergo conformational transition upon binding to the preformed nucleus, and iii) to transit to the species with high degree of flexibility, which are prone to a more rapid assembly in the presence of template (nucleus).

Parallel to NCC model, Kamishira et al. proposed a three-step mechanism for calcitonin fibrillization which involves 1) formation of protein micelles, 2) micelle transformation into the nucleus, and 3) elongation of the nucleus to the fibrils [30]. Using the assumptions that i) the monomeric and micellar states are in an equilibrium and produce the same measurable signal, ii) the total concentration of monomers is always higher than critical micelle concentration, and iii) each protofibril may act as a catalytic site for the conversion from monomeric to fibrillar state, the expression for the fraction of fibrillized protein was obtained.

In 2001 Pallito and Murphy described the “off-pathway folding” model (OFF) that postulates multi-step route for fibril formation [31]. Specifically, amyloid growth was assumed to start from the denaturated state and proceeds via 1) refolding of the denaturated monomers into either stable monomer/dimer or less stable intermediate, 2) cooperative assembly of the intermediates into the nucleus, 3) elongation of the nucleus into the filaments via repeated attachment of dimeric intermediate, 4) lateral association of the filaments into fibrils, 5) fibril elongation through the end-to-end association.

Next model worth of mentioning is the “dipole-assembly” model proposed by Xu et al. in 2001. This model provides qualitative description and predicts two-step route of amyloid growth [32]. First step involves the nucleation of protein monomers into the granular or spheroid oligomers controlled presumably by hydrophobic interactions. Next, if the nucleation units possess the intrinsic dipole moment, attractive forces between the dipoles would promote further linear aggregation of the nuclei. Nucleus size is governed by the magnitude of the dipole moment which is necessary to overcome the repelling forces between the nuclei or between the nucleation units and the fibril. The overall dipole moment of the fiber enables the templating mechanism, giving rise to the binding of new nucleation units to the ends of already existing fibrils. In addition, asymmetrical distribution of the charges within the protein structure may promote dipole-dipole interactions, triggering the formation of the amyloid fibers.

In 2003 Chiti et al. developed the phenomenological approach to describe the process of protein aggregation based on the physicochemical properties of polypeptide chain [33]. Initially this concept was applied to predict the difference in aggregation propensity between the wild and mutated amyloid proteins. In 2004 this powerful scientific team extended this approach and derived an absolute protein aggregation rate as a function of both intrinsic and extrinsic (condition-dependent) physicochemical characteristics of polypeptide chain.

In the following we refer ourselves to the theoretical methodology of van Gestel and Leeuw, that describes protein fibrillization in a dilute solution [34]. Specifically, the authors suggested the statistical-mechanical model which predicts the following fibril parameters: 1) mean degree of polymerization, 2) fibril length, 3) weight fractions of the

aggregates in solution. To derive the equations of the model, three variants of boundary conditions were considered: i) *case A* – amyloid fibers have the “loose ends”, i.e. small linear stretches of the proteins in a non- $\beta$  conformation, ii) *case B* – fibrils consist only of the proteins in cross- $\beta$ -sheet conformation, i.e. the last monomer in the assembly is in a  $\beta$ -strand configuration, iii) *case C* – fibril ends may attain both of the above states.

Lattice model for fibril formation was developed in 2008 by Li et al. by combining the main principles of TA, NP and NCC models [35]. Within the concept of this model, the assembly of fibrillar aggregates takes place via three steps. The first one is the “burst phase”, characterized by the formation of mobile oligomers of different size. At the second step these oligomers are assembled to form compact, but still disordered structure with multiple intramolecular contacts. The transition into the aggregation-prone state occurs during this phase. Finally, the last stage involves the fiber formation via incorporation of the small oligomeric clusters into the large ones which act as a template. The key conclusions drawn from the analysis of protein aggregation behavior in terms of the lattice model, may be formulated as: i) the kinetics of amyloid formation is characterized by three distinct regimes, with the “burst phase” being the most rapid and the amyloid growth being the slowest one; ii) the growth kinetics depends on the temperature according to Arrhenius law; iii) the growth of the mature fibers during the last stage proceeds via Lifshitz-Slyuzov mechanism, implying that the larger clusters grow at the expense of smaller ones. This suggests that seeding of the solution with pre-formed fibrils would accelerate the fiber growth.

