

## Изучение фибриллообразования амилоидогенных участков рибосомных белков S1

Дерюшева Е.И.<sup>1</sup>, Мачулин А.В.<sup>2</sup>, Селиванова О.М.<sup>3</sup>, Гришин С.Ю.<sup>3</sup>, Глякина А.В.<sup>3</sup>, Сурин А.К.<sup>3,4</sup>, Галзитская О.В.<sup>3</sup>

<sup>1</sup>Институт биологического приборостроения РАН, ФИЦ ПНЦБИ РАН, Пуццино

<sup>2</sup>Институт биохимии и физиологии микроорганизмов им. Г.К. Скрыбина РАН, ФИЦ ПНЦБИ РАН, Пуццино

<sup>3</sup>Институт белка РАН, Пуццино

<sup>4</sup>Филиал института биоорганической химии им. М.М. Шемякина и Ю.А. Овчинникова РАН, Пуццино

[evgenia.deryusheva@gmail.com](mailto:evgenia.deryusheva@gmail.com), [ogalzit@vega.protres.ru](mailto:ogalzit@vega.protres.ru)

РНК-связывающий структурный домен S1, принадлежащий к суперсемейству олигосахарид/олигонуклеотид-связывающих доменов, широко распространён и консервативен для прокариот, архей и эукариот. Несмотря на известную склонность к агрегации рибосомных белков S1, способность к фибриллообразованию у амилоидогенных участков белков S1 мало изучена. В данной работе мы объединили биоинформатические и экспериментальные подходы для изучения особенностей амилоидогенных областей в этом семействе белков. Биоинформатические инструменты были использованы нами для оценки склонности к образованию фибрилл на примере 1331 последовательности рибосомного белка S1 и для идентификации таких областей в различных структурных доменах. Мы показали, что доля амилоидогенных участков для полноразмерных белков снижается с увеличением размера белков. Внутри отдельных доменов амилоидогенные области чаще предсказываются на N- и C-концевых участках. Для выбранных нами и синтезированных 11 амилоидогенных пептидов из рибосомных белков S1 *Escherichia coli* и *Thermus thermophilus* мы показали их способность формировать фибриллы методом электронной микроскопии.

*Ключевые слова:* белковый домен S1, амилоидогенные участки, рибосомный белок S1, фибриллы, FoldAmyloid.

## Investigation of Fibrillation by Amyloidogenic Regions of the Ribosomal S1 Proteins

Deryusheva E.I.<sup>1</sup>, Machulin A.V.<sup>2</sup>, Selivanova O.M.<sup>3</sup>, Grishin S.Yu.<sup>3</sup>, Glyakina A.V.<sup>3</sup>, Surin A.K.<sup>3,4</sup>, Galzit'skaya O.V.<sup>3</sup>

<sup>1</sup>Institute for Biological Instrumentation, Federal Research Center “Pushchino Scientific Center for Biological Research of the Russian Academy of Sciences, Pushchino

<sup>2</sup>Skryabin Institute of Biochemistry and Physiology of Microorganisms, Russian Academy of Sciences, Federal Research Center “Pushchino Scientific Center for Biological Research of the Russian Academy of Sciences, Pushchino

<sup>3</sup>Institute of Protein Research, Russian Academy of Sciences, Pushchino

<sup>4</sup>The Branch of the Institute of Bioorganic Chemistry, Russian Academy of Sciences, Pushchino,

RNA-binding structural S1 domain, belonging to the oligosaccharide/oligonucleotide-binding domain superfamily, is widely distributed and conserved in prokaryotes, archaea, and eukaryotes. Despite the known tendency to aggregation of ribosomal S1 proteins, the tendency of fibrillation by amyloidogenic regions of this family is poorly understood. In this paper, we combined bioinformatics and experimental approaches to study the features of amyloidogenic regions in this family of proteins. Bioinformatics tools were used to assess the propensity for amyloidogenesis on the example of 1331 sequences of ribosomal S1 protein and to identify such regions in various structural domains. We showed that the fraction of amyloidogenic regions for the full-size proteins decreases with increasing protein size. Within individual domains amyloidogenic regions are more often predicted in the C- and N-terminal parts. For the selected and synthesized 11 amyloidogenic peptides from the *Escherichia coli* and *Thermus thermophilus* ribosomal S1 proteins, we showed their ability to form amyloid-like fibrils by electron microscopy.

