Joint PHD Thesis

Between Wrocław University of Science and Technology, Department of Biomedical Engineering
and Université de Lorraine, École doctorale Chimie - Mécanique - Matériaux-Physique

Submitted to obtain the titles of
PhD in **Biomedical Engineering** at WUST
and of PhD in **Chemistry** at l'Université de Lorraine

Intrinsic and extrinsic determinants of the aggregation process of amyloid proteins

Déterminants intrinsèques et extrinsèques du processus d'agrégation des protéines amyloïdes

Natalia SZULC

Supervisors:

Prof. Małgorzata Kotulska, PhD, Eng	Wroclaw University of Science and Technology, Wrocław, Poland
Dr. Mounir Tarek	Université de Lorraine, Nancy, France

Science and everyday life cannot and should not be separated.

Science, for me, gives a partial explanation of life ...

Rosalind Franklin

Doctoral thesis co-financed by the European Union under the European Social Fund,
project no POWR.03.02.00-00-I003/16 and Bourses du Gouvernement Français – Doctorat cotutelle

Table of contents

Acknowledgements

Firstly, I would like to thank my supervisors: Prof. Małgorzata Kotulska and Dr. Mounir Tarek, for the support and guidance in my journey. Each of you, by sharing your knowledge, scientific and life experience with me, helped me become a researcher. Thank you for the trust that I was given to explore various projects independently. It was a pleasure to work with you.

Marlena thank you for co-supervising the experimental part of the studies. I really enjoyed that part. Thank you for all the fruitful discussions, especially those after your working hours. I want to thank you for being with me during wet lab times.

I am also very grateful to my colleague Monika, who brought me into the peptide synthesis world. Thank you for sharing your experience and knowledge.

I would like to thank my colleague Michał for helping me to better understand R, bioinformatics tools and sharing his knowledge regarding CsgA protein.

I am very thankful to my colleague Grégory Francius from Laboratoire de Chimie Physique et Microbiologie pour les Matériaux et l'Environnement, who is an expert in atomic force microscopy.

I also want to thank my friends Daniel, Aurianne, Vishal, Valentin, and Kuba, who helped and supported me during that journey, especially when I was living in France. I am a lucky person to be surrounded by such nice people.

To all the colleagues at the Department of Biomedical Engineering and at the Laboratoire de Physique et Chimie Théoriques, we thank you for your help and kindness.

Finally, I thank my family for all the support and motivation to push further. I am particularly grateful to my parents and my husband for their enormous support.

Funding

The research in the thesis was funded by:

- InterDok – Interdisciplinary Doctoral Studies Projects at Wroclaw University of Science and Technology, co-financed by the European Union under the European Social Fund no. POWR.03.02.00-00-I003/16.
- Bourses du Gouvernement Français – Doctorat cotutelle fellowship, Campus France.
- The National Science Centre, grant no. OPUS18 2019/35/B/NZ2/03997.

List of publications

Some results of this thesis are included in the publications:

1. **N. Szulc**, M. Burdukiewicz, M. E. Gąsior-Głogowska, J. W. Wojciechowski, J. Chilimoniuk, P. Mackiewicz, T. Šneideris, V. Smirnovas & M. Kotulska, (2021) *Bioinformatics methods for identification of amyloidogenic peptides show robustness to misannotated training data*, Scientific Reports, 11(1), 8934.

2. **N. Szulc,** M. Gąsior-Głogowska, J. W. Wojciechowski, M. Szefczyk, A. M. Żak, Michał Burdukiewicz & M. Kotulska, (2021) *Variability of amyloid propensity in imperfect repeats of csga protein of salmonella enterica and escherichia coli*, International Journal of Molecular Sciences 22(10), 5127.

3. M. E. Gąsior-Głogowska, **N. Szulc**, M. Szefczyk, (2022) *Challenges in experimental classification of amyloid peptides*, edited by Mai Suan Li, Marek Cieplak, Maksim Kouza and Andrzej Kloczkowski, Computer Simulations of Aggregation of Proteins and Peptides, Springer.

4. M. Burdukiewicz, D. Rafacz, A. Barbach, K. Hubicka, L. Bąkała, A. Lasota, J. Stecko, N. Szymanska, J. W. Wojciechowski, D. Kozakiewicz, **N. Szulc**, J. Chilimoniuk, I. Jęśkowiak, M. Gąsior-Głogowska, M. Kotulska, (2023) *AmyloGraph: A comprehensive database of amyloid-amyloid interactions*, Nucleic Acids Research, 51 (D1), D352–D357.

Other publications:

1. M. Szefczyk, **N. Szulc**, M. Gąsior-Głogowska, A. Modrak-Wójcik, A. Bzowska, W. Mastrzyk, M. Taube, M. Kozak, T. Gotszalk, E. Rudzińska-Szostak,. & Ł. Berlicki, (2021) *Hierarchical approach for the rational construction of helix-containing nanofibrils using α,β-peptides*, Nanoscale, 13(7), 4000–4015.

2. D. Wiczew, **N. Szulc**, M. Tarek, (2021) *Molecular dynamics simulations of the effects of lipid oxidation on the permeability of cell membranes*, Bioelectrochemistry, 141.

3. W. Dyrka, M. Gąsior-Głogowska, M. Szefczyk, **N. Szulc**, (2021) *Searching for universal model of amyloid signaling motifs using probabilistic context-free grammars*, BMC Bioinformatics, 22(1), 1-28.

4. J. W. Wojciechowski, E. Tekoglu, M. Gąsior-Głogowska, V. Coustou, **N. Szulc**, M. Szefczyk, M. Kopaczyńska, S. J. Saupe, W. Dyrka, (2022) *Exploring a diverse world of effector domains and amyloid signaling motifs in fungal NLR protein*s, PLoS Comput Biol 18(12): e1010787.

Conference presentations

1. **N. Szulc**, M. Gąsior-Głogowska, J. W. Wojciechowski, M. Szefczyk, A. M. Żak, M. Burdukiewicz, M. Kotulska, *The effect of deuterium oxide on the aggregation process of CsgA fragments*, XXII Polish Conference on Biocybernetics and Biomedical Engineering, 2021, oral contribution.
2. **N. Szulc**, M. Gąsior-Głogowska, P. Żyłka, J. W. Wojciechowski, M. Szefczyk, A. M. Żak, W. Dyrka, A. Kaczorowska, M. Burdukiewicz, M. Tarek, M. Kotulska, *Structural effects of charged mutations on amyloids explored by Contact Potential Difference method*, The Biochemistry Global Summit, 2022, Lisbon, Portugal, poster presentation.
3. **N. Szulc**, M. Gąsior-Głogowska, P. Żyłka, J. W. Wojciechowski, M. Szefczyk, A. M. Żak, W. Dyrka, A. Kaczorowska, M. Burdukiewicz, M. Tarek, M. Kotulska, *Structural effects of charge mutations in amyloid fragments of CsgA – quantitative analysis involving contact potential difference method*, The Leuven Protein Aggregation Meeting, 2022, Leuven, Belgium, poster presentation (received conference grant).

List of abbreviations

μIR – infrared microscopy
AD – Alzheimer's disease
AFM – atomic force microscopy
APP – amyloid-β precursor protein
APR – aggregation prone regions
ATR-FTIR – Fourier transform infrared spectroscopy with attenuated total reflectance
Aβ42 – amyloid-β
BACE1 – β-secretase
CD – circular dichroism
CPD – contact potential difference
C_{pep} – peptides' concentration
CPK – Corey-Pauling-Koltun coloring
c_R – critical concentration
CSF – cerebrospinal fluid
D_2O – deuterium oxide/heavy water
DIC – differential interference contrast
DMF – dimethylformamide
DMSO – dimethyl sulfoxide
DOPC – 1,2-dioleoyl-sn-glycero-3-phosphocholine
DPPC – dipalmitoylphosphatidylcholine
EM – Electron microscopy
ER4 – R4 fragment of *Escherichia coli*
FT-Raman – Fourier Transform-Raman spectroscopy
HEPES – 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HEWL – hen egg white lysozyme
HFIP – 1,1,1,3,3,3-hexafluoro-2-isopropanol
hIAPP – human islet amyloid polypeptide, called also amylin
HPLC – high-performance liquid chromatography
IDP – intrinsically disordered protein
IR – infrared
IRE – internal refractive element
k_{app} – growth rate
KPFM – Kelvin probe force microscopy
MCT–A – mercury cadmium telluride
MD – molecular dynamics simulations
MS – mass spectroscopy
MSA – multiple sequence alignment
Nd:YAG – neodymium-doped yttrium aluminum garnet
OXYMA – ethyl 2-cyano-2-(hydroxyimino)acetate

PBS – Phosphate-buffered saline
PC – Principal Components
PCA – principal component analysis
PDB – protein data bank
pH – potential of hydrogen
pI – isoelectric point
pK_a – negative base-10 logarithm of the acid dissociation constant (Ka) of a solution
POPC – phosphatidylcholine
PVDF – polyvinylidene difluoride
RDF – radial distribution function
RFU – relative fluorescence units
ROI – region of interest
SKP – Scanning Kelvin Probe
SLB – supported lipid bilayers
SNR – signal to noise ratio
SR4 – R4 fragment in *Salmonella enterica*
SUV – small unilamellar vesicle
$t_{1/2}$ – the time needed to reach halfway through the elongation phase
TEM – transmission electron microscopy
TFA – 2,2,2-trifluoroacetic acid
ThT – thioflavin T
TIS – trialkylsilane
WF – work function
WT – wild type

Abstract

Amyloids are a group of peptides and proteins that fold into assemblies of insoluble fibrils of very regular and tightly packed β-cross structures, which resemble a steric zipper. They are associated with many civilization diseases, such as type 2 diabetes, and wide range of neurodegenerative diseases. However, due to their self-assembly properties and unique physicochemical characteristics, as well as favorable mechanical and optical features, amyloids can and are being used in different fields of material and life sciences.

Amyloids are very sensitive to various experimental conditions in which their aggregation process is studied. Moreover, experimental conditions have a great influence on the amyloid characteristics and classification. In many cases, data collected from experiments using different techniques and performed by various scientific groups, do not include all the information about the specifics regarding experimental conditions. This has a strong impact on the ability, for instance of bioinformatics methods, to predict the amyloid propensity or the structural models of aggregation.

This dissertation focuses on the influence of selected internal and external factors affecting the aggregation process of amyloidogenic peptides. In the study, we consider in particular the length of the amino acid sequence, the effect of mutations, the influence of the solvent including, type and salt concentration, the interactions with other peptides, as well as the proximity and interaction with a cell wall (lipid environment). This examination concerned selected pathological and functional amyloids and their most relevant fragments, and relied on results from experimental and computational investigations.

First, we set a reference protocol to study the aggregation properties of short peptides, necessary to identify possible amyloidogenic amino acid sequences. Hence, we selected hexapeptides using AmyloGram, a well-established bioinformatics prediction tool. Then we examined whether and to what extent, poorly annotated training data can influence the accuracy of bioinformatics methods. Finally, we investigated whether longer sequences can be treated in the same way as short ones.

Second, we examined experimentally long amino acid sequences (up to 23 amino acids) from the functional amyloid CsgA protein (consisting of R1–R5 imperfect fragments) in order to check their amyloid propensity. We investigated whether such peptides are more sensitive than pathological amyloids to selected internal and external factors, such as: point mutations, solvent, type of ions, ions' concentration. To do so we compared the effects of these parameters on the aggregation properties of homologous repeats from two bacteria *Escherichia coli* and *Salmonella enterica*. This study included the influence of mutations, which has been validated by computational and theoretical studies.

We finally conducted a study on pathological amyloids (sequences up to 42 amino acids), namely Aβ42 and hIAPP. To mimic physiological conditions, experiments and simulations were carried out in the presence of a lipid membrane. Therefore, we investigated the influence of lipid membrane and interaction with another peptide on aggregation of these amyloids.

The results of the research present that it is important to collect coherent data from experiments, which could be disrupted by non-identical conditions. Additionally, different experimental methods may deliver somehow different results regarding the same peptides. This effect is more significant in case of some sequences.

We found that short hexapeptide sequences exhibit distinct aggregation propensities in response to external factors, e.g., the solvent used, compared to longer sequences (up to 23 amino acid). The flexibility of six-amino acid sequences allows them to adopt specific conformations. Moreover, the symmetry-breaking transitions phenomenon plays a crucial role in this process.

Additionally, we showed that the choice of solvent influences on the aggregation process. The usage of deuterium oxide might alter the classification, as observed for the R2 fragment of *S. enterica*. Fibrils were detected in the presence of heavy water but not in phosphate buffered saline. The dominant peptide conformation in D_2O was attributed to intermolecular aggregates, a signature typical of amyloid structures. Based on these findings, the R2 fragment's classification as amyloid or non-amyloid depends on the conditions applied.

Next, we showed that R4 fragment from *S. enterica* (SR4) has a larger tendency to aggregate compared to the R4 fragment from *E. coli* (ER4). Theoretical sequence analysis revealed that SR4 has higher hydrophobicity and electric charge values than ER4. Those factors are known to increase protein aggregation. Additionally, SR4 has more significant variations in gatekeeper residues, which play a crucial role in regulating amyloid formation. The results indicate that specific amino acid substitutions can significantly affect the propensity of functional amyloid proteins to form amyloid structures, with potential functional consequences.

The research highlighted the influence of the phosphate buffer ions on the peptides' morphology, affecting the local electrostatic interactions involving the polypeptide chains. The molecular dynamics simulations revealed that the interactions of positively charged amino acids with negatively charged phosphate moieties in the buffer determined the morphology. Non-specific ion enrichment may occur in the proximity of protein moieties that bear charges opposite to those of the ions. Hence, the dimers structures observed in the simulation may have contributed to the snail-like conformation of the M4 peptide.

In the last phase of the study, the stability of Aβ42 in a lipid membrane was examined through molecular dynamics simulations and atomic force microscopy. The results suggested that Aβ42

remains embedded in the lipid membrane, indicating that peptide aggregation may occur within the membrane. It was also found that Aβ42 interacting with hIAPP exhibited a higher β-sheet content when in close proximity, compared to the analysis of Aβ42 alone. Atomic force microscopy investigations showed that supported lipid bilayers were more affected in the presence of both Aβ42 and hIAPP.

Overall, the study sheds light on the potential challenges arising from the ambiguity of experimental outcomes in the context of amyloid investigations. Our research has shown that even small changes in experimental conditions can alter the properties of amyloid peptides and proteins, which can be the cause of obtaining incorrect models and predictions of bioinformatics tools, based on incompatible learning data. However, as shown in the study, minor deviations are not detrimental to these tools. Bioinformatics methods are fairly robust to incompatibilities in the data, but their performance can be disrupted if the influence of experimental conditions and seemingly insignificant differences in sequences of homological sequences are not taken into the consideration. The results described above, confirm that in the process of planning experimental research is important to choose proper conditions according to the studied object. Based on that choice, we can effectively modulate the peptide tendency to aggregate.