In 2011 NCC model was revisited by Schmit et al. who put forward the thermodynamic theory yielding the free energy landscape of different protein states involved in fibril formation [36]. This theory considers several equilibrium states of the protein: i) monomer, ii) oligomers consisted of a few polypeptide chains, iii) nucleus with  $\beta$ -sheet structure, iv) post-critical nucleus with high level of  $\beta$ -structures, v) protofilament, described as long assembly of  $\beta$ -structures, consisting of two  $\beta$ -sheets in face-to-face orientation, vi) fibril comprising a bundle of protofilaments.

The role of electrostatics in fibril formation was examined by Foderà et al. in 2013 [37]. Specifically, the authors developed the theoretical model, describing the aggregation behavior of a charged protein. The model predicts the assembly of protein monomers into the multifractal complexes based on the calculation of the free energy of spherical cluster with a certain radius  $R$ . The main conclusion reached is that the geometry of aggregate growth is controlled by the electrostatic interactions between the single proteins. At decreased values of fractal dimensions,  $d_f$ , free energy minimum shifts towards the higher radii of the cluster. The changes in growth morphology (i.e. alterations in  $d_f$ ) results in multifractal free energy profile with several minima and more compact structures being energetically favorable at the early stages of aggregation process.

In their another work this group analyzed the impact of electrostatics on the amyloid fibrils formation by describing the protein fibrillization process as a conformational transition between the soluble ( $\Pi$ ) and amyloidogenic ( $\beta$ ) states [38]. Modelization of the protein aggregation potential within the framework of this approach showed that the temporal course of  $\Pi \rightarrow \beta$  conversion is strongly slowed with increasing charging state of the hydrophilic parts of the protein.

Last but not least model worthy of mention in the present contribution is the extended lattice model of Auer [39]. This model suggests that protein  $\beta$ -sheets have weakly hydrophobic (WH) and strongly hydrophobic (SH) surfaces. Each  $\beta$ -strand is represented by the right rectangular prism. The fibril is modeled as 2D aggregate in  $m, i$  plane, where  $m, i$  denote the number of  $\beta$ -strands and  $\beta$ -sheets, respectively. The building blocks of the fiber ( $\beta$ -strands) are assumed to be arranged in 2D lattice with rectangular symmetry. Based on this approach, the equations for the nucleation work, fibril solubility, equilibrium concentration of the nuclei, and fibril nucleation rate as a function of supersaturation of protein solution, have been derived. Thermodynamic analysis of protein polymerization revealed the existence of two possible fibrillization routes. According to the first one, the process of fiber growth begins with the formation of 1 $\beta$ -sheet, that further transforms its shape into 2 $\beta$ -sheets with two SH surfaces, which then convert into 3 $\beta$ -sheets with one SH and one WH surfaces, and subsequently into 4 $\beta$ -sheets with two SH surfaces, etc. Along the second route, 1 $\beta$ -sheet transits into 2 $\beta$ -sheets with two WH surfaces, then into 3 $\beta$ -sheets with one SH and one WH surfaces, then into 4 $\beta$ -sheets with two WH surfaces and so on. SH bonds provide the arrangement of the  $\beta$ -strands into the  $\beta$ -sheets which stretch along  $m$ -axis (lengthening axis), while WH bonds cause  $\beta$ -strands to assemble along the  $i$ -axis (thickening axis).

## CONCLUSIONS

In the present work we tried to structurize the available information about the molecular routes of protein fibrillization, and mathematical models proposed so far to describe this process. The models point to a scenario in which protein fibrillization proceeds via four main stages: misfolding, nucleation, elongation, amyloid fibril growth. Validated by many experimental results, this mechanism is valid for various types of proteins, and for fibrillation processes occurring under different environmental conditions.

## ACKNOWLEDGEMENTS

This work was supported by the grant № 0116U000937 for Young Scientists from the Ministry of Science and Education of Ukraine.