*Key words:* S1 domain, amyloidogenic regions, ribosomal S1 protein, fibrils, FoldAmyloid.

## 1. Введение

Многофункциональный рибосомный белок S1 является частью 30S субъединицы рибосомы и играет важную роль в инициации трансляции мРНК, участвует в элонгации, а также выполняет ряд внерибосомных функций [1–4].

Биоинформатический анализ последовательностей рибосомных белков S1 позволил нам продемонстрировать, что количество доменов в белках S1 является отличительной характеристикой для филогенетической классификации бактерий по основным отделам [5, 6]. Количество S1 доменов сильно варьируется в различных семействах эукариотических, архейных и бактериальных белков [7, 8]. Нами было показано, что центральные части в мультидоменных белках S1 имеют самый низкий процент гибкости. Кроме того, стабильный фрагмент белка S1 (49 кДа), менее склонен к образованию агрегатов, по сравнению с полноразмерным белком из *T. thermophilus* [9]. Полученные результаты в целом коррелируют с наличием повторов домена S1 в белках, содержащих домен S1, от одного (бактериальный и архейный) до пятнадцати (эукариотический), что, по-видимому, обусловлено эволюционной необходимостью увеличения специфичности связывания отдельных белков с их лигандами [7].

Несмотря на все эти наблюдения, поиск амилоидогенных областей белков S1 и исследование их фибриллообразования еще не проводились. Мы проанализировали тенденцию к фибрилlogenезу бактериальных рибосомных белков S1 с использованием репрезентативного набора данных, а также способность отдельных амилоидогенных областей образовывать различные фибриллярные структуры. Процесс формирования амилоидов может существенно зависеть от буферных условий [10–12]. Известно, что многие пептиды и белки способны образовывать агрегаты амилоидного типа при нейтральных значениях pH буферного раствора [13–15]. В то же время изменение pH раствора может активировать разворачивание и фибрилlogenез белков, а также влиять на морфологию фибрилл [16–18]. Изменение pH с 7.5 до 2.0 не приводило к образованию фибрилл для целого белка S1 (неопубликованные данные). Но, как показано в нашем исследовании, оно влияло на фибрилlogenез пептидов, синтезированных на основе последовательности S1 из *E. coli* и *T. thermophilus* [19].

## 2. Материалы и методы

### 2.1. Создание коллекции белковых последовательностей и анализ метаданных

Репрезентативный набор данных был выбран, как описано в [5]. Исследуемый набор состоял из 1331 записи. Алгоритмы поиска, сбора,

сопоставления, представления и анализа данных были реализованы с помощью языка программирования Python 3 (<https://www.python.org/>) в версии PyCharm v.2017 (<https://www.jetbrains.com/pycharm/>). Результаты данных были визуализированы с помощью библиотеки Matplotlib [20].

### 2.2. Предсказание амилоидогенных регионов

Для предсказания амилоидогенных регионов в белках S1 мы использовали следующие биоинформатические онлайн-инструменты: FoldAmyloid (<http://bioinfo.protres.ru/fold-amyloid/>) [21], Waltz (<http://waltz.switchlab.org>) [22, 23], PASTA 2.0 (<http://protein.bio.unipd.it/pasta2/>) [24, 25] и Aggrescan (<http://bioinf.uab.es/aggrescan/>) [26]. Алгоритмы этих программ позволяют выявлять аминокислотные последовательности, склонные к амилоидогенезу, как описано в [27].

Уникальность предсказанных амилоидогенных областей проверяли с помощью сервера BLAST (<https://blast.ncbi.nlm.nih.gov/Blast.cgi>) [28].