Résumé

Les amyloïdes sont un groupe de peptides ou de protéines qui se replient en assemblages de fibrilles insolubles de structure β-croiséss très régulière et étroitement embriqués, ressemblant à une fermeture à glissière. Ils sont associés à de nombreuses maladies, telles que le diabète de type 2, ainsi qu'à un large éventail de maladies neurodégénératives. Cependant, en raison de leurs propriétés d'auto-assemblage, de leurs caractéristiques physico-chimiques uniques, ainsi que de leurs caractéristiques mécaniques et optiques particulieres, les amyloïdes sont utilisés dans différents domaines des sciences des matériaux et sciences de la vie.

Les amyloïdes sont très sensibles aux différentes conditions expérimentales dans lesquelles leur processus d'agrégation est accompli. De plus, les conditions expérimentales ont une grande influence sur les caractéristiques et la classification des amyloïdes. Dans de nombreux cas, les données collectées à partir d'expériences utilisant différentes techniques et réalisées par différents groupes recherche n'incluent pas toutes les informations sur les spécificités relatives aux conditions expérimentales. Cela a un fort impact sur la capacité, par exemple des méthodes de bioinformatique, à prédire la propension de l'amyloïde à s'agréger.

Cette thèse se concentre sur l'influence de certains facteurs internes et externes sélectionnés qui affectent le processus d'agrégation de peptides amyloïdogènes. Dans cette étude, nous avons particulièrement considéré la longueur des sequences d'acides aminés, les'effets de mutations, l'influence du solvant, dont son type et la concentration saline, les interactions avec d'autres peptides, ainsi que la proximité et l'interaction avec une paroi cellulaire (environnement lipidique). Cette analyse a porté sur une sélection d'amyloïdes pathologiques et fonctionnels et sur leurs fragments les plus pertinents. Elle s'est appuyée sur des simulations numériques et résultats expérimentaux.

Tout d'abord, nous avons établi un protocole de référence pour étudier les propriétés d'agrégation de peptides courts, nécessaire pour identifier les séquences d'acides aminés amyloïdogènes possibles. Pour ce faire, nous avons sélectionné des hexapeptides à l'aide d'AmyloGram, un outil de prédiction bioinformatique bien établi. Ensuite, nous avons examiné si et dans quelle mesure les données d'entraînement peu annotées peuvent influencer l'exactitude des méthodes de bioinformatique. Enfin, nous avons étudié si les séquences plus longues peuvent être traitées de la même manière que les séquences courtes.

Ensuite, nous avons examiné expérimentalement de longues séquences d'acides aminés (jusqu'à 23 acides aminés) de la protéine amyloïde fonctionnelle CsgA (composée de fragments R1–R5) afin de vérifier leur propension à former des amyloïdes. Nous avons étudié si de tels peptides sont plus sensibles que les amyloïdes pathologiques à une sélection de facteurs internes et externes. Les facteurs

choisis sont les mutations ponctuelles, le solvant, le type d'ions et la concentration en ions. Pour ce faire, nous avons comparé les effets de ces paramètres sur les propriétés d'agrégation de sequences homologues de deux bactéries *Escherichia coli* et *Salmonella enterica*. Cette étude a inclus l'influence des mutations, qui a été validée par des études computationnelles et théoriques.

Enfin, nous avons mené une étude sur des amyloïdes pathologiques (séquences jusqu'à 42 acides aminés), à savoir Aβ42 et hIAPP. Pour imiter les conditions physiologiques, des expériences et des simulations ont été réalisées en présence d'une membrane lipidique. Nous avons donc étudié l'influence de la membrane lipidique et de l'interaction avec un autre peptide sur l'agrégation de ces amyloïdes.

Les résultats de la recherche montrent qu'il est important de collecter des données cohérentes à partir d'expériences, qui pourraient être perturbées par des conditions non identiques. De plus, différentes méthodes expérimentales peuvent fournir des résultats quelque peu différents pour les mêmes peptides. Cet effet est plus significatif dans le cas de certaines séquences.

Nous avons découvert que la propension à s'agréger des courtes séquences hexapeptidiques se distingue de celle des séquences plus longues (jusqu'à 23 acides aminés) en fonctions de facteurs externes comme le solvant utilize. La flexibilité des séquences de six acides aminés leur permet d'adopter des conformations spécifiques. De plus, le phénomène de transition de brisure de symétrie joue un rôle crucial dans ce processus.

De plus, nous avons montré que le choix du solvant influence le processus d'agrégation. L'utilisation d'oxyde de deuterium (D_2O) peut modifier la classification, comme observé pour le fragment R2 de *S. enterica*. Des fibrilles ont été détectées en présence d'eau lourde mais pas en solution tampon phosphate. La conformation dominante du peptide en D_2O a été attribuée à des agrégats intermoléculaires, une signature typique des structures amyloïdes. Sur la base de ces résultats, la classification du fragment R2 comme amyloïde ou non-amyloïde dépend des conditions appliquées.

Ensuite, nous avons montré que le fragment R4 de *S. enterica* (SR4) a une tendance plus forte à l'agrégation par rapport au fragment R4 d'*E. coli* (ER4). L'analyse théorique de la séquence a révélé que SR4 avait des valeurs d'hydrophobicité et de charge électrique plus élevées que ER4. Ces facteurs sont connus pour augmenter l'agrégation des protéines. De plus, SR4 présente des variations plus importantes dans les résidus dit « gate Keeper », qui jouent un rôle crucial dans la régulation de la formation amyloïde. Les résultats indiquent que des substitutions d'acides aminés spécifiques peuvent affecter significativement la propension des protéines amyloïdes fonctionnelles à former des structures amyloïdes, avec des conséquences fonctionnelles potentielles.

Nos investigations ont mis en évidence l'influence des ions tampons phosphate sur la morphologie des peptides, en affectant les interactions électrostatiques locales impliquant les chaînes

polypeptidiques. Les simulations de dynamique moléculaire ont révélé que les interactions des acides aminés chargés positivement avec les groupes phosphate chargés négativement dans le tampon déterminent la morphologie des fibrilles. Une concentration d'ions non spécifiques peut se produire à proximité des parties protéiques portant des charges opposées à celles des ions. Ainsi, les structures dimères observées dans la simulation ont peut-être contribué à la conformation en forme d'escargot du peptide M4.

Dans la dernière section de ce travail, la stabilité de Aβ42 dans une membrane lipidique a été examinée à travers des simulations de dynamique moléculaire et l'utilisation d'un microscope à force atomique. Les résultats suggèrent que Aβ42 reste intégré dans la membrane lipidique, indiquant que l'agrégation du peptide peut se produire dans la membrane. Il a également été constaté que Aβ42 interagissant avec hIAPP présente une teneur en feuillet β plus élevée lorsqu'ils sont en proximité étroite, par rapport à l'analyse de Aβ42 seul. Les observations sous microscope à force atomique ont montré que les bicouches lipidiques supportées étaient plus affectées en présence de Aβ42 et hIAPP.

Dans l'ensemble, l'étude met en lumière les défis découlant de la forte dépendance des résultats en fonction des conditions expérimentales des propriétés des amyloïdes. Dans l'ensemble, nous avons montré que les conditions expérimentales jouent un rôle prépondérant dans la détermination des propriétés des amyloïdes. Ne pas en tenir compte lors de l'utilisation d'outils de bioinformatique peut induire des erreurs dans les modèles et predictions. Toutefois, comme nous l'avons montré, les écarts mineurs ne sont pas préjudiciables. Les méthodes de bioinformatique sont assez robustes aux incompatibilités dans les données, mais leurs performances peuvent être perturbées si l'influence des conditions expérimentales et les différences apparemment insignifiantes dans les séquences de séquences homologues ne sont pas prises en considération. En conclusion, la planification du processus de la recherche expérimentale, requiert de choisir les bonnes conditions en fonction de l'objet étudié. Sur la base de ce choix, nous pouvons moduler efficacement la tendance du peptide à s'agréger.

Streszczenie

Amyloidy to grupa peptydów i białek, które tworzą nierozpuszczalne włókna, o bardzo regularnej i ciasno upakowanej strukturze β-cross, przypominającą strukturę zamka błyskawicznego. Są one związane z wieloma chorobami cywilizacyjnymi, takimi jak cukrzyca typu 2 i choroby neurodegeneracyjne. Pomimo tego, dzięki swoim unikalnym cechom, takim jak zdolność do agregacji oraz korzystne właściwości fizykochemiczne, mechaniczne i optyczne, amyloidy mogą i są wykorzystywane w różnych dziedzinach naukowych.

Amyloidy są bardzo wrażliwe na różne warunki eksperymentalne, w których badany jest proces agregacji. Warunki eksperymentu mają duży wpływ na własności i klasyfikacje amyloidów. W wielu przypadkach dane zebrane z eksperymentów wykorzystujących różne techniki i przeprowadzane przez różne grupy naukowe, nie zawierają wszystkich informacji o istotnych szczegółach dotyczących warunków eksperymentalnych. Fakt ten może mieć istotny wpływ na skuteczność stosowanych metod bioinformatycznych, których celem jest przewidywanie skłonności amyloidów do agregacji.

Prezentowana praca doktorska skoncentrowana jest na opisie wpływu wybranych czynników wewnętrznych i zewnętrznych środowiska eksperymentalnego, na proces agregacji peptydów amyloidowych. W badaniach rozważono przede wszystkim: długość sekwencji aminokwasów, wpływ mutacji, wpływ rozpuszczalnika, a w tym rodzaju i stężenia soli, interakcje z innymi peptydami oraz bliskość i interakcję ze ścianą komórkową (środowisko lipidowe). Opisane badania dotyczyły wybranych amyloidów patologicznych i funkcjonalnych oraz ich najistotniejszych fragmentów. Badania przeprowadzono metodami eksperymentalnymi i obliczeniowymi.

Pierwszy etap badań dotyczył ustanowienia standardowego protokołu badania właściwości agregacyjnych krótkich peptydów, koniecznego do identyfikacji możliwych sekwencji aminokwasowych o charakterze amyloidogennym. Wykorzystano w tym celu narzędzie bioinformatyczne AmyloGram. Następnie zbadano, czy i w jakim stopniu niewłaściwie opisane dane treningowe mogą wpłynąć na wyniki metod bioinformatycznych. Ostatecznie, sprawdzono, czy dłuższe sekwencje mogą być traktowane w taki sam sposób jak krótsze.

W drugiej części badań wykorzystano metody eksperymentalne, zbadano długie sekwencje aminokwasowe (do 23 aminokwasów) pochodzące z funkcjonalnego amyloidu CsgA (składającego się z fragmentów R1–R5), w celu sprawdzenia czy posiadają skłonność do tworzenia amyloidów. Zbadano czy takie peptydy są bardziej wrażliwe niż amyloidy patologiczne na wybrane czynniki wewnętrzne i zewnętrzne, takie jak: mutacje punktowe, rozpuszczalnik, rodzaj jonów, stężenie jonów. Porównano wpływ tych parametrów na właściwości agregacyjne homologicznych fragmentów pochodzących

z dwóch bakterii: *Escherichia coli* i *Salmonella enterica*. W tym przypadku uwzględniono również wpływ mutacji, który został zweryfikowany przez badania obliczeniowe i teoretyczne.

W końcowej części przeprowadzono badania na amyloidach patologicznych: Aβ42 i hIAPP (sekwencje do 42 aminokwasów). W celu naśladowania warunków fizjologicznych, eksperymenty i symulacje zostały przeprowadzone w obecności błony lipidowej. W tym celu zbadano wpływ błony lipidowej i interakcji z innym peptydem na agregację tych amyloidów.

Wyniki badań wskazują na istotność zbierania spójnych danych z eksperymentów przeprowadzonych w nieidentycznych warunkach. Potwierdzono, że różne metody eksperymentalne mogą dostarczać różne wyniki w odniesieniu do tych samych peptydów. Różnica ta jest bardziej widoczna w przypadku niektórych sekwencji.

Stwierdzono, że krótkie sekwencje heksapeptydów wykazują wyraźne skłonności do agregacji w odpowiedzi na czynniki zewnętrzne, takie jak stosowany rozpuszczalnik, w porównaniu z dłuższymi sekwencjami (do 23 aminokwasów). Elastyczność sekwencji heksapeptydów pozwala im na przyjmowanie określonych konformacji, a koncepcja przejść łamiących symetrię odgrywa istotną rolę w tym procesie.

Dodatkowo wykazano, że wybór rozpuszczalnika ma wpływ na proces agregacji. Użycie rozpuszczalnika D_2O może zmienić, na przykład, klasyfikację fragmentu R2 pochodzącego z bakterii *S. enterica*. Włókna wykryto w obecności ciężkiej wody, ale nie w roztworze NaOH+PBS. Dominująca konformacja w D_2O została przypisana do agregatów międzymolekularnych, sygnatury typowej dla struktur amyloidowych. Na podstawie tych wyników klasyfikacja fragmentu R2 jako amyloidowego lub nieamyloidowego zależy od zastosowanych warunków eksperymentalnych.

Badania wykazały, że fragment R4 pochodzący z *S. enterica* (SR4) ma większą skłonność do agregacji niż fragment R4 pochodzący z *E. coli* (ER4). Teoretyczna analiza sekwencji wykazała, że SR4 ma wyższą hydrofobowość i wartości ładunku niż ER4. Czynniki te są znane z przyspieszania agregacji białek. Dodatkowo, peptyd SR4 posiada znaczące różnice w aminokwasach regulujących proces agregacji, tzw. „gatekeeper", które odgrywają istotną rolę w regulowaniu formowania amyloidów. Wyniki wskazują, że konkretne substytucje aminokwasów mogą znacząco wpływać na skłonność funkcjonalnych białek amyloidowych do tworzenia struktur amyloidowych, co może mieć potencjalne konsekwencje funkcjonalne.

Badania podkreśliły wpływ jonów buforowych fosforanów na morfologię peptydów, wpływając na lokalne oddziaływania elektrostatyczne między łańcuchami polipeptydowymi. Symulacje dynamiki molekularnej pokazały, że interakcje aminokwasów naładowanych dodatnio z ujemnie naładowanymi resztami fosforanowymi w buforze przyczyniły się do powstałej morfologii agregatu.

Zauważono, że mutacje jak i obecność jonów, dla peptydu M4, mogą przyczynić się do powstania kulistego agregatu, który przedstawia włókno zwinięte w sferyczną strukturę.

W ostatniej fazie, badana była stabilność peptydu Aβ42 w błonie lipidowej za pomocą symulacji dynamiki molekularnej i badań mikroskopii sił atomowych. Wykazano, że peptyd Aβ42 jest stabilny w błonie lipidowej, co wskazuje, że agregacja peptydów może zachodzić wewnątrz błony. Stwierdzono również, że Aβ42 w interakcji z hIAPP wykazywała wyższą zawartość β-struktur przy bliskim odziaływaniu peptydów, w porównaniu do peptydu Aβ42 samodzielnie. Badania mikroskopii sił atomowych wykazały, że dwuwarstwy lipidowe były również bardziej uszkodzone w obecności obu peptydów (Aβ42 i hIAPP).