## REFERENCES

1. Chiti F., Dobson C.M. Protein Misfolding, Amyloid Formation, and Human Disease: A Summary of Progress Over the Last Decade // *Annu. Rev. Biochem.* -2017.-Vol. 86.- P. 27-68.
2. Harrison R.S., Sharpe P.C., Singh Y., Fairlie D.P. Amyloid peptides and protein in review // *Rev. Physiol. Biochem. Pharmacol.* – 2007. – Vol. 159. – P. 1-77.
3. Pham C.L.L., Kwan A.H., Sunde M. Functional amyloid: widespread in nature, diverse in purpose // *Essays Biochem.* – 2014. – Vol. 56. – P. 207-219.
4. Maji S., Schubert D., Rivier C., Lee S., Rivier J., Riek R. Amyloid as a depot for the formulation of long-acting drugs // *PLoS Biol.* – 2008. – Vol. 6. - P. e17.
5. Mankar S., Anoop A., Sen S., Maji S. Nanomaterials: amyloid reflect their brighter side // *Nano Rev.* – 2011. – Vol. 2. – P. 6032-6043.
6. Luthey-Schulten Z., Wolynes P. Theory of protein folding: the energy landscape perspective // *Annu. Rev. Phys. Chem.* – 1997. – Vol. 48. – P. 545-600.
7. Ferreira D., Hegler J., Komives E., Wolynes P. On the role of frustration in the energy landscapes of allosteric proteins // *Proc. Natl. Acad. Sci. USA.* – 2011. – Vol. 108. – P. 3499-3503.
8. Straub J.E., Thirumalai D. Toward a molecular theory of early and late events in monomer to amyloid fibril formation // *Annu. Rev. Phys. Chem.* – 2011. – Vol. 62. – P. 437-463.
9. Bowerman C., Ryan D., Nissan D., Nilsson B. The effect of increasing hydrophobicity on the self-assembly of amphipathic beta-sheet peptides // *Mol. Biosyst.* – 2009. – Vol. 5. – P. 1058-1069.
10. Doran T., Kamens A., Byrnes N., Nilsson B. Role of amino acid hydrophobicity, aromaticity, and molecular volume on IAPP (20-29) amyloid self-assembly // *Proteins.* – 2012. – Vol. 80. – P. 1053-1065.
11. Lim K.H., Naqchowdhuri P., Rathinavelan T., Im W. NMR characterization of hydrophobic collapses in amyloidogenic unfolded states and their implications for amyloid formation // *Biochem. Biophys. Res. Commun.* – 2010. – Vol. 396. – P. 800-805.
12. Ramakrishna D., Prasad M., Bhuyan A. Hydrophobic collapse overrides Coulombic repulsion in ferricytochrome *c* fibrillation under extremely alkaline condition // *Arch. Biochem. Biophys.* – 2012. – Vol. 528. – P. 67-71.
13. Marek P., Abedini A., Song B., Kanungo M., Johnson M., Gupta R., Zaman W., Wong S., Raleigh D. Aromatic interactions are not required for amyloid fibril formation by islet amyloid polypeptide but do influence the rate of fibril formation and fibril morphology // *Biochemistry.* – 2007. – Vol. 46. – P. 3255-3261.
14. Gazit E. A possible role for pi-stacking in the self-assembly of amyloid fibrils // *FASEB J.* – 2002. – Vol. 16. – P. 77-83.
15. Marshall K., Morris K., Charlton D., O'Reilly N., Lewis L., Walden H., Serpell L.C. Hydrophobic, aromatic, and electrostatic interactions play a central role in amyloid fibril formation and stability // *Biochemistry.* – 2011. – Vol. 50. – P. 2061-2071.
16. Girych M., Gorbenko G., Trusova V., Adachi E., Mizuguchi C., Nagao K., Kawashima H., Akaji K., Lund-Katz S., Philips M., Saito H. Interaction of thioflavin T with amyloid fibrils of apolipoprotein A-I N-terminal fragment: resonance energy transfer study // *J. Struct. Biol.* – 2014. – Vol. 185. – P. 16-124.
17. Guest W., Cashman N., Plotkin S. Biochem. Electrostatics in the stability and misfolding of the prion protein: salt bridges, self-energy, and solvation // *Cell Biol.* – 2010. – Vol. 88. – P. 371-381.
18. Yun S., Urbanc B., Cruz L., Bitan G., Teplow D., Stanley H. Role of electrostatic interactions in amyloid  $\beta$ -protein (A $\beta$ ) oligomer formation: a discrete molecular dynamics study // *Biophys. J.* – 2007. – Vol. 92. – P. 4064-4077.
19. Gilliam J., MacPhee C. Modelling amyloid fibril formation kinetics: mechanisms of nucleation and growth // *J. Phys. Condens. Matter.* – 2013. – Vol. 25. – P. 373101-373120.
20. Oosawa F., Asakura S., Hotta K., Nobuhisa I., Ooi T. G-F transformation of actin as a fibrous condensation // *J. Polymer Sci.* – 1959. – Vol. 37. – P. 323-336.
21. Jarrett J., Lansbury P. Seeding “one-dimensional crystallization” of amyloid: a pathogenic mechanism in Alzheimer’s disease and scrapie? // *Cell.* – 1993. – Vol. 73. – P. 1055-1058.
22. Flyvbjerg H., Jobs E., Leibler S. Kinetics of self-assembling microtubules: an “inverse problem” in biochemistry // *Proc. Natl. Acad. Sci. USA.* – 1996. – Vol. 93. – P. 5975-5979.
23. Lomakin A., Chung D., Benedek G., Kirschner D., Teplow D. On the nucleation and growth of amyloid  $\beta$ -protein fibrils: detection of nuclei and quantification of rate constants // *Proc. Natl. Acad. Sci. USA.* – 1996. – Vol. 93. – P. 1125-1129.
24. Griffith J. Self-replication and scrapie // *Nature.* – 1967. – Vol. 215. – P. 1043-1044.
25. Ferrone F., Hofrichter J., Ferrone F.A., Hofrichter J., Sunshine H.R., Eaton W.A. Kinetic studies on photolysis-induced gelation of sickle cell hemoglobin suggest a new mechanism // *Biophys. J.* – 1980. – Vol. 32. – P. 361-377.
26. Prusiner S. Molecular biology of prion diseases // *Science.* – 1991. – Vol. 252. – P. 1515-1522.
27. Kodaka M. Requirements for generating sigmoidal time-course aggregation in nucleation-dependent polymerization model // *Biophys. Chem.* – 2004. – Vol. 107. – P. 243-253.
28. Ferrone F. Analysis of protein aggregation kinetics // *Methods Enzymol.* – 1999. – Vol. 309. – P. 256-274.
29. Serio T., Cashikar A., Kowal A., Sawicki G., Moslehi J., Serpell L., Arnsdorf M., Lindquist S. Nucleated conformational conversion and the replication of conformational information by a prion determinant // *Science.* – 2000. – Vol. 289. – P. 1317-1321.
30. Kamishira M., Naito A., Tuzi S., Nosaka A., Saito H. Conformational transitions and fibrillation mechanism of human calcitonin as studied by high-resolution solid-state  $^{13}\text{C}$  NMR // *Protein Sci.* – 2000. – Vol. 9. – P. 867-877.
31. Pallito M., Murphy R. A mathematical model of the kinetics of beta-amyloid fibril growth from the denatured state // *Biophys. J.* – 2001. – Vol. 81. – P. 1805-1822.