### 2.3. Синтез амилоидогенных пептидов

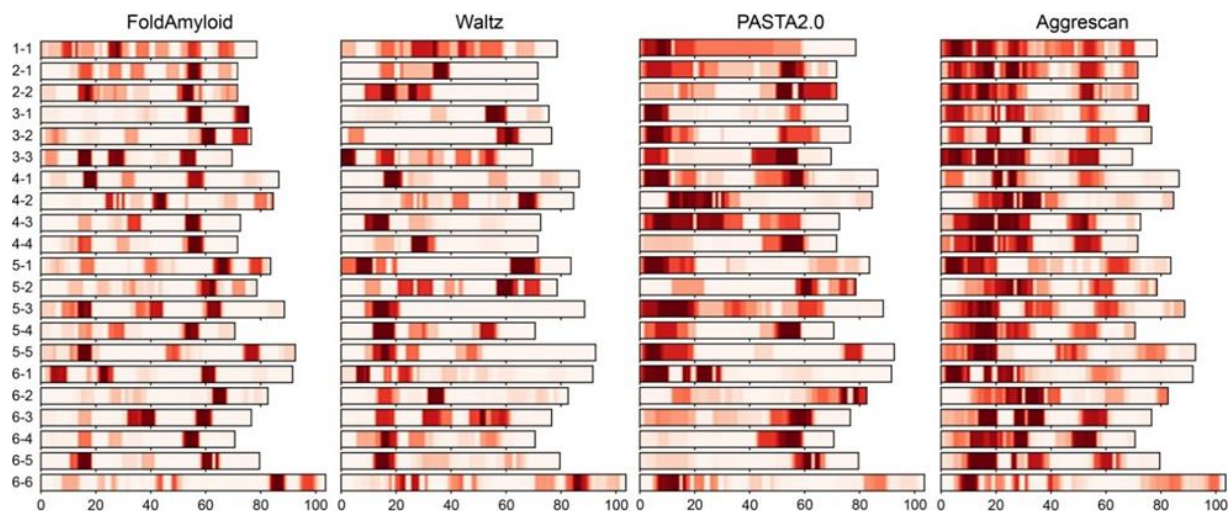
Для экспериментальной проверки были выбраны предсказанные амилоидогенные пептиды, полученные для белков S1 из модельных организмов *T. thermophilus* и *E. coli*. Соответственно четыре пептида (VVEGTVVEVT, DFGVFNLG, VTDFGVFVEI, EMEVVVLNID – для *T. thermophilus*) и семь пептидов (IVRGVVVAID, DEITVKVLKF, TDYGCVEIE, VNVGDVVEV, DFGIFGLDG, VHLSDISWNV, EIAAVVLQVD – для *E. coli*) получены твердофазным синтезом [29]. Синтез пептидов VVEGTVVEVT и VTDFGVFVEI осуществлялся в ФИБХ РАН (Россия), DFGVFNLG и EMEVVVLNID – в «ELABSCIENCE» (Китай), остальные пептиды – в «OOO IQChemical» (Россия). Соответствие пептидов проверяли с помощью масс-спектрометра Orbitrap Elite, «Thermo Scientific» (Германия).

### 2.4. Электронная микроскопия амилоидогенных пептидов

Для электронно-микроскопических исследований отбирали препараты пептидов с концентрацией 0.2 мМ после инкубации при 37 °С в буферах с pH 7.5 (50 мМ Tris-HCl, 150 мМ NaCl, 1 % ДМСО) и pH 2.0 (20 %-уксусная кислота, 150 мМ NaCl, 1 % ДМСО). 10 мкл препарата наносили на покрытые формваром медные сеточки, «Electron Microscopy Sciences» (США). После 5 минут абсорбции сеточки обрабатывали 1 %-раствором уранил ацетата в течение 1.5–2 минут. Полученные препараты были проанализированы с помощью трансмиссионного электронного микроскопа JEM-100С, «JEOL» (Япония), с ускоряющим напряжением 80 кВ. Изображения были получены на пленке SO-163, «Kodak» (USA).