Badania przedstawione w pracy doktorskiej wskazały na potencjalne wyzwania wynikające z niejednoznaczności wyników eksperymentów w kontekście badań amyloidów. Pokazały, że nawet niewielkie różnice w warunkach eksperymentalnych mogą zmienić właściwości peptydów i białek amyloidowych, co może prowadzić do otrzymywania niepoprawnych modeli i prognoz narzędzi bioinformatycznych opartych na niespójnych danych uczących. Jednak, jak pokazano w badaniu, niewielkie błędy w danych uczących nie są szkodliwe dla tych narzędzi. Metody bioinformatyczne są dość odporne na niezgodności w danych, ale ich wydajność może zostać zaburzona, jeśli nie uwzględni się wpływu warunków eksperymentalnych i nieznacznych różnic w sekwencjach homologicznych. Opisane powyżej wyniki potwierdzają, że w procesie planowania badań eksperymentalnych ważne jest wybranie odpowiednich warunków zgodnie z badanym obiektem. Na podstawie tego wyboru możemy skutecznie modulować skłonność peptydów do agregacji.

Introduction

1.1. Amyloids

Protein misfolding events might result in the creation of insoluble aggregates, called amyloids [1]. The latter consist of very characteristic cross-β-sheet structures (see Figure 1), built up from stacked β-strands running perpendicularly to the long axis of the fibril [2]. Recent research revealed however, that amyloids structure can also exhibit cross-α structures, as shown in the case of aggregation of phenol-soluble modulin α3 (PSMα3) [3].

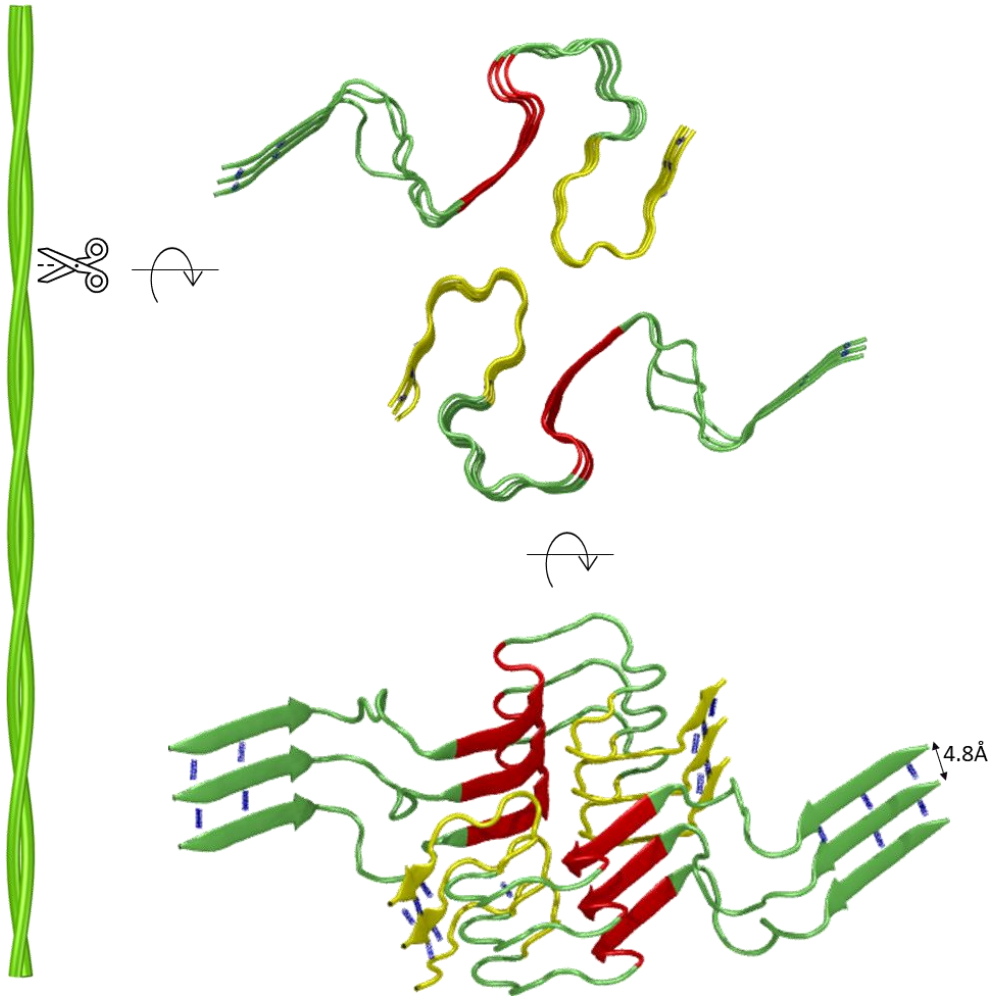


Figure 1 Schematic representation of tertiary amyloid fibril structure (dimers) from solution Nuclear Magnetic Resonance of amyloid-β(42) (PDB: 2NAO). In blue the hydrogen bonds between beta-sheets.

Thanks to their hierarchical and ordered structure, amyloids exhibit unique characteristics. Amyloid fibrils are elongated, unbranched protofilaments typically of ~5–10 nanometers width and a few micrometers length [4]. These fibrils possess important property as well: polymorphism [5], which relates to the fact that a given protein can adopt different structures [6].

The general definition of amyloid term is based on the structural and biophysical properties of these assemblies. There are hence three branches of amyloid groups: pathological, artificial and functional [7]. The first group relates to amyloids associated with neurodegenerative disorders. Currently, about 37 peptides or proteins [8] are associated with human pathologies e.g.: Alzheimer's disease (AD), type 2 diabetes (T2D), Parkinson's disease, Creutzfeldt-Jacob disease [9] etc. Artificial amyloids are a group of (globular) protein e.g. myoglobin that form amyloid-like fibrils under denaturing conditions [10]. Functional amyloids are those that play beneficial physiological roles. Thanks to evolutionary adaptations, the aggregation process of functional amyloids is well controlled [11] to fulfil their role in a biological machinery. They share the same structural cross-β structure with pathological amyloids, but they differ in their activity and stability. Five main classes of functional amyloids are known, among which are those which maintain the shape or support, such as: chaplin fibrils or curli fibrils. Thanks to these evolutionary adaptations, functional amyloids are fully controlled biological machinery: their aggregation process is fully controlled [11].

1.2. Aggregation process

Proteins are dynamic structures, which undergo many reversible conformational changes. However, due to misfolding, genetic mutations, failure of chaperons machinery or undesired proteolytic cleavage they may lose part or all of their specific three-dimensional conformations [12]. As a consequence, proteins in their native state misfold into non-functional and cytotoxic protein aggregates [13]. In which their native tertiary but also secondary structure changes. Chiti and Dobson [8] reported in their recent review that in unordered peptide/proteins the characteristic transition from α-helix to β-sheet in the secondary structure conformation must occur. Experimental investigation showed that it is the case for the Aβ42 amyloid [14]. However, such a structural transition is often a complicated, multistage process, as the balance between the various transition forms must be maintained. Remarkably, the amyloid state has been reported to be more thermodynamically advantageous then the native one [15]. The energy landscape of amyloidogenic peptides (see Figure 2) has a smoother funnel-shaped surface in comparison to larger proteins, which have a more rough landscape [16]. Figure 2 depicts several local minima and dominant global free energy minima, which correspond to amyloid fibrils. These dominant global minima consist of several minima that correspond to the polymorphic nature of fibril structures. The lower the value of free energy and sharper the minimum, the more durable and stable are the structures [17], [18]. High energy barriers characterize the transitions between oligomeric and fibril state in the aggregation process

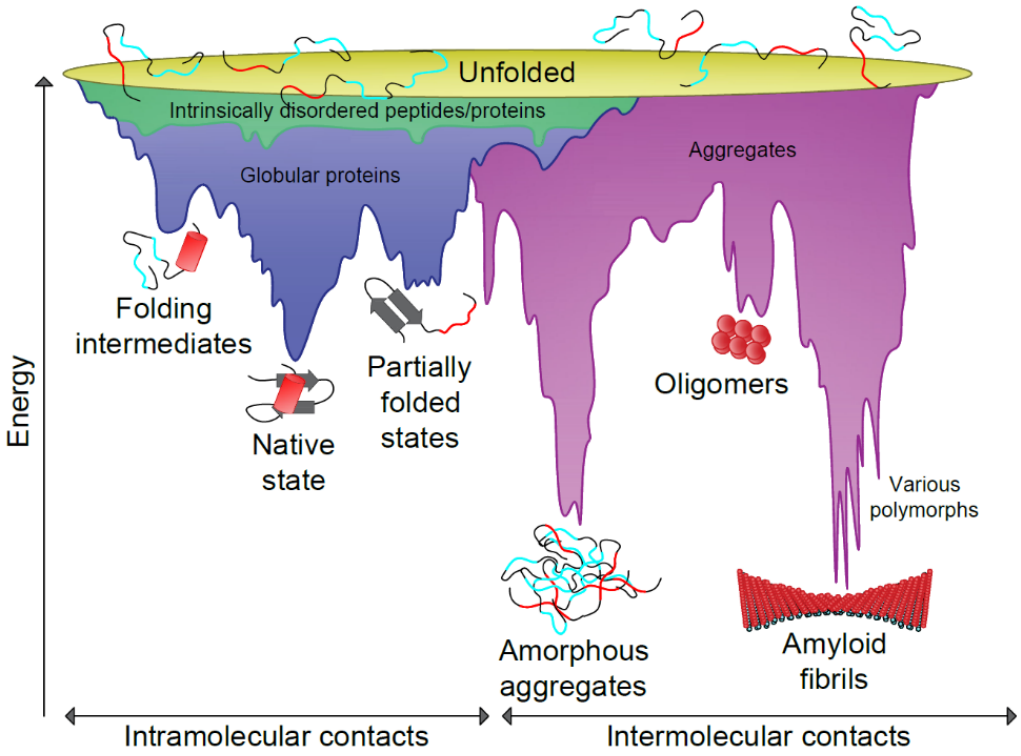


Figure 2 Schematic representation of the folding energy landscape and aggregation process. The intramolecular contacts correspond to the amino acid interactions within single protein chain of globular proteins (dark blue surface), of intrinsically disordered peptide/proteins (green surface). The intermolecular contacts involve interactions between various polypeptide chains (pink surface). Based on [19].

The phase transitions noticed on Figure 2, correspond to different stages of amyloid formation. The amyloid aggregation process starts with the formation of monomeric forms. Then trough association of molecules, the oligomeric state can occur. Oligomers can be formed by dimers, or multimers, which adopt roughly globular in shape [20]. The oligomeric structures are stabilized by additional cross interactions between the peptides/proteins amino acid side chains (pink area in Figure 2), favoring intermolecular contacts. These include formation of numerous noncovalent molecular interactions, specifically hydrogen bonds, π–π stacking and hydrophobic ones [21]. The latter characterize the thermodynamic signature of amyloid fibril growth [22]. All additional intermolecular interactions among aggregates cause destabilization of the native structures and changes in the hydrogen bonding network. Water is squeezed out between chains involved and a dry steric-zipper is formed [23]. The amyloid fibrils stability is hence not only due to specific interactions between amino acid side chains from adjacent peptide chains but also mediated by specific, well defined protein-solvent interactions. Protein or peptide aggregation involves serval phases, which are thermodynamically and kinetically favorable to form amyloid fibrils [24].

Noticeably, mutations and environmental factors, may affect as well the landscape and the pathways towards various energy minima.

1.3. Kinetics of the aggregation process

The aggregation process can involve several mechanisms and depends on various experimental conditions [9]. The mechanisms involved are either linear polymerization (nucleation-independent) or nucleation-dependent polymerization [25]. The nucleation-dependent polymerization is the most frequent model, in which the aggregation process is controlled by peptide/protein concentration and time. The most important characteristic of this mechanism is that no aggregation of peptide/protein occurs below a critical concentration (c_R). Second, whenever the concentration exceeds the c_R value, a lag phase is observed, during which addition of seeds (fibrils, other peptides/proteins) effects the polymerization process [26].

The nucleation-dependent mechanism starts with a nucleation phase, a stochastic process, in which nuclei are formed (see Figure 3). This process is the slowest one during the whole aggregation [26]. Nuclei consists of randomly attached monomers and are known as oligomers [27]. These intermediates are considered as the most toxic ones [28], [29] and play an essential role in the patho-mechanism of a neurodegenerative disease [30].

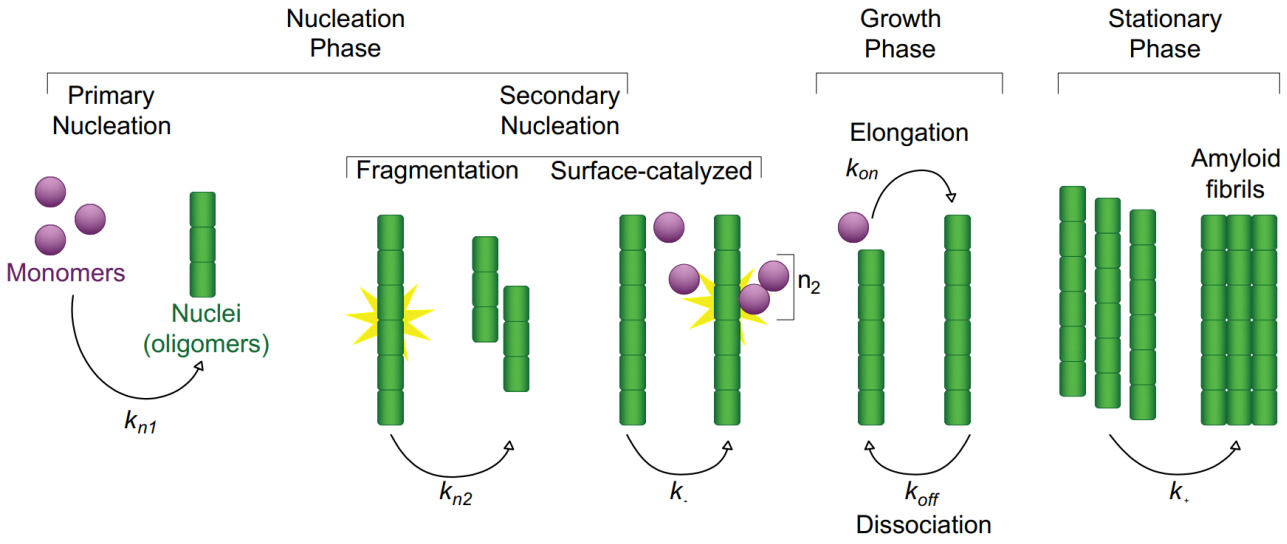


Figure 3 Schematic representation of the aggregation mechanism, where k_{n1}, k_{n2}, k_-, k_{on}, k_{off}, k_+ represent rate constants. Adapted from [9], [24].

The nucleation phase is followed by an elongation phase, also called growth, polymerization, or fibrillation phase. During this phase, oligomers continue to self-assemble into ordered fibrillar structure termed protofibrils. The last step is equilibrium (stationary or saturation phase). In this phase, protofibrils interact and self-associate into mature fibrils.

Furthermore, during the nucleation phase a primary and secondary nucleation can be distinguished (see Figure 3). The primary nucleation pathway can be homogenous or heterogenous. In the

former, monomers aggregate in bulk solution, whereas in the latter, the aggregation of monomers is catalyzed by the surface of an object [31], [32], other protein or peptide [33], or a lipid membrane [34]. In the secondary nucleation process, the aggregation process is seeded by the presence of pre-formed aggregates [35] or by fragmentation [36]. In the fragmentation process only fibrils are present, however, in secondary nucleation both monomers and fibrils are involved in the mechanism.