32. Xu S., Bevis B., Arnsdorf M. The assembly of amyloidogenic yeast sup35 as assessed by scanning (atomic) force microscopy: an analogy of linear colloidal aggregation? // *Biophys. J.* – 2001. – Vol. 81. – P. 446-454.
33. Chiti F., Stefani M., Taddei N., Ramponi G., Dobson C. Rationalization of the effects of mutations on peptide and protein aggregation rates // *Nature.* – 2003. – Vol. 424. – P. 805-808.
34. van Gestel J., de Leeuw S. A statistical-mechanical theory of fibril formation in dilute protein solutions // *Biophys. J.* – 2006. – Vol. 90. – P. 3134-3145.
35. Li M., Klimov D., Straub J., Thirumalai D. Probing the mechanisms of fibril formation using lattice models // *J. Chem. Phys.* – 2008. – Vol. 129. – P. 175101-175110.
36. Schmit J., Ghosh K., Dill K. What drives amyloid molecules to assemble into oligomers and fibrils? // *Biophys. J.* – 2011. – Vol. 100. – P. 450-458.
37. Foderà V., Zaccone A., Lattuada M., Donald A. Electrostatics controls the formation of amyloid superstructures in protein aggregation // *Phys. Rev. Lett.* – 2013. – Vol. 111. – P. 108105-108109.
38. Di Michele L., Eiser E., Foderà V. Minimal model for self-catalysis in the formation of amyloid-like elongated fibrils // *J. Phys. Chem. Lett.* – 2013. – Vol. 4. – P. 3158-3164.
39. Auer S. Amyloid fibril nucleation: effect of amino acid hydrophobicity // *J. Phys. Chem. B.* – 2014. – Vol. 118. – P. 5289-5299.