### 3. Результаты и обсуждение

#### 3.1. Анализ распределения амилоидогенных участков в рибосомных белках S1



**Рис. 1.** Распределение амилоидогенных областей в отдельных (от 1 до 6) доменах S1 в соответствии с программами FoldAmyloid, Waltz, PASTA 2.0 и Aggrescan. Цветовая шкала соответствует частотному распределению амилоидогенных областей вдоль белковой цепи (выровненные последовательности белков). Размер каждого структурного домена составляет около 70 аминокислотных остатков (а.о.).

Как видно из рисунка 1, частота распределения амилоидогенных областей в доменах S1 позволяет выявить характерные особенности для каждого отдельного домена. Результаты программ, использованных для предсказания амилоидогенных участков белков S1, сильно различаются. Программы PASTA 2.0 и Aggrescan не обнаружили четко определенных (со строгими границами) амилоидогенных участков вдоль белковой цепи. В то же время программы FoldAmyloid и Waltz предсказали определенные области с высокой склонностью к фибриллообразованию. Наше недавнее исследование продемонстрировало низкий процент идентичности между отдельными доменами для выровненных последовательностей белков S1 [5]. Этот факт объясняет отсутствие строгих границ амилоидогенных областей, предсказываемых программой Aggrescan, которая основана на агрегационных свойствах отдельных аминокислотных остатков. PASTA 2.0 предсказывает формирование  $\beta$ -листов в белках в процессе самосборки кросс- $\beta$ -структур в амилоидных агрегатах. Однако, отдельные амилоидогенные области и  $\beta$ -листы в процессе агрегации могут вести себя по-разному. Программа Waltz основана на экспериментально существующих библиотеках амилоидов, поэтому

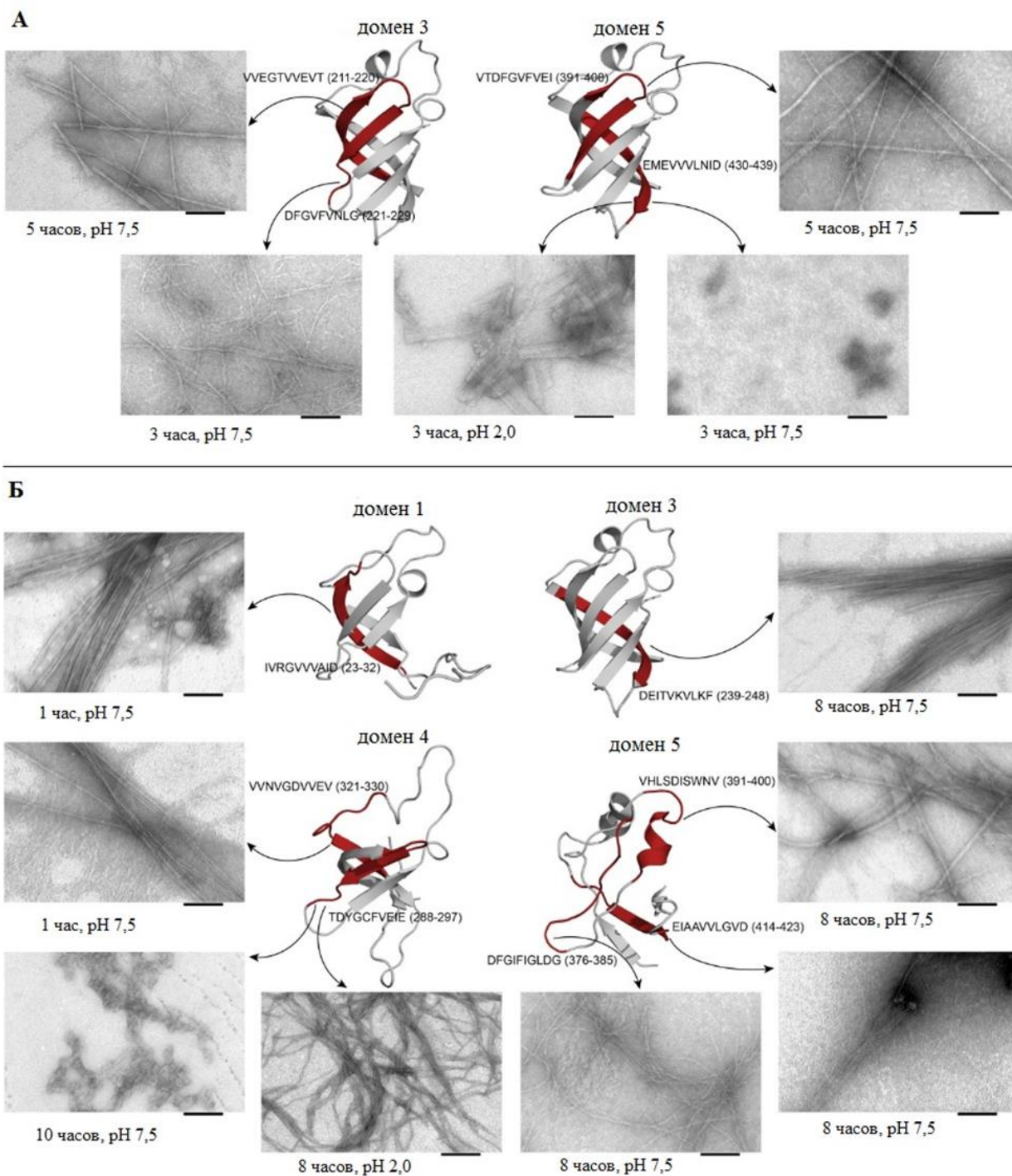
Частотные распределения амилоидогенных областей в отдельных доменах S1 показаны на рисунке 1.