The second type of aggregation process is the nucleation-independent mechanism. During this process, the lag phase is missing, and there is no critical c_R(monomer in the equilibrium). Additionally, the rate constants are identical in each step [25]. In turn, in the nucleation-dependent process, the smaller aggregates (oligomers) elongate in a significantly slower phase than amyloid fibrils [37].

The gold standard method to follow the an aggregation process, and determine the mechanism it undergoes, is the thioflavin T (ThT) assay [38], in which, various modes of binding between the thioflavin fluorophore and the aggregates can be traced. The ThT dye can attach to the grooves in fibrils formed along β-sheets or to the monomeric or dimeric state of aggregates [39]. The quantity of fibrils formed during an aggregation reaction is measured as a function of time (note that the kinetics aggregation is highly dependent on initial protein or peptide concentration). The process of amyloid formation shows the increase of the total aggregate mass concentration, which is proportional to fluorescence. The results often display a sigmoidal kinetics reflecting processes related to three distinct phases: nucleation, elongation, and equilibrium (see Figure 4).

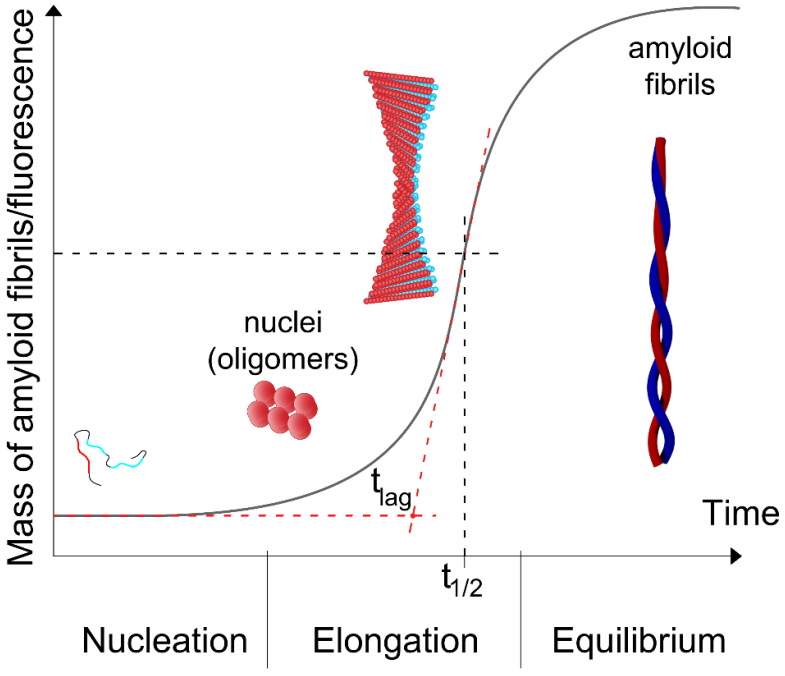


Figure 4 The sigmoidal curve of kinetics of the aggregation process. Based on [40].

By fitting the sigmoidal curve to registered data, various parameters of the kinetics of the aggregation process can be estimated: t_{lag} (time value based on the intersection of extension of the curve's linear part and the line tangent to the curve's beginning), $t_{1/2}$ the time needed to reach halfway through the elongation phase and k_{app} the growth rate [38] (see Figure 4). Thanks to such analysis, the kinetics under various experimental conditions can be compared.

1.4. Intrinsic and extrinsic factors affecting amyloids' aggregation

The initial concentration of peptides is one of the critical factors that influence the aggregation kinetics. However, in some cases, even starting from the same initial concentration, protein or peptide behaves differently in different samples. This is due to several other experimental conditions, such as: type of solvent, pH, presence of the lipid membrane, *etc*. This raises questions onto how various environmental or intrinsic factors influence the aggregation process and can one modulate the peptide tendency to aggregate.

1.4.1. Amino acids sequence

The propensity of the aggregation process depends on amino acids sequence. Generally, in proteins we can distinguish the primary structure, thanks to their specific amino acid order in polypeptide chain. Amino acids are linked by a covalent (peptide) bond involving two backbone atoms. Chiti et al. reported that hydrophobicity of the sequence contributes positively to the aggregation process as it, reduces the total net charge of the peptide/protein [41]. This intrinsic factor is also essential for understanding which amino acid sequence can particularly favor a peptide or protein to aggregation and how the order of amino acids influence the amyloidogenicity.

To help amylogenic fragments identification, various computational bioinformatics methods are being developed. These methods enable prediction of the so-called aggregation-prone regions (APR) [42], which are likely to trigger the aggregation process of a peptide or a whole protein. Based on the amino acid sequence, various physicochemical properties, such as: the total hydrophobicity, the charge state, the secondary structure propensity and the aggregation propensity can be calculated [43], [44]. The bioinformatics approach allow as well prediction of the effects of mutations on aggregation propensity [45]–[49]. While it has been found in some instances that a single mutation can favor protein aggregation [45], [46], it is important to identify not only the amino acids that catalyze aggregation processes but also those that prevent it. These amino acids are called gatekeeper residues [50], [51]. The well-known gatekeeper residues in the fragments of CsgA are, for example the aspartic acid (D) and the glycine (G) residues [15]. Another fact worth noticing is that functional amyloids, e.g. CsgA or the amyloid-like fimbriae protein, contain conserved residues in their repeats. In turn, those conserved residues are not characteristic of pathological amyloids. However, the order of amino acids in

the case of pathological amyloids, which are mostly intrinsically disordered proteins (IDPs), underlines a special feature. Contrary to Anfinsen's dogma, one sequence does not correspond to only one structure [52]. This property is called polymorphism (see section 1.1. Finally, the length of the polypeptide chain is another important factor: shorter sequences are easier to study, due to their low complexity, they do not present numerous steric collisions in the structure [53] and their synthesis is less troublesome.

1.4.2. Concentration

Concentration is the most important, yet somewhat disputable, factor in determining the ability of a peptide/protein to aggregate. This aspect directly affects the aggregation rate in experiments (see Section 1.3. The higher the initial concentration of the peptide, the shorter are t_{lag} and $t_{1/2}$ (Figure 4) [54]. Based on the characteristic features of nucleation-dependent polymerization, the c_R of each studied peptide/protein differs, and depends on the general solution conditions [55]. We recall that c_R corresponds to concentration of peptide/protein species required to form nuclei (see Figure 3), meaning that below c_R, peptide/protein species are in the monomeric forms. As an example, Wang et al. showed that R2 and R4 fragments of the CsgA protein (two out of its five fragments) do not fibrillate at all at a concentration of 0.2 mg/ml. However, fibrils were observed at a concentration of 2 mg/ml, in TEM micrograms [56].

1.4.3. Solvents

Amyloids are generally water insoluble [57]. To enhance their solubility, different media are used. The most popular approach to enhance amyloid solubility is the use of organic solvents helping. the peptides to remain in monomeric form and, more or less, unfolded [58]. Depending on the sequence of the peptide, various organic solvents are used, the most popular being 1,1,1,3,3,3-hexafluoro-2-isopropanol (HFIP) [38], dimethyl sulfoxide (DMSO) [59], 2,2,2-trifluoroacetic acid (TFA) [60], 2,2,2-Trifluoroethanol [61]. All of the aforementioned solvents vary greatly in their polarity. HFIP, the apolar solvent, is widely used in the monomerization of Aβ peptide, because it stabilizes the helical structures [62]. Note that such process may influence the peptide/protein's kinetics of aggregation [63], [64].

Frequently, deuterium oxide/heavy water (D_2O) is used as a solvent, especially in vibrational spectroscopy or nucleic magnetic resonance measurements, i.e. due to experimental requirements. D_2O causes a distortion in hydrogen bonding network, due to N-D bond lengthening [65], [66]. It is well known that in the case of larger proteins, such as tubulin, β-lactoglobulin A, actin, etc., heavy water stabilizes proteins, makes them more compact and less flexible than in H_2O, and increases the level of aggregation. However, how D_2O influences peptides it is not very well characterized.

1.4.4. pH and net charge

Factors such as pH and charge have a significant effect on the kinetics of the aggregation process. Peptides/proteins possess amphoteric properties, so they behave like acids or bases, depending on the pH of the solution. pH modulates the distribution of surface charges, which is crucial in protein folding, as well as in peptide-protein interactions. In extreme pH, peptides/proteins are heavily charged, and the denaturation process is enhanced. One of the approaches is to lower pH using HCl [67] or increase of pH/isoelectric point (NaOH, $NH_3 \cdot H_2O$) [68]. However, fibrillar forms may still be present, especially at high concentrations or in the case of highly hydrophobic peptides such as Aβ42. The fibrillization can be modulated by pH, for example, in acidic pH, hIAPP, and insulin are in monomeric forms [69], [70].

Isoelectric point (pI) is a characteristic pH value where peptide carries no net electrical charge or is electrically neutral in the statistical mean, so the number of positive and negative charges on the surface of the molecule balances with each other. In this case, the peptide/protein solubility is minimal. For amyloids, it is characteristic to possess a low net charge [61], however, it also depends on the pH of the solution. The lack of charged residues in peptides/proteins should limit interaction with ions.

1.4.5. Ions' contribution

To mimic more physiological conditions (control pH) in experimental setup, many distinct buffers are used in the peptide/protein study. It should be emphasized that various buffers have diverse buffering capacity [71]. The greatest buffering range is around their pK_a value of the acid (used to make this solution). Additionally, buffers can also differ in molarity and ionic strength. The ionic strength, related to the concentration and charge of ions in solution affects the solubility of peptides. Moreover, both positive (cation) and negative (anion) ions in solution can electrostatically interact with the peptide/protein surface. This may affect not only conformation, but also kinetics and morphology of formed fibrils [72].

Hence, the use of appropriate salts can stabilize or destabilize peptides/proteins and influence the aggregation process. The Hofmeister series ranks ions depending on their kosmotropic and chaotropic properties [73]. Kosmotrope ions (water structure maker) have a propensity to stabilize the native fold leading to salt-out effect of proteins. In contrast, chaotrope ions (water structure breaker) tend to facilitate protein denaturation and unfolding [74]. For instance, salts e.g. Na_2SO_4, Na_2HPO_4, $NaNO_3$, and $NaClO_4$ have an important impact on the rate constant of the aggregation reaction of human insulin and chicken egg white lysozymes. The ions' nature and concentration also control the formation of a large number of amyloid-like forms from nano to micro scales [72]. The type of ions

influences not only the aggregation process but also the ionic strength. For example, ionic strength has an impact on the lag phase and the growth of hIAPP amyloid formation [75].

Mikalauskaite et al. have shown that the ionic strength has a significant impact on both the binding affinity to amyloids' fibrils and fluorescence intensity of ThT [76]. ThT is a widely used amyloid-specific dye in the study of the kinetics of aggregation processes (see 1.5.5.). Klement et al. showed that cationic series such as $MgCl_2$ promote mainly the aggregation process of Aβ40 [77]. In the presence of salts, fibrils have smaller diameter and narrower crossover distance. Note that the molarity, i.e. the ions activity/influence and impact depend on the temperature of the samples. In the following section we discuss more thoroughly the influence of the latter.

1.4.6. Temperature

Temperature is the most important physical/environmental factor that alters the behavior of particles in solution. Change in temperature influences the vibrational motions and the overall diffusion of peptides/protein particles [78]. As a result, a decrease or increase in temperature can destabilize peptides/proteins and trigger an aggregation process. Generally, increasing the temperature leads to faster aggregation due to creation of additional hydrophobic interactions [79] and increased diffusion of peptides/molecules [80]. It also influences the aggregation kinetics [81]. The temperature at which the peptide/protein undergoes secondary structure changes depends on the thermostability of the studied specimen. Generally, the aggregation process in amyloids is described as non-Arrhenius behavior [82]. As an example, insulin, which at 37°C exists in its native conformation (α/β structure under ambient conditions), forms a cross-β sheet structure with increasing temperature ($\geqslant 37$), but above ~140°C transforms to a random coil structure. Random coil formation was found to be partially reversible. *O*w et al. showed that the hen egg white lysozyme (HEWL) aggregates into fibrils at approximately 65°C, and above this temperature, more amorphous aggregates have been observed [83]. Moreover, prion proteins which are infectious amyloids associated with mad cow disease, scrapie, and others, have higher thermal stability. For example, PrP protein possesses a native structure at 100°C [84]. In contrast, increasing the temperature from 37°C to 39°C, promotes aggregation of Aβ40 and Aβ42 (a higher level of Aβ in the group of examined patients [85]).

Presumably, temperature is one of the means of controlling the aggregation process, based on the regulatory mechanism of fungal, bacterial, or functional amyloids.

1.4.7. Presence of another peptide or protein

The kinetics of the aggregation is influenced by physical, chemical, and also physiological factors, such as interaction with other peptides/proteins. In the human body, various cellular compartments coexist and interact with each other to support various biological machinery. Researchers discovered

common denominators between different diseases, such as: Parkinson's, Alzheimer's, or even 2 type diabetes [86]. Some disorders presumably result from interactions between different amyloid proteins, a process or phenomenon called cross-seeding (or cross-talk) [87], [88].

It is well established from *in vitro* and *in vivo* studies that one amyloid protein can trigger the aggregation of another one. The phenomenon of cross-talk appears to be possible mainly between proteins with a high percentage of identical residues (identity) and a high percentage of residues conserved with similar physicochemical properties (similarity). In some cases, cross-amyloid interactions may lead to inhibition of aggregation [89], [90]. Such inhibition may be useful to exploit in potential therapies. The cross-talk between amyloid proteins involves cross-sequence interaction between two various samples (monomers and peptides in the various state). These different scenarios explain possible pathways of interactions between amyloid proteins.

Many parameters, such as: buffer concentration, pH, temperature, *etc*., may influence the process. Different cross-interactions involving interaction between Aβ and α-synuclein [91], Aβ and tau [92], Aβ and hIAPP [93] were reported in the literature. Various interactions, collected in the AmyloGraph database [94], are displayed in Figure 5.