полученные результаты можно считать надежными. FoldAmyloid использует ожидаемое количество контактов и вероятность образования водородных связей между аминокислотными остатками. В результате для нашего изученного набора данных FoldAmyloid предсказал регионы со строгими границами в каждом отдельном домене.

Наличие строго ограниченных амилоидогенных участков (по программам FoldAmyloid и Waltz) для каждой дополнительной группы S1 доменов (содержащих разное количество структурных доменов) и между этими группами позволяет рассматривать эти участки как уникальные и наиболее актуальные для дальнейших экспериментальных исследований, изучения их склонности к фибриллообразованию.

#### 3.2. Экспериментальная проверка амилоидогенности синтезированных пептидов

Мы изучили индуцированную нагреванием агрегацию пептидов рибосомных белков S1 из *E. coli* и *T. thermophilus*. Длительная инкубация при значениях pH 7.5 или 2.0 привела к образованию агрегатов с фибриллярной амилоидоподобной морфологией, видимой под электронным микроскопом (рис. 2).



**Рис. 2.** 3D-структуры доменов S1 из *T. thermophilus* (А) и *E. coli* (Б). Красным цветом выделены экспериментально изученные амилоидогенные участки (указаны положение и аминокислотная последовательность). 3D-структуры доменов 3 и 5 из *T. thermophilus* и домена 3 из *E. coli* были предсказаны с помощью сервера Robetta. Остальные 3D-структуры доменов из *E. coli* были получены из PDB (Protein Data Bank, <https://www.rcsb.org/>): домен 1 –2MFI (название файла); домен 4 –2KHI; домен 5 –5XQ5. Размер шкалы сравнения соответствует 100 нм.

Нами показано, что пептиды длиной 10 аминокислотных остатков, синтезированные на основе последовательности белка S1, могут формировать протяженные фибриллы (рисунок 2). В то же время короткий белок (один домен) может вызывать характерное для амилоидов увеличение

интенсивности флуоресценции тиофлавина Т, формирует сферо-подобные агрегаты, но не формирует фибриллы (например, S1 из *M. mobile*; данные не представлены). Недавно мы изучили способность к образованию амилоидоподобных фибрилл белка S1 (пять доменов) из *T. thermophilus*. Мы показали, что белок S1 (60 кДа) и его

стабильный фрагмент S1 (49 кДа) не формирует амилоиды, а только аморфные агрегаты при pH 7.5 и различной ионной силе [9]. То есть можно предположить, что с увеличением размера белка тенденция к фибрилlogenезу уменьшается, что определяется структурными особенностями (наличие нескольких S1 доменов) и функциями этого белка, в частности, образованием гомодимеров и связыванием с рибосомными белками в 30S рибосомной субъединице [1].

### 3. Благодарности

Использовались средства электронной микроскопии ЦКП ПНЦБИ РАН (№ 670266). Масс-спектрометрические измерения были выполнены при участии ЦКП ИБ РАН (№ 584307).

Исследование финансировалось РФФ, грант № 18-14-00321.