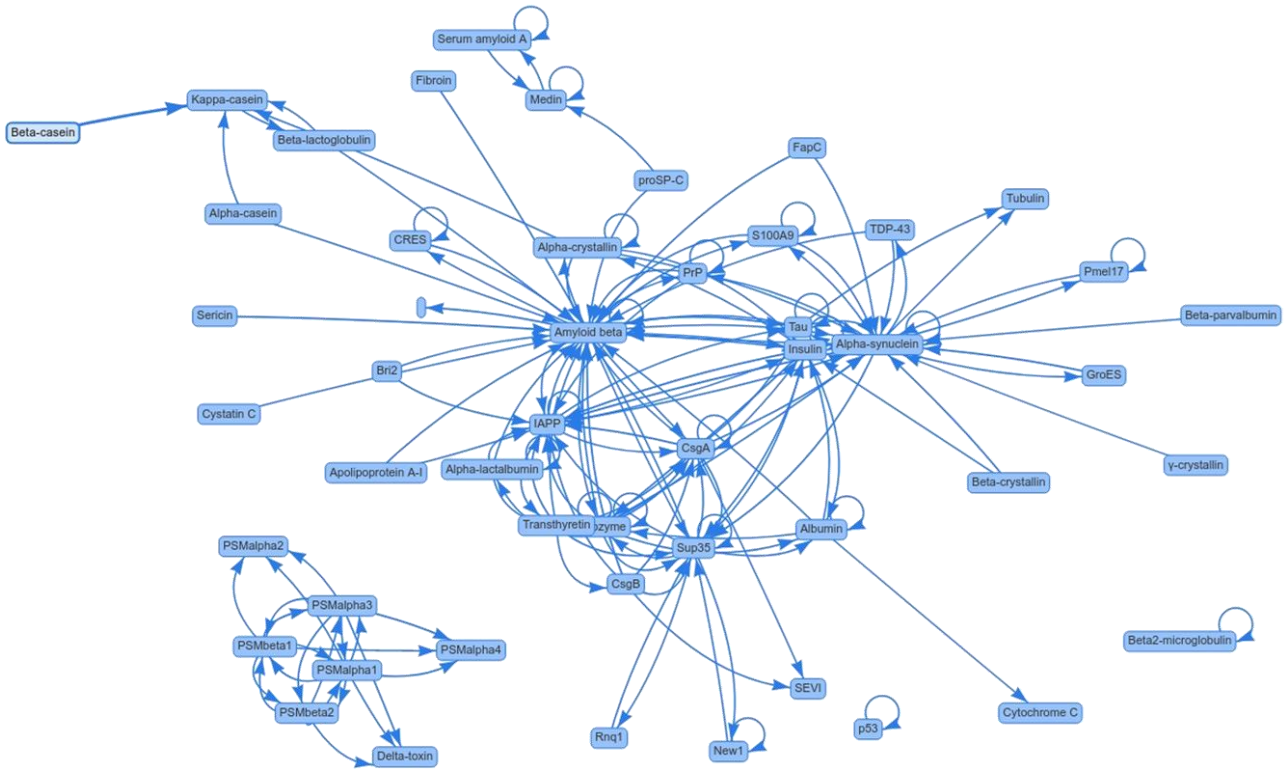


Figure 5 Experimentally investigated amyloid cross-interactions. The number of arrows correlates with a number of publications (also regarding different motifs). Based on AmyloGraph, comprehensive database of interactions between amyloid proteins.

Figure 5 shows that the most frequently studied amyloid protein is Aβ, which is linked to AD [95]. This protein is the main constituent of amyloid fibrils. Aβ peptide is a 40 or 42 amino acid long sequence generated by proteolysis of a precursor within the transmembrane domain (TMD). Formation of Aβ peptide, via the amyloidogenic pathway, requires two cleavages of the amyloid-β precursor protein (APP), caused by β-secretase (BACE1) and γ-secretase [96]. The underlying causes and mechanisms leading to the development and progression of AD remain still elusive. Deposits of Aβ peptides are mainly observed in the region of hippocampus and neocortex, as well as in the cerebral-vasculature. It has been often shown that aggregation process is accelerated due to the presence of biological membrane [97]. However, the exact mechanisms of interactions between the amyloid peptide and cell membrane *in vivo* are far from being fully understood.

Nowadays, the interaction between functional and pathological amyloids is of special interest. It is hypothesized that microbiome amyloids could affect the progression of neurodegenerative diseases by influencing the rate of aggregation of pathological amyloidogenic proteins [98], the effect called the brain-gut axis [99]. Discovering the interactions between amyloids is important for better understanding of etiology and mechanism of amyloidosis.

1.4.8. Protein-membrane interaction

It is important to study the aggregation process not only in bulk solution but also in the presence of more complex systems, such as phospholipid bilayer. Interactions between peptides and cell or organelle's membrane are different from their aggregation behaviors in solution. The roles played by lipid-peptide/protein interaction in the aggregation process can be various. Specific amyloid peptides can adsorb on or insert into and destabilize membranes, due to direct interactions between them [34][100]. On the other hand, oligomers can interact directly with membranes [101], disrupting the integrity, permeability, and functions, leading for instance to ionic homeostasis, changing signaling pathways, causing oxidative injury, and, ultimately, the cell death. First, the lipid membrane can act as an interface, that triggers the aggregation process. Besides, the interaction with amyloidogenic peptides/proteins can disrupt the cell membrane via direct interaction [102]. Additionally, based on the amyloid cascade hypothesis for Aβ peptide [103], we can assume that toxic aggregates are formed by assembly of the Aβ peptides that left the membrane, after the γ-secretase cleavage of APP (see Figure 6). However, the mechanism of the amyloid cascade hypothesis remains elusive and brings in the question of whether the emergent Aβ peptides remain associated within the membrane, why and how it would leave the membrane toward intra or extracellular median.

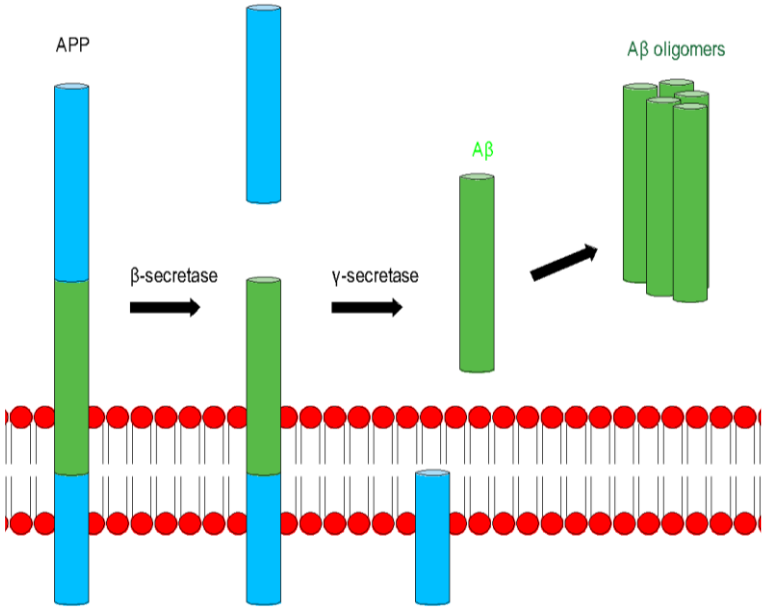


Figure 6 Schematic representation of the classic proteolytic process of APP, the Aβ precursor, leading to the amyloid cascade hypothesis.

1.4.9. Other factors

Many other factors, not described above, can influence the aggregation. One of them is the peptide/protein source (if it is synthetic or recombinant). In the case of synthetic peptides, the problem starts in solid-phase synthesis due to the high presence of hydrophobic residues. Synthesis of sequences with more than 20 amino acids, for example Aβ42, is more challenging because these peptides are more prone to aggregation [104]. Aggregation can also be influenced by various impurities from the synthesis, purification, or monomerization processes [38]. Other factors include pressure, agitation, stirring speed [105], storage conditions, irradiation, interaction with the test tube [78], [106][107]. Additionally, another important aspect that affects the local stability of peptides and proteins is chemical degradation. It mainly includes processes like deamidation of amide groups (asparagine (N), aspartic acid (D), glutamic acid (E), and glutamine (Q)) or oxidation of amino acid side chains (mainly aromatic amino acids). Furthermore, even chemical degradation such as hydrolysis [108], racemization [107], or gelation [109] can occur.

All the above factors affect the reproducibility of the experiments, so a detailed description is essential and an integral part of the conducted examination. All the extrinsic factors can also lead to various structural conformations of amyloid fibrils [110].

1.5. Experimental methods in amyloid studies

To understand and follow the aggregation process of peptides/proteins, multiple experimental techniques can be implemented. According to the MIRRAGGE – Minimum Information Required for Reproducible AGGregation Experiments [27] protocol, to obtain reliable information about amyloid

assemblies, a good practice is to combine techniques [111]. It is also bound with the fact that each of the methods provides various information (kinetics, secondary structure, morphology). Additionally, obstacles in studying amyloid peptides/proteins (insolubility, rapid aggregation) and other limiting aspects, such as extrinsic factors (see section 1.4.), effect in different results from various methods. Furthermore, experimental methods differ in requirements, such as form of the sample. In the following, I will describe in a nutshell the most popular methods, with a focus on the techniques implemented and used in this thesis.

1.5.1. Circular dichroism spectroscopy

Circular dichroism (CD) spectroscopy is an analytical method, used to obtain information about the secondary structure of proteins [112]. It is a quick and cheap method for the primary determination of whether the aggregation process occurs [113]. The negative and positive minima of the curve provide information about secondary structure changes (see Table 1).

Table 1 The assignment of the curve's minima and maxima to secondary peptide/protein structures.

Wavenumber [nm]	Assignment of the structure
broad negative minimum at **215–220** nm and positive at about **196** nm	β-sheet
positive peak at around **212** nm and a negative minimum at about **195** nm	random coil
two negative bands at **222** nm and **208** nm	α-helical

As a result, it allows one to observe the transition from α/random coil to β-sheet structure over time, which is characteristic for amyloids. The CD spectroscopy results obtained for the peptides and proteins containing β- or α/β- structures always need to be carefully interpreted and the general conclusions about the given protein/peptide conformation should be drawn with the support of other characterization techniques.

Technique requirements: solution, peptide/protein concentration 0.1–2 mg/mL.

1.5.2. Infrared spectroscopy

1.5.2.1. Attenuated Total Reflection-Fourier Transform Infrared

Infrared (IR) spectroscopy is an alternative method to CD spectroscopy for protein secondary structure analysis. This vibrational technique is based on the absorption of infrared radiation. The most useful technique, from the perspective of the biological sample measurement, is ATR-FTIR [111], [114]. It is one of the most commonly used tools in protein analysis [115], including that of amyloids [116]. Application of the ATR-FTIR technique enables the measurement of both dry and fully hydrated samples. In contrast to the transmission FTIR technique, in ATR the signal of water is weakened due

to the reflection law. As a result, it solves the problem of high-water's absorbance band at approximately 1634 cm^{-1}, which overlaps with Amide I band, crucial for secondary structure analysis. Usually, the sample is placed on the internal refractive element (IRE) crystal surface in the form of a hydrated solid film/air-dried film. IRE is made from a material with a high refractive index (for example diamond). Unfortunately, the aggregation propensity of the studied peptide/protein can be accelerated by additional interactions between proteins and the hydrophobic diamond surface [117].

The Amide I is the most intensive band in protein spectra. Observed in the spectral range of 1700–1600 cm^{-1}, it mainly corresponds to the stretching vibrations of the C=O groups of amide bonds. Vibrational frequencies of amide bonds depend on the strength of hydrogen bonding interactions, backbone conformation, hydration level, as well as the length of the peptide chain. Based on the position of the maximum of Amide I, the secondary structure of peptide can be distinguished, according to the assignments presented in Table 2. The presence of the characteristic narrow and relatively intense absorption band in the IR spectra allows to discriminate aggregated peptides/proteins from non-aggregated. The native β-sheet rich peptides/protein possess local maximum at about 1635 cm^{-1} [118], while for the mature amyloid fibrils/intermolecular aggregates, this maximum is shifted to about 1627–1610 cm^{-1} [115]. Moran et al. showed that the shift in the band to lower wavenumbers indicates more rigid and ordered structure [116]. Furthermore, ATR-FTIR enables one to distinguish oligomers from fibrils in the amide band analysis. Oligomers with anti-parallel β-sheet structure usually possess two characteristic local maxima, one at about 1695 cm^{-1} with lower intensity and the second main at about 1630 cm^{-1} [118]. To estimate the participation of each of components in secondary structure, the deconvolution of the Amide I band has to be made, based on second derivative minima.

Table 2 The assignment of Amide I bands components to secondary structures of proteins in water and heavy water. Based on [118]–[121].

Band position Wavenumber [cm^{-1}]		**Assignment of the structure**
H_2O	**D_2O**	
1696–1683	1690–1680	anti-parallel β-sheet
1685–1660	1675–1660	β-turns
1670–1660	1667–1660	parallel β-sheet
1658–1647	1655–1648	α-helix
1652–1643	1648–1640	random/extended, loops
1640–1630	1637–1625	β-helix
1638–1627	1635–1615	β-sheet
1627–1610	1620–1600	intermolecular aggregates

One of the advantages of the method is the ability to reveal the presence of different β-structures, which allows to differentiate between structurally distinct types of amyloid fibrils *in vitro* [122], [123]. This fact is of particular importance in the study of amyloid polymorphism.

Technique requirements: hydrated solid film, peptide/protein concentration: 1 mg/mL or less, depending on spectrometer configuration and measurement parameters.

1.5.2.2. Infrared Microscopy

IR microscopy (μIR) combines FTIR with microscopy and allows sample mapping and imaging in ATR, reflection, and transmission modes. In transmission mode, in μIR, infrared radiation penetrates the sample and is absorbed by it. In this method, the sample solution (about 10 μL) is applied on the surface of the window made from IR transparent material *e.g.,* barium fluoride (BaF_2) and allowed to dry out. As a result, a coffee-ring [53] is formed. When peptides/proteins assemble, particles distribute linearly, according to their mass, at the air-liquid interface. Based on this, the particles with the highest mass are placed on the edge of the drop (see Figure 7). It is important to form a thin layer, otherwise the thicker samples will not be visualized using transmission mode [124].

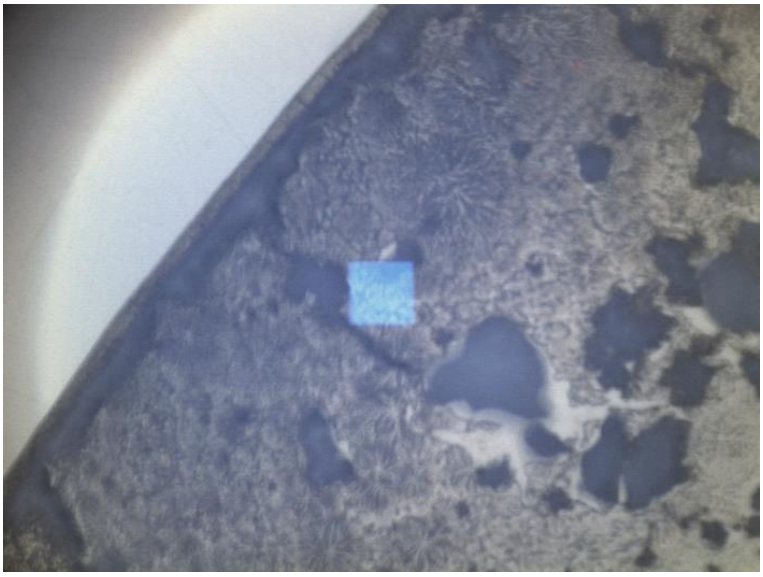


Figure 7 Exemplary image of the edge of the coffee-ring formed by the peptide SFLIF.

IR microscopy allows for mapping the probe with a step of 10 μm or even less with fast acquisition time and good signal to noise ratio (SNR), due to utilization of sensitive detectors (for example, germanium nitrogen cooled). In addition, implementation of a linear array or a focal plane array enables: fast scanning of large areas and choosing a region of interest (ROI), which is significant for non-homogeneous deposition patterns. The analyses of the spectra are the same as described in section 1.5.2.1.

Technique requirements: hydrated solid film on the infrared transparent material, peptide/protein concentration about 0.5 mg/mL.