### 4. Список литературы

1. Wower I.K., Zwieb C.W., Guven S.A., Wower J. Binding and cross-linking of tmRNA to ribosomal protein S1, on and off the Escherichia coli ribosome. *EMBO J.* 2000. V. 19. P. 6612–6621. doi: [10.1093/emboj/19.23.6612](https://doi.org/10.1093/emboj/19.23.6612).
2. Hajnsdorf E., Boni I.V. Multiple activities of RNA-binding protein S1 and Hfq. *Biochimie.* 2012. V. 94. P. 1544–1553. doi: [10.1016/j.biochi.2012.02.010](https://doi.org/10.1016/j.biochi.2012.02.010).
3. Kutlubaeva Z.S., Chetverina H.V., Chetverin A.B. The Contribution of Ribosomal Protein S1 to the Structure and Function of Q $\beta$  Replicase. *Acta Naturae.* 2017. V. 9. P. 26–30. doi: [10.32607/20758251-2017-9-4-26-30](https://doi.org/10.32607/20758251-2017-9-4-26-30).
4. Lund P.E., Chatterjee S., Daher M., Walter N.G. Protein unties the pseudoknot: S1-mediated unfolding of RNA higher order structure. *Nucleic Acids Res.* 2020. V. 48. P. 2107–2125. doi: [10.1093/nar/gkz1166](https://doi.org/10.1093/nar/gkz1166).
5. Machulin A.V., Deryusheva E.I., Selivanova O.M.; Galzitskaya O.V. The number of domains in the ribosomal protein S1 as a hallmark of the phylogenetic grouping of bacteria. *PLoS ONE.* 2019. V. 14. e0221370. doi: [10.1371/journal.pone.0221370](https://doi.org/10.1371/journal.pone.0221370).
6. Deryusheva E.I., Machulin A.V., Selivanova O.M., Galzitskaya O.V. Taxonomic distribution, repeats, and functions of the S1 domain-containing proteins as members of the OB-fold family. *Proteins: Structure, Function and Bioinformatics.* 2017. V. 85. P. 602–613. doi: [10.1002/prot.25237](https://doi.org/10.1002/prot.25237).
7. Deryusheva E.I.; Machulin A.V., Matyunin M.A., Galzitskaya O.V. Investigation of the relationship between the S1 domain and its molecular functions derived from studies of the tertiary structure. *Molecules.* 2019. V. 24. 3681. doi: [10.3390/molecules24203681](https://doi.org/10.3390/molecules24203681).
8. Amir M., Kumar V., Dohare R., Islam A., Ahmad F., Hassan M.I. Sequence, structure and evolutionary analysis of cold shock domain proteins, a member of OB fold family. *J. Evol. Biol.* 2018. V. 31. P. 1903–1917. doi: [10.1111/jeb.13382](https://doi.org/10.1111/jeb.13382).
9. Grishin S.Y., Dzhus U.F., Selivanova O.M., Balobanov V.A., Surin A.K.; Galzitskaya O.V. Comparative Analysis of Aggregation of Thermus thermophilus Ribosomal Protein bS1 and Its Stable Fragment. *Biochemistry (Moscow).* 2020. V. 85. P. 344–354. doi: [10.1134/S0006297920030104](https://doi.org/10.1134/S0006297920030104).
10. Swaminathan R., Ravi V.K., Kumar S., Kumar M.V., Chandra N. Lysozyme: a model protein for amyloid research. *Adv. Protein Chem. Struct. Biol.* 2011. V. 84. P. 63–111. doi: [10.1016/B978-0-12-386483-3.00003-3](https://doi.org/10.1016/B978-0-12-386483-3.00003-3).
11. Brudar S., Hribar-Lee B. The Role of Buffers in Wild-Type HEWL Amyloid Fibril Formation Mechanism. *Biomolecules.* 2019. V. 9. 65. doi: [10.3390/biom9020065](https://doi.org/10.3390/biom9020065).
12. Shigeto S., Chang C.F., Hiramatsu H. Directly Probing Intermolecular Structural Change of a Core Fragment of  $\beta$ 2-Microglobulin Amyloid Fibrils with Low-Frequency Raman Spectroscopy. *J. Phys. Chem. B.* 2017. V. 121. P. 490–496. doi: [10.1021/acs.jpcc.6b10779](https://doi.org/10.1021/acs.jpcc.6b10779).
13. Moriarty G.M., Olson M.P., Atieh T.B., Janowska M.K., Khare S.D., Baum J. A pH-dependent switch promotes  $\beta$ -synuclein fibril formation via glutamate residues. *J. Biol. Chem.* 2017. V. 292. 16368–16379. doi: [10.1074/jbc.M117.780528](https://doi.org/10.1074/jbc.M117.780528).
14. Esbjörner E.K., Chan F., Rees E., Erdelyi M., Luheshi L.M., Bertocini C.W., Kaminski C.F., Dobson C.M., Kaminski Schierle G.S. Direct observations of amyloid  $\beta$  Self-assembly in live cells provide insights into differences in the kinetics of A $\beta$ (1-40) and A $\beta$ (1-42) aggregation. *Chem. Biol.* 2014. V. 21. P. 732–742. doi: [10.1016/j.chembiol.2014.03.014](https://doi.org/10.1016/j.chembiol.2014.03.014).
15. van der Wateren I.M., Knowles T.P.J., Buell A.K., Dobson C.M., Galvagnion C. C-terminal truncation of  $\alpha$ -synuclein promotes amyloid fibril amplification at physiological pH. *Chem. Sci.* 2018. V. 9. P. 5506–5516. doi: [10.1039/c8sc01109e](https://doi.org/10.1039/c8sc01109e).
16. Dyuysekina A.E., Dolgikh D.A., Samatova E.N., Tiktupulo E.I., Balobanov V.A., Bychkova V.E. pH-induced equilibrium unfolding of apomyoglobin: substitutions at conserved Trp14 and Met131 and non-conserved Val17 positions. *Biochemistry (Moscow).* 2008. V. 73. P. 693–701. doi: [10.1134/s0006297908060102](https://doi.org/10.1134/s0006297908060102).
17. Katina N.S., Ilyina N.B., Kashparov I.A., Balobanov V.A., Vasiliev V.D., Bychkova V.E. Apomyoglobin Mutants with Single Point Mutations at Val10 Can Form Amyloid Structures at Permissive Temperature. *Biochemistry*