1.5.3. Fourier Transform-Raman spectroscopy

Raman spectroscopy is a method complementary to infrared (IR) spectroscopy. In contrast to IR spectroscopy, this technique is based on inelastic scattering of electromagnetic radiation. In case of Fourier Transform-Raman spectroscopy (FT-Raman), infrared lasers, typically neodymium-doped yttrium aluminum garnet (Nd:YAG) are used as the excitation source. FT-Raman also allows secondary structure analysis and examination of the aggregation process [125], [126]. However, it is not frequently used for studying amyloids, due to long acquisition time and possible excessive heating of the sample through laser radiation. To enhance weak Raman signal via surface-enhanced Raman spectroscopy effect, gold coated substrates can be used [127], [128].

In the FT-Raman spectra of proteins, the Amide I (1725–1575 cm^{-1}) and Amide III (1375–1185 cm^{-1}) are the most conclusive for the interpretation of the components of the secondary structure, due to their highest intensity. Interestingly, secondary structure components are better separated in the Amide III spectrum. As a result, interpretation of two amide bands enables increasing the certainty of structure assignment (see Table 3). The spectral characteristic of aggregates can be distinguished by narrower Amide I band with higher intensity at 1670 cm^{-1}, and additionally, the higher intensity can be observed at about 1235 cm^{-1}. Additionally, it differentiates the exposure of aromatic amino acid residues to an aqueous environment [129].

Table 3 General assignment of amide bands to secondary structures in FT-Raman spectrum. Based on [130], [131].

Band position Wavenumber [cm^{-1}]		**Assignment of the structure**
Amide I	**Amide III**	
1659–1655	1310–1280	α-helix
~1680	1270–1250	unordered
1695–1680	1240–1210 1270–1280	β-turn
~1670	1240–1210	β-sheet

Technique requirements: hydrated solid film on the gold slide, peptide/protein concentration about 1 mg/mL.

1.5.4. Chemometric analysis

Nowadays, the analysis of spectral data is combined with chemometric tools [132], *e.g.* principal component analysis (PCA) [133], [134]. PCA is a dimensionality reduction technique, which uses

linear combination of the original variables to produce Principal Components (PC), which capture most of the variability in the data. This method enables visualization of high dimensional data such as spectra, using the first two or three PCs and provides insight into data set structure. Such analysis allows fast and more robust identification of amyloids, as well as distinguishing between homo and hetero aggregates. Furthermore, it enables detection of regions of spectra that differentiate different classes, by examining wavelengths that vary the most in the dataset. The use of chemometric techniques accelerates the analyses and makes it possible to find additional relationships.

Technique requirements: spectra as input data.

1.5.5. Thioflavin-T

The benzothiazole derivative thioflavin-T is a widely used amyloid staining fluorescent dye. It has a wide variety of applications, from histology to *in vitro* studies [38], [135]. When ThT binds to amyloid fibrils, its excitation maximum shifts from 385 to 450 nm and the emission maximum from 445 to 482 nm. The fairly good selectivity of this process and the intensity of the fluorescence increase, make ThT an excellent amyloid sensor. The enhancement of fluorescence depends on fibril structure, which was proved by [136]. This technique is used to follow the kinetics of the aggregation process in time (see 1.3.).

Technique requirements: monomerized solution, peptide/protein concentration about 0.5 mg/mL or even lower, depends on the peptide/protein.

1.5.6. Electron microscopy

Electron microscopy (EM) is the first microscopic technique, that enables observation of proteins and amyloids morphology [137]. It is based on the elastic scattering of the electron beam with a studied sample [138]. To enhance contrasts on EM images, samples are stained with uranyl acetate (called negative staining). There are various types of EM, here I will focus on transmission electron microscopy (TEM). In this mode, the electron beam penetrates through the sample on a metal mesh grid. As the measurement is in the high vacuum, the specimen has to be properly prepared. TEM provides important information about the basic dimensions of amyloid fibrils (length and width).

Technique requirements: solution applied on the metal mesh grid, staining with uranyl acetate, peptide/protein concentration about 0.5 mg/mL or even lower depending on the peptide/protein.

1.5.7. Atomic force microscopy

AFM is a high-resolution method, which enables acquisition of topographic images of the sample area in real time. It is based on the interaction of the nanomechanical tip with the scanned surface. This method is widely used in amyloid studies [139], due to its high resolution. The vertical resolution

of AFM is about 1.5 nm and horizontal about 0.1 nm [140]. Peptides/proteins are most frequently studied using the tapping™ mode (TM), which means intermittent contact of the tip with the scanned surface. In this method, various measurement modes, such as: in the air on dried samples, in liquid or even in the lipid environments, are possible. The main advantage of this technique is the possibility to characterize interaction between peptide/protein and a lipid membrane [141], not only the peptide/protein morphology [142]. Additionally, the oligomers can be easily distinguished from mature fibrils. The analyses of topographic images take into account the registered cross-section images, based on which we can measure the registered specimens' height.

Technique requirements: solution, peptide/protein concentration about 0.1 mg/mL or even lower.

1.5.8. Scanning Kelvin Probe

Kelvin probe is a non-contact, capacitively coupled voltage measuring technique capable of sensing potential difference between the probe conductive tip and the examined material surface. The probe is often used as an attachment to AFM, called Kelvin probe force microscopy (KPFM), also known as scanning surface potential microscope. In this case, the electrostatic force is sensed at a nanometric distance scale between the Kelvin probe tip (affixed to the AFM cantilever) and the studied surface. Because of the atomic-scale tip-object interaction, the energy bands tend to be locally distorted. As a result, this fact influences the reliability of the registered values of contact potential difference (CPD) and work function (WF) [143]. CPD is commonly measured using a modulation technique proposed already in 1932 by Zisman [144], combined with nulling of the electric field in the tip-object gap by an external bias voltage (Vb), which is equal to CPD when nulling condition occurs. Off-null detection technique, which was proposed by Baikie et al. [145], provides additional measures to increase sensitivity and keep the tip-surface separation constant, furnishing CPD values with single mV resolution. However, irrespectively on measurement approach, if the tip-surface CPD is gauged and the WF of the tip material (WF_{tip}) is known, it is possible to calculate WF of the analyzed specimen material (WF_{spec}), local to its region, beneath the probe tip, employing a simple formula $WF_{spec} - WF_{tip} = eVb$ (where e is the electron elementary charge). As the tip may be reallocated over the analyzed object surface (while keeping the tip-object gap constant).

KPFM has been employed in amyloid studies to map CPD distribution in single peptides and amyloid fibrils [146], [147], showing the surface charge distributions. However, it has not yet been implemented to calculate WF of macroscopic peptides' deposits and to correlate WF values to structural differences of the examined peptides.

This is a novel approach in characterization of amyloid fibrils and order of the aggregates (highly structured or not). The Kelvin probe can provide quantitative measures of the molecular

characteristics of amyloid aggregates. CPD results from a difference in the WF of the probe tip material and the investigated material. The method is very sensitive not only to changes in the surface chemical composition and its contamination, but also to real charges located at the surface or (in case of dielectric specimens) in the bulk of examined material. SKP can deliver profiles and 2D maps of CPD and WF distribution with spatial resolution comparable to the tip dimensions.

Technique requirements: hydrated solid film on the gold slide, peptide/protein concentration about 1 mg/mL.

Table 4 sums up all advantages, disadvantages and specifies the application of each of the methods.

Table 4 Glossary of techniques described above for amyloid studies. Based on [111].

Method	Applications	Strengths/limitations
Staining		
Thioflavin T	histochemical examinations *in vitro/in vivo* experiments kinetic studies	+ time-efficient + sensitive - nonspecific - affected by pH, ions, dye concentration - problem with reproducibility of results - demands monomerized sample - autofluorescence of amyloids
Analytical methods		
Fourier transform infrared spectroscopy	secondary structure conformational changes	+ real-time monitoring + time-efficient + no sample preparation + highly reproducible results - one average single spectra obtained from a small area (ATR-FTIR) - high water absorption - hydrophobic diamond surface can accelerate aggregation
Circular dichroism (CD) spectroscopy	secondary structure conformational changes	+ real-time monitoring + time-efficient + low sample concentration - homogeneous solution
Raman	secondary structure conformational changes	+ aromatic amino acids orientation and exposure to local environment + information about fibers' orientation in the case of polarization measurement - long acquisition time - lack of amyloid-specific band - low signal-to-noise ratio - sample heating through the laser radiation
High resolution methods		

Atomic force microscopy (AFM)	morphology nanomechanical properties	+ minimal sample preparations + measurements in physiological conditions + different conditions + mechanical properties - peptide immobilization - artifacts - time-consuming images acquisition
Electron & cryo-microscopy (EM & cryo-EM)	morphology three-dimensional (3D)	+ fibrils visualization + different aggregates + no need for crystals - difficult sample preparations - no real-time imaging - oligomers are difficult to visualization - time-consuming - costly
Others		
Scanning Kelvin Probe (contact potential difference (CPD) and work function (WF))	as a correlation method to other techniques	+ minimal sample preparations + time-efficient - no buffer, due to crystallization

To fully understand the aggregation of peptides/proteins, it is recommended to follow the time evolution of its process. Experimental studies can be supported by bioinformatics and computational methods, due to their time-consuming nature and ambiguity of experimental results. Additionally, it is difficult to perform experiments on large sample sets.

1.6. Computational methods in amyloid studies

1.6.1. Bioinformatics predictors

There is a wide range of various amyloid sequence predictors available [148]. They are applied to estimate the ability to form amyloid aggregates considering the amino acid sequence. Predictors are based on various approaches (see Figure 8).

The most popular methods are founded on the physicochemical properties of amino acids in the sequence. The other ones rely on structure-based and statistics-based models. The latter approach includes machine learning techniques, which are of great interest to researchers. Most of these methods analyze protein sequences using a sliding window of six amino acid length. This number is historically linked to the development of the amyloid databases. Nevertheless, researchers are not only interested in studying short fragments, but also in examining complete proteins. Most amyloidogenic predictors combine physiochemical and statistical approaches, for example: Pasta 2.0 [149], FoldAMyloid [150] or Aggrescan 3D [151].

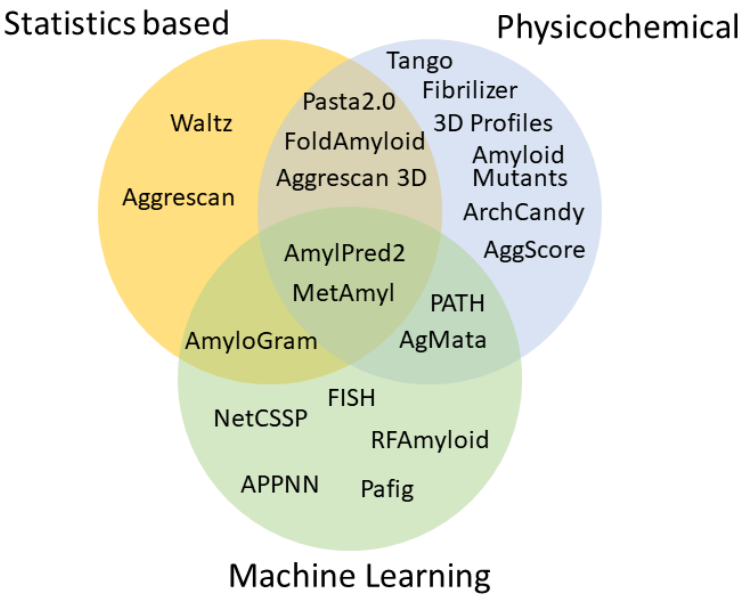


Figure 8 Bioinformatic approaches and methods for amyloid predictions. Adapted from [148].

1.6.2. Molecular dynamics simulations

Molecular dynamics simulations (MD) allows insight into dynamics of biological macromolecules at the atomistic and mesoscopic scale which is not readily accessible through experimental measurements. It allows us to explore peptide/protein conformational changes, interactions with another peptide/protein, phospholipid bilayers, ions, and other compounds *etc*., and ultimately the peptide/protein self-assembly process. In this paragraph, I will focus on the all-atom conventional MD. Based on solving Newton's laws of motion equation, the interactions and motions of atoms can be studied. Integration of Newton's equation by consecutive MD steps provides the position and velocity of atoms in every time frame. Gathering subsequent particle positions forms a system trajectory that allows for further calculation or analysis [152].

The main disadvantage of this method is that amyloidogenic peptides/proteins have complicated free energy landscape (see Figure 2) and tend to be trapped in one of the local minimum states of free energy landscape in MD. To overcome this limitation, other algorithms have been used to study the aggregation process, for example: replica-exchange [153], simulated tempering [154]. These methods allow easier crossing of the energy barrier and studying the broader conformational space.

Another issue is setting proper force field parameters (the forces between atoms, molecules, and total energy of the system) for amyloid proteins. The CHARMM36m force field is parametrized particularly to study IDPs, due to refined backbone peptide/protein parameters and the tendency towards the creation of ordered structures [155]. Depending on which force field is used, various conformations for Aβ in water can be obtained [156].

The next challenge is time scale [157]: the aggregation process of some peptide/proteins experimentally takes hours or even days. In turn, MD simulations can generate at most only a fraction of a second of trajectories, which are not enough to observe a potential conformational change from native to fibrillar state of peptide/protein.

Despite all current challenges in studying amyloid proteins and peptides by MD, this technique allows characterization of the phenomena at the molecular level, that is likely to provide key information when combined with experimental studies.

1.7. Introduction summary

All of the above factors influence the reliability of the reference data and can negatively affect the accuracy and the learning process of the computational methods. The efficiency of such methods, especially based on supervised machine learning, is highly dependent on an accurate annotation of the reference training data obtained from *in vitro* experiments. It should be emphasized that experiments are conducted on proteins under different environmental conditions. This raises the question, whether the fact that protein creates an aggregate structure, for example at acidic conditions, is enough to classify it as an amyloid for the purpose of a database of bioinformatics methods. Unambiguous data from the various experimental methods are a source of biased bioinformatics predictions.

This dissertation is focused on selected internal and external factors which may affect the aggregation process of amyloidogenic peptides. Additionally, it raises potential problems of the ambiguity of results from experimental techniques in amyloid studies. I endeavored to study the effectiveness of the protocols and the dependence of the results on the experimental conditions.

The research was conducted with a focus on the following:

- Solvent and influence of ions: presence, type, and concentration
- Sequence determinants (mutations)
- The lipid membrane environment
- Interaction with another peptide

The studies were carried out considering:

- Various variants of hexapeptides from well-known proteins, which were the base data for very many bioinformatics predictors
- Fragments of the CsgA functional amyloid protein originating from *E. coli* and *S. enterica* bacteria
- Mutants of R4 fragments of the CsgA protein from *E. coli* and *S. enterica* bacteria
- Amyloid-β 42 (Aβ42), and human islet amyloid polypeptide (hIAPP) peptides originating from human species.

Research hypotheses

Amyloids have specific segments called amyloidogenic regions or hot spots, that trigger the protein's aggregation. These regions, can be predicted using various bioinformatics methods. However, the latter rely heavily on experimentally confirmed reference data, which can be ambiguous due to differences in various factors, such as protein sequence properties, external factors, and biological surfaces. Thus, the following research hypotheses were formulated based on the objectives presented.