- (Moscow). 2011. V. 76. P. 555–563. doi: [10.1134/S0006297911050051](https://doi.org/10.1134/S0006297911050051).
18. Iannuzzi C., Borriello M., Portaccio M., Irace G., Sirangelo I. Insights into Insulin Fibril Assembly at Physiological and Acidic pH and Related Amyloid Intrinsic Fluorescence. *Int. J. Mol. Sci.* 2017. V. 18. 2551. doi: [10.3390/ijms18122551](https://doi.org/10.3390/ijms18122551).
  19. Grishin S.Y., Deryusheva E.I., Machulin A.V., Selivanova O.M., Glyakina A.V., Gorbunova E.Y., Mustaeva L.G., Azev V.N., Rekstina V.V., Kalebina T.S., Surin A.K., Galzitskaya O.V. Amyloidogenic Propensities of Ribosomal S1 Proteins: Bioinformatics Screening and Experimental Checking. *Int. J. Mol. Sci.* 2020. V. 21. 5199. doi: [10.3390/ijms21155199](https://doi.org/10.3390/ijms21155199).
  20. Hunter J.D. Matplotlib: A 2D graphics environment. *Comput. Sci. Eng.* 2007. V. 9. P. 90–95. doi: [10.1109/MCSE.2007.55](https://doi.org/10.1109/MCSE.2007.55).
  21. Garbuzynskiy S.O., Lobanov M.Y., Galzitskaya O.V. FoldAmyloid: A method of prediction of amyloidogenic regions from protein sequence. *Bioinformatics.* 2010. V. 26. P. 326–332. doi: [10.1093/bioinformatics/btp691](https://doi.org/10.1093/bioinformatics/btp691).
  22. Oliveberg M. Waltz, an exciting new move in amyloid prediction. *Nat. Methods.* 2010. V. 7. P. 187–188. doi: [10.1038/nmeth0310-187](https://doi.org/10.1038/nmeth0310-187).
  23. Maurer-Stroh, S., Debulpaep M., Kuemmerer N., de la Paz M.L., Martins I.C., Reumers J., Morris K.L., Copland A., Serpell L., Serrano L. et al. Exploring the sequence determinants of amyloid structure using position-specific scoring matrices. *Nat. Methods.* 2010 V. 7. P. 237–242. doi: [10.1038/nmeth.1432](https://doi.org/10.1038/nmeth.1432).
  24. Walsh I., Seno F., Tosatto S.C.E., Trovato A. PASTA 2.0: An improved server for protein aggregation prediction. *Nucleic Acids Res.* 2014. V. 42. P. 301–307. doi: [10.1093/nar/gku399](https://doi.org/10.1093/nar/gku399).
  25. Trovato A., Seno F., Tosatto S.C.E. The PASTA server for protein aggregation prediction. *Protein Eng. Des. Sel.* 2007. V. 20. P. 521–523. doi: [10.1093/protein/gzm042](https://doi.org/10.1093/protein/gzm042).
  26. Conchillo-Solé O. de Groot, N.S.; Avilés F.X.; Vendrell J., Daura X., Ventura S. AGGRESCAN: A server for the prediction and evaluation of “hot spots” of aggregation in polypeptides. *BMC Bioinformatics.* 2007. V. 8. 65. doi: [10.1186/1471-2105-8-65](https://doi.org/10.1186/1471-2105-8-65).
  27. Surin A.K., Grishin S.Y., Galzitskaya O.V. Determination of amyloid core regions of insulin analogues fibrils. *Prion.* 2020. V. 14. P. 149–162. doi: [10.1080/19336896.2020.1776062](https://doi.org/10.1080/19336896.2020.1776062).
  28. Shirayev S.A., Papadopoulos J.S., Schaffer A.A., Agarwala R. Improved BLAST searches using longer words for protein seeding. *Bioinformatics.* 2007. V. 23. P. 2949–2951. doi: [10.1093/bioinformatics/btm479](https://doi.org/10.1093/bioinformatics/btm479).
  29. Beyermann M., Henklein P., Klose A., Sohr R., Bienert M. Effect of tertiary amine on the carbodiimide-mediated peptide synthesis. *Int. J. Pept. Protein Res.* 2009. V. 37. P. 252–256. doi: [10.1111/j.1399-3011.1991.tb00737.x](https://doi.org/10.1111/j.1399-3011.1991.tb00737.x).