The work carried out within this thesis aims at addressing the following research hypotheses:

1. Minor deviations in experimental conditions strongly influence the validity of a protein classification as an amyloid.
2. Aggregation propensity of closely related homologous functional amyloids may be different.
3. The presence of a lipid membrane affects aggregation of the native Aβ42 peptide.
4. The presence of another peptide/protein impacts the general fibrilization process.

Materials and Methods

1.8. Peptides

1.8.1. Hexapeptides

Hexapeptide sequences were extensively studied experimentally [158] and some were shown to form amyloid-like fibrils [159] with characteristic steric zippers. Such short peptides are easier to synthesize and are essential for bioinformatic predictors as input data. Among these, sequences reported in Table 5 are taken from the well-known Waltz-DB/2.0 amyloid database [50], [160], which contains hexapeptides tested for their amyloidogenic properties. All 34 sequences were studied in this work.

Table 5 List of hexapeptides studied in this thesis. In the parentheses, the proteins' accession number in the UniProt database.

No.	Sequence	Protein source	Database
1	FNPQGG	Eukaryotic peptide chain release factor GTP-binding subunit (P05453)	WALTZ-DB 2.0
2	FTFIQF	NA-binding protein RO60 (P10155)	WALTZ-DB 2.0
3	ISFLIF	Major prion protein (P04156)	WALTZ-DB 2.0
4	KPAESD	Eukaryotic peptide chain release factor GTP-binding subunit (P05453)	WALTZ-DB 2.0
5	LVFYQQ	Centrosome-associated protein ALMS1 (Q8TCU4)	WALTZ-DB 2.0
6	NPQGGY	Eukaryotic peptide chain release factor GTP-binding subunit (P05453)	WALTZ-DB 2.0
7	SFLIFL	Major prion protein (P04156)	WALTZ-DB 2.0
8	TKPAES	Eukaryotic peptide chain release factor GTP-binding subunit (P05453)	WALTZ-DB 2.0
9	YLLYYT	Beta-2-microglobulin (P61769)	WALTZ-DB 2.0
10	YTVIIE	unknown	WALTZ-DB 2.0
11	ALEEYT	Apolipoprotein A-I (P02647)	WALTZ-DB 2.0
12	ASSSNY	Ribonuclease pancreatic (P61823)	WALTZ-DB 2.0
13	DETVIV	Flocculation protein FLO1 (P32768)	WALTZ-DB 2.0
14	ELNIYQ	Major curli subunit (P28307)	WALTZ-DB 2.0
15	FGELFE	unknown	WALTZ-DB 2.0
16	FQKQQK	Eukaryotic peptide chain release factor GTP-binding subunit (P05453)	WALTZ-DB 2.0
17	FTPTEK	Beta-2-microglobulin (P61769)	WALTZ-DB 2.0
18	HGFNQQ	CAAX prenyl protease 1 homolog (O75844)	WALTZ-DB 2.0
19	HLFNLT	4-galactosyl-N-acetylglucosaminide 3-alpha-L-fucosyltransferase 9	WALTZ-DB 2.0

		(Q9Y231)	
20	HSSNNF	unknown	WALTZ-DB 2.0
21	MIENIQ	unknown	WALTZ-DB 2.0
22	MIHFGN	Major prion protein (P04925)	WALTZ-DB 2.0
23	MMHFGN	Major prion protein (P04273)	WALTZ-DB 2.0
24	NIFNIT	unknown	WALTZ-DB 2.0
25	NNSGPN	Major curli subunit (P28307)	WALTZ-DB 2.0
26	NTIFVQ	RNA-binding protein FUS (P35637)	WALTZ-DB 2.0
27	QANKHI	unknown	WALTZ-DB 2.0
28	QEMRHF	unknown	WALTZ-DB 2.0
29	SHVIIE	unknown	WALTZ-DB 2.0
30	STTIIE	unknown	WALTZ-DB 2.0
31	STVVIE	unknown	WALTZ-DB 2.0
32	SWVIIE	unknown	WALTZ-DB 2.0
33	WSFYLL	Beta-2-microglobulin (P61769)	WALTZ-DB 2.0
34	YYTEFT	Beta-2-microglobulin (P61769)	WALTZ-DB 2.0

1.8.2. CsgA fragments

The subsequent step was to study longer amino acid sequences, that mimic more physiological situations. CsgA protein is functional amyloid [161] and main component of biofilm in enteric bacteria, such as *Escherichia coli* (*E. coli*) and *Salmonella enterica* (*S. enterica*) spp. CsgA is highly hydrophobic, and it consists of 151 amino acids. Its synthesis and purification is technically very difficult. Additionally, its structure is not resolved, there are only predicted models [162]. Accordingly, we did not study the full-length protein at once, but five imperfect fragments from either *S. enterica* or *E. coli* (see Table 2). These fragments (R1–R5) build up the whole protein.

Table 6 Sequence and source information of CsgA fragments

Name	Sequence	Source organism
R1	STLSIYQYGSANAALALQSDARK	*S. enterica*
R2	SETTITQSGYGNGADVGQGADN	
R3	STIELTQNGFRNNATIDQWNAKN	
SR4	SDITVGQYGGNNAALVNQTASD	
R5	SSVMVRQVGFGNNATANQY	
R1	SELNIYQYGGGNSALALQTDARN	*E. coli*
R2	SDLTITQHGGGNGADVGQGSDD	
R3	SSIDLTQRGFGNSATLDQWNGKN	

ER4	SEMTVKQFGGGNGAAVDQTASN	
R5	SSVNVTQVGFGNNATAHQY	

1.8.3. Mutants of R4 fragments of CsgA

We have considered mutant peptides of the R4 fragments from of CsgA from *S. enterica* and *E. coli*. This was done in consideration of previous findings [163], which revealed differences in the aggregation tendency of the latter. The wild type R4 fragment of *E. coli* (ER4) in CsgA protein P28307 (UniProt accession code), is located at 91–112 amino acid positions, and the R4 fragment in *S. enterica* protein (SR4), A0A5X9F0D5 (UniProt accession code) at 91–112. In the following, the amino acid positions will be reported as in the peptides i.e., 1–22. In this study, a set of mutants was designed and synthesized based on the SR4 peptide: M4 (G6K), M2 (L15A), M3 (G6K, N17D, D22N), and the ER4 peptide: M1 (D11N, N22D), see Table 7.

Table 7 Sequence and source information of CsgA fragments (the introduced mutations are bolded and underlined)

Name	Sequence	Source organism	Mutation
ER4	SEMTVKQFGGGNGAAVDQTASN	*E. coli*	WT
M1	SEMTVKQFGGGNGAAV**N**QTAS**D**	*E. coli*	D17N N22D
SR4	SDITVGQYGGNNAALVNQTASD	*S. enterica*	WT
M2	SDITVGQYGGNNAA**A**VNQTASD	*S. enterica*	L15A
M3	SDITV**K**QYGGNNAALV**D**QTAS**N**	*S. enterica*	N17D G6K D22N
M4	SDITV**K**QYGGNNAALVNQTASD	*S. enterica*	G6K

1.8.4. Pathological peptides

We then turned to examine pathological amyloids (see Table 4). The first one is Aβ42 and the second one is hIAPP. Both originate from *homo sapiens*.

Table 8 Sequence and source information of the pathological amyloids.

Name	Sequence	Source organism
Aβ42	NH_2-DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVIA-COOH	*Homo sapiens*
hIAPP	NH_2-KCNTATCATQRLANFLVHSSNNFGAILSST-$CONH_2$	*Homo sapiens*

1.9. Synthesis

1.9.1. Hexapeptides

All hexapeptide sequences selected for experimental validation were provided by CASLO (CASLO ApS, Denmark) (see Table 5). The purity of synthesized peptides was in the range between 95% and 99.6%.

1.9.2. CsgA fragments

CsgA *S. Enterica* fragments sequences were also provided by CASLO (CASLO ApS, Denmark) and R4, R2 *E.coli* fragments were synthesized "in-house". The purity of synthesized peptides was in the range between 95% and 99.6% (see SI [163]).

1.9.3. Mutants of R4 fragments of CsgA

Peptides M1 to M4 were synthesized "in-house", according to the following procedure. Peptides were obtained with an automated solid-phase peptide synthesizer (Liberty Blue, CEM) using rink amide resin (loading: 0.59 mmol/g). Fmoc deprotection was achieved using 20% piperidine in dimethylformamide (DMF) for 1 min at 90 ºC. A double-coupling procedure was performed with 0.5 M solution of differential interference contrast (DIC) and 0.25 M solution of ethyl 2-cyano-2-(hydroxyimino)acetate (OXYMA) (1:1) in DMF for 4 min at 90 °C. Cleavage of the peptides from the resin was accomplished with the mixture of TFA/TIS/H_2O (95:2.5:2.5) after 3 hours of shaking or, for a peptide with Met, TFA/thioanisole/ethane-1,2-dithiol /anisole (90:5:3:2) after 2 hours of shaking. The crude peptide was precipitated with ice-cold Et_2O and centrifuged (9 000 rpm, 15 min, 2 °C). Peptides were purified using preparative high-performance liquid chromatography (HPLC, Knauer Prep) with a C18 column (Thermo Scientific, Hypersil Gold 12 µ, 250 mm x 20 mm) with water/acetonitrile (0.05% TFA) eluent system. After that, HPLC and mass spectroscopy (MS) were performed.

1.9.4. Pathological peptides

Aβ42 and hIAPP, were purchased from ProteoGenix (Schiltigheim, France) with a purity of 95.32% and 95.27%, respectively.

1.10. Dissolving protocols

1.10.1. Hexapeptides

The lyophilized hexapeptides were dissolved and vortexed in 0.1 M NaOH. Next, phosphate-buffered saline (PBS) (50 mM, pH 7.2) was added to obtain pH = 7. The samples were diluted to the final concentration of 4 mg/ml (*ca.* 500 µM) with Milli-Q® water. Then, they were incubated at 37 °C for one month [124].

1.10.2. CsgA fragments

1 mg of peptide was dissolved in 490 µl of 0.01 M NaOH and vortexed for one min. 450 µl of 50 mM PBS of pH 7.2, which contains 0.0027 M potassium chloride and 0.137 M sodium chloride, and 60 µl of Milli-Q® water (pH 6.9) was added to the final peptide concentration of ca. 500 µM and pH 7.4. To obtain monomers, each sample was filtered through a 0.2 µm polyvinylidene difluoride (PVDF) syringe filter [163].

1.10.3. Mutants of R4 fragments of CsgA

1 mg of peptide was dissolved in 490 µl of 0.01 M NaOH and vortexed for one min. Then 450 µl of 50 mM PBS of pH 7.2, which contains 0.0027 M potassium chloride and 0.137 M sodium chloride, and 60 µl of Milli-Q® water (pH 6.9) was added to the final peptide concentration of ca. 500 µM and pH 7.4. To obtain monomers, each sample was filtered through a 0.2 µm PVDF syringe filter.

1.10.4. Pathological peptides

First, 1 mg of Aβ and hIAPP, respectively, was mixed with 1 ml of HFIP. The dissolved peptides were occasionally mixing for 1 hour 30 minutes and then sonicated for 15 minutes. Then the HFIP was evaporated under N_2 stream and put under vacuum for 3 hours. After the monomerization process peptides were studied.

1.11. Supported lipids bilayers

Supported lipid bilayers (SLB) were composed of 0.95 mg of phosphatidylcholine (POPC) and 2.75 mg of dipalmitoylphosphatidylcholine (DPPC) (final concentration 1 mM). Lipids mixture was dissolved in $CHCl_3$ in a glass flask (4.5 mL). Then the lipid mixture 0.5 mL of monomerized Aβ42 peptide dissolved in $CHCl_3$ was added and mixed (final peptide concentration 21 µM). $CHCl_3$ was evaporated with a N_2 flow and dried overnight in a desiccator under vacuum. The lipid-Aβ42 were then resuspended from the walls of the glass flask by vigorous vortexing for 10 minutes in 150 mM HEPES (4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid) (20 mM)-$CaCl_2$ (2mM) buffer (10 mL at 1 mM final total lipids concentration). To obtain small unilamellar vesicles (SUVs), the suspension was sonicated to clarity (2 cycles of 2 min) using a Branson Digital Sonifier 450 (Branson Ultrasonics Corporation, Danbury, USA) at 60% of maximal power while keeping the suspension in an ice bath to limit heating. The suspension was finally filtered on 0.2 µm nylon filters (Whatman Inc., USA) to eliminate titanium particles from the ultrasound probe. The SLBs were obtained according to the vesicle fusion method. For AFM experiments, the SUV solution was put into contact with a freshly cleaved mica substrate for 60 min at 65 °C. Then add NaCl (150 mM)+HEPES (20 mM) to stop the fusion of the lipids, wait until system cooled down to room temperature. Finally, the samples were carefully rinsed

to remove SUV excess using NaCl (150 mM)+HEPES (20 mM) buffer. The samples were transferred into the AFM fluid cell while avoiding dewetting. The hIAPP was added in the concentration of about 20 μM on the top of SLB, the peptide was dissolved in a mixture of NaCl (150 mM)+HEPES (20 mM). Then after 5 minutes of incubation at RT, the buffer was rinsed a few times in order to remove aggregates which hinder visualization with AFM.

1.12. Experimental methods

1.12.1. Circular dichroism

The CD spectra were recorded with JASCO J-815 at 20 °C between 250 and 190 nm in the reference buffer from section 2.8, with the following parameters: 0.2 nm resolution, 1.0 nm bandwidth, 20 mdeg sensitivity, 0.25 s response, 50 nm/min scanning speed, and 0.02 cm cuvette path length. The sample concentration was 500 μM. The CD spectra of the solvent alone was recorded and subtracted from the raw data. The CD intensity is given as mean residue molar ellipticity (θ [deg × cm2 × dmol−1]). Spectra were smoothed using Origin 2020b software (Origin Lab Corporation, Northampton, Massachusetts, United States).

1.12.2. Vibrational spectroscopy

1.12.2.1. Attenuated Total Reflection—Fourier Transform Infrared

The spectra were collected using a Nicolet 6700 FTIR spectrometer (Thermo Scientific, USA) with attenuated total reflectance accessory and heated diamond top-plate (PIKE Technologies), continuously purged with dry air. Each sample of 10 μL of peptide aqueous solution was dropped directly on the diamond surface and allowed to dry out. All ATR-FTIR spectra were obtained in the range of 3600–400 cm^{-1}. For each spectrum, 512 interferograms were co-added with 4 cm^{-1} resolution. All measurements were conducted at a constant temperature of 22 °C [71.6 F]. Directly before sampling, the background spectrum of diamond/air was recorded as a reference (512 scans, 4 cm^{-1}). We used 500 μM or 4mg/ml (hexapeptides) concentration of the samples, which was essential to obtain a good signal-to-noise ratio. All spectra were analyzed with OriginPro 2020b (Origin Lab Corporation, Northampton, Massachusetts, United States). The ATR-FTIR spectra preprocessing included baseline correction and smoothing using the Savitzky-Golay filter (polynomial order 2, widow size 19). Spectra were normalized to the Amide I band. Spectra in the amide bands region (1750–1490 cm^{-1}) were deconvoluted into subcomponents using the Lorentz function based on its second derivative spectra.

1.12.2.2. Infrared microscopy

The spectra from IR microscopy were recorded using a microscope FTIR Nicolet iN10 (Thermo Scientific, USA). Samples were measured with a liquid nitrogen cooled mercury cadmium telluride

(MCT-A) detector, at the spatial resolution of 10 μm. The microscope was continuously purged with dry air. An area of 450 μm x 450 μm was first selected with the upper aperture (100/5 = 50 μm), then the data were collected. All spectra were recorded in the wave number range from 4000 to 500 cm^{-1}; 64 interferograms per sample at the resolution of 4 cm^{-1} were collected. The volume of 10 μl of the solution was applied to the barium fluoride window cell and allowed to dry out. The measurements were carried out at room temperature. For each spectral map the average spectrum was calculated.

1.12.3. FT-Raman

The Raman spectra were recorded using a Nicolet NXR 9650 FT-Raman spectrometer (Thermo Scientific, USA) equipped with MicroStage extension, Nd:YVO4 laser (1064 nm, 500 mW) as an excitation source and InGaAs detector. A drop of 10 μL of each sample was deposited on the gold surface and dried under laser irradiation. All FT-Raman spectra were acquired in the range of 3700–0 cm^{-1} with 4 cm^{-1} resolution by averaging 1024 scans. Peptide concentration was 500 μM. All spectra were analyzed with OriginPro 2020b (Origin Lab Corporation, Northampton, Massachusetts, United States). FT-Raman spectra were baseline corrected, smoothed (SG, 2, 35) and normalized to band at ca. 1450 cm^{-1} attributed to deformation vibrations of methylene (CH_2) groups.

1.12.4. Thioflavin T fluorescence assay

The fluorescence of each well was read by a microplate reader CLARIOstar Plus, BMG LABTECH at 25 °C with shaking before each cycle, every 5 minutes, during 24 hours of measurements. The ThT dye was dissolved in Milli-Q® water to 2 mM concentration and filtered. Then ThT stock was dissolved in 1 M citric acid to a final concentration of 20 μM, pH 7.4. The pH was adjusted with HCl and the solution was filtered. The samples were dissolved in a 20 μM buffer to final peptide concentration about 220 μM and sonicated for ten minutes. Each well in a 96-well plate (BRAND® microplate BRANDplates®, pureGrade) was filled with 100 μL of peptide solution. The excitation wavelength was set at 440±10 nm and emission at 480±10 nm. Each group of experiments contained six parallel samples and the data were averaged after measurements.

1.12.5. Microscopy

1.12.5.1. Atomic Force Microscopy

AFM images were recorded using a Bioscope Resolve (Bruker France SAS, Palaiseau, France) operating in PeakForce-Quantitative Nano-Mechanical imaging mode. Silicon nitride probes with a nominal spring constant of 0.12 N/m and nominal tip radius of 1 nm were purchased from Bruker (PeakForce-HIRS-SSB tips, Bruker France SAS). Topography images were performed in liquid

(HEPES buffer) at 25°C with a resolution of 512 × 512 pixels, a scan rate of 0.6 Hz, and a maximal applied force up to 1.2 nN for lipid/peptide system studying.

1.12.5.2. Transmission Electron Microscopy

Imaging was performed using a transmission electron microscope Hitachi H-800, using 150 kV accelerating voltage. Negative stained samples were prepared by applying a 4 µl drop of solution, containing 0.5 µM peptide in water, on glow discharged carbon on copper grid (Agar S160). After 1 minute of adhesion, an excess of the material was blotted, and 2% uranyl acetate was applied for 1 minute before blotting. The samples were allowed to dry under normal conditions for at least 1 hour.

1.12.6. Off-null detection scanning Kelvin probe

The peptides were dissolved in water due to buffer crystallization problems, interfering with Scanning Kelvin Probe (SKP) off-null detection algorithm. 200 µL of sample aliquot was applied to Au-plated glass slides, allowed to dry completely in 22 °C [71.6 F] and 28–33% RH. CPD measurements were performed under similar laboratory conditions (21–24°C, 28–37% RH) using the environmental SKP with an off-null detection system (SKP5050, KP Technology, Scotland). All measurements were carried out at 40 a.u. amplitude and 67.0 Hz probe oscillation, while the probe backing voltage V_b (routed to the probe tip) varied between +5.0 and -5.0 V. The current gain was 4×10^7 V/A and tenfold signal averaging was used, while the gradient was maintained by the off-null detection algorithm at 150 a.u. A flat stainless steel probe tip (KP Technology, Scotland) with 1 mm end diameter was rinsed with absolute ethyl alcohol and air-dried before operation. The tip WF calibration performed using Au-evaporated reference specimen (in which the mean work function WF_{Au} of Au was assumed to be 5.25 eV[164]) provided WF_{tip}= 4.83 eV ± 0.085 eV (which flawlessly agrees with work function values WF_{316L}= 4.76–4.96 eV provided in the literature for 316L stainless steel surface with randomly-oriented crystal grains[165]]. The dried peptide deposits were scanned with 0.5 mm spatial resolution to provide 2D CPD maps. CPD measurement conditions allow to estimate the effective specimen area analyzed with the Kelvin probe of ca. 1 mm^2 (thus, CPD determined at a single point corresponded to the average CPD value in the area of approximately 1 mm^2 under the probe tip). The mean CPD in the area of the deposit was then calculated and it was finally converted to the corresponding WF value using the above calibration-determined WF_{tip}. All WF uncertainties were expressed at 95% confidence intervals.

1.13. Computational methods

1.13.1. Bioinformatics predictors

1.13.1.1. Hexapeptides

The hexapeptides were classified by bioinformatics methods, such as AmyloGram [94], PATH [166], FoldAmyloid [150], and PASTA 2.0 [149]. AmyloGram is a tool based on machine learning methods, FoldAmyloid and PASTA 2.0 are based on physical models, whereas PATH is our latest method combining physicochemical properties and structural-modeling. All predictors, excluding PASTA 2.0, were used with their default parameters available in their tools. In PASTA 2.0 the *peptide* option was chosen to set the thresholds. The presented statistics of classification results include accuracy (*Acc*) calculated as the ratio of correctly assigned data labels, sensitivity (*Sn*) denoting the ratio of correctly identified true positives versus actual positives, and specificity (*Sp*) meaning the ratio of true negatives versus actual negatives [124].

1.13.1.2. CsgA fragments

The aggregation propensity of studied the peptides was assessed using nine bioinformatics methods: AmyloGram [94], PATH [166], Pasta2.0 [149], Waltz [167], AmylPred2 [168], FoldAmyloid [150], MetAmyl [169], Tango [43], and ArchCandy [170]. Each predictor was used with its default parameters.

1.13.1.3. Mutants of R4 fragments of CsgA

The aggregation propensity of studied peptides was assessed using nine bioinformatics methods: AmyloGram [94], PATH [166], Pasta2.0 [149], Waltz [167], AmylPred2 [168], FoldAmyloid [150], MetAmyl [169], Tango [43], and ArchCandy [170]. Each predictor was used with its default parameters.

1.13.2. Theoretical and computational simulations

1.13.2.1. Mutants of R4 fragments of CsgA – structure modeling

Multiple Sequence Alignment (MSA) of non-redundant CsgA family was based on UniProt [171] (gene name “csgA”, 1058 sequences), and generated with Mafft [172], [173] in the auto mode, trimmed only to the R4 fragments. Structural models of wild type (WT) and mutant R4 octamers were generated with AlphaFold 2 [174] advanced notebook at ColabFold [175], based on previously obtained MSA. The number of generated monomers was increased to 4 or 8. Otherwise, the default options were used. The best models (all with predicted local distance difference test [174], [176] values above 84, and typically around 95) were relaxed with the Amber-Relax option.

1.13.2.2. Mutants of R4 fragments of CsgA – MD

The modeled systems were composed of three octamers obtained from AlphaFold 2 software. The protonation state of each amino acid was set to pH 7.4 [177]. The simulated systems were solvated, using the three-site transferable intermolecular potential (TIP3P) for modeling water molecules, with NaCl or K_3PO_4 ions added, at the physiological concentration of 50 mM (K_3PO_4) and 100 mM (NaCl, K_3PO_4), in the CHARMM-GUI interface [178]. The total number of atoms in the modeled system was ca. 50 000 atoms and the box size was 10×8.2×6.1 nm. The all-atom MD were performed using OpenMM 7.7 modeling framework [179] along with CHARMM36 force field [180]. First, each system was energy-minimized using the L-BFGS algorithm to remove atomic clashes. This was followed by equilibration using the isothermal–isobaric (NPT) ensemble at 310 K. We used LangevinMiddleIntegrator, and the LFMiddle discretization [181], with hydrogen mass repartition HMR=3 [182]. Additionally, MonteCarloBarostat was implemented [183]. The particle mesh Ewald (PME) method was applied for calculating long range electrostatic interactions. The Lennard-Jones potential was evaluated with a cut-off at 1.2 nm. The radial distribution function (RDF) was calculated by the MDAnalysis Python library [184].

1.13.2.3. Pathological peptides – MD

The studied Aβ42 peptide corresponds to the APP's transmembrane domain (PDB: 2LLM), which contains 28 amino acids on the N-terminal part of Aβ42 [185]. The 14 additional amino acids of the C-terminal were modeled with SWISS-MODEL [186]. The protonation state of each amino acid was set to pH 7.4 [177]. The model system of the Aβ42 peptide embedded in 1,2-dioleoyl-sn-glycero-3-phosphocholine (DOPC) bilayer was built using the CHARMM-GUI interface [178], then solvated in a NaCl water solution at both sides of the bilayer within a rectangular box in the Z-direction. The water thickness in all simulations was taken as ~ 30 Å. More information about system details can be found in Table 9.

Table 9 The number of molecules and the initial size of the simulation box in the investigated systems.

System	DOPC	Water	Na	Cl	Ca^{2+}	Initial box size (nm^3)	Force field	Simulation time (μs)
Aβ42	150	15213	43	40	0	7.15 x 7.15 x 12.7	CHARMM 36	3.75
amylin solution	0	11871	34	37	0	7.13 x 7.13 x 7.13	CHARMM 36	8.12
4 amylin	0	40650	116	128	0	10.74 x 10.74 x 10.74	CHARMM 36	3.83
amylin above	128	10038	0	55	26	6.79 x 6.79 x 11.60	CHARMM 36	8.52
amylin above	128	10128	26	29	0	9.37 x 9.37 x 11.99	CHARMM 36	5.24
amylin insert	256	16293	42	45	0	9.31 x 9.31 x 9.49	CHARMM 36	5.1
cross native + external force	256	28721	78	234	78	9.51 x 9.51x 14.46	CHARMM 36	2.83
cross native	256	28721	78	234	78	9.51 x 9.51x 14.46	CHARMM 36	5.41
cross native + external force	256	28721	78	234	78	9.51 x 9.51x 14.46	CHARMM 36m	1.2
cross native	256	28721	78	234	78	9.51 x 9.51x 14.46	CHARMM 36m	3.59
cross prion	256	26329	71	218	70	9.52 x 9.52 x 13.86	CHARMM 36m	3.57

The simulation of the embedded Aβ42 peptide was performed using the GROMACS 2021.2 software considering 3-d periodic boundary conditions [187]. All the systems were firstly minimized using the steepest descent algorithm to remove interatomic clashes, then equilibrated at constant temperature T (310 K), and constant (semi-isotropic) pressure P (1 atm). For the temperature coupling, the Noose-Hover thermostat [188] and the Parinello-Rahman barostat [189] were used. Bonds between heavy atoms and hydrogens were constrained using the LINCS algorithm [190] allowing the use of a 2 fs MD timestep. The long-range electrostatic interactions were evaluated using the Particle Mesh Ewald method [191]. The Fourier grid spacing for the Particle Mesh Ewald method was optimized at the beginning of simulation. The visualizations were rendered in VMD software [192].

In case of amylin, the peptide structure was obtained experimentally by nucleic magnetic resonance (NMR) (PDB: 2L86), resulting in structure of human amylin in sodium dodecyl sulfate (SDS) micelles at pH 7.3 [193]. The protonation state of each amino acid was set to pH 7.4 [177]. Various scenarios were simulated, the peptide was placed above the DOPC lipid bilayer, inserted into the DOPC, and in a water solution. Additionally, the behavior of amylin was simulated in presence of two types of ions. More details in Table 9. All simulations of these studied systems were performed using OpenMM 7.7 modeling framework, according to the protocol described in 1.13.2.2. However, due to higher complexity of the studied systems, additional external forces (position restrains) on the phosphate heavy atoms of lipid membrane and tetrahedral restrain to a protein (torsion forces), were added.

The systems for studying cross-seeding of Aβ42 and amylin were taken from the last frame in Aβ42 embedded in DOPC simulation (solvated in 0.15 M $NaCl_2$) and from simulation in solution (solvated in 0.15 M NaCl) in case of amylin. Amylin was placed above the DOPC near the Aβ42 peptide, to accelerate the conformational changes during the simulation time. Some systems in Table 9 have annotation "+external force", which means that external force (CustomTorsionForce with pull force constant equal 10) between Aβ42 and amylin was added. As a result, if one peptide went too far from the other, the force pulled it back. The last system, called as "cross prion", denotes the simulation of embedded Aβ42 in DOPC and monomeric cross-β structure of amylin (PDB: 6Y1A). The fibrillar structure of amylin was built up from the dimer structure of 13–37 residues cryo-EM in pH 6.0, since this condition is characteristic to intracellular secretory granules [194].

Results

Chapter 1 Hexapeptides

The initial examination concerned a set of 6 amino acid peptides (see Table 5). The choice was inspired by the typical length of amino acid sequences on which classical computational predictors are usually based and trained.

Proper classification based on the experimental results is crucial for the bioinformatics tools, which rely on it. Therefore, in this work, AmyloGrams' predictions were confirmed with other three bioinformatics predictors, based on different approaches. Other predictors were used in order to compare the classification reliability. This way, the resulting data was divided into two parts. The *reference dataset* consisted of 10 peptides with high percentage of classification consent between AmyloGram and the AmyLoad and Waltz databases. The *test dataset* contained 24 outlier sequences, which were found to have different annotations in the experimental-based databases of AmyLoad and Waltz. This raised concerns about the compatibility of the data within these databases and the potential presence of annotation errors. The classification of these sequences was determined through the examination of different experimental studies conducted under various environmental conditions. However, there is currently no widely accepted protocol for assessing the amyloidogenic propensity of peptides and proteins. The main reason for that, is the fact that their various physicochemical properties influence inter alia the dissolving properties. Additional factor could be a potential human error.

The primary objective of this initial study was to investigate the impact of experimental conditions on the different classification of the same amyloid, with the special emphasis on the IR spectroscopy technique. In this examination, two IR modes were compared (transmission and attenuated total reflectance). The results point out the usefulness of such data in case of bioinformatics predictors, which are based on the experimental input data. The outcomes from various predictors were compared, which showed their sensitivity to improperly classified amyloids. Through this examination, a tentative protocol for studying the aggregation properties of peptides was established and potential amyloidogenic amino acid sequences were identified. Additionally, the study evaluated the effect of poorly annotated training data on learning of bioinformatics methods.

The validation of protocol started with the choice of the ***solvent*** and the peptide ***concentration*** to be used in IR spectroscopy techniques. The hexapeptides were dissolved according to the protocol in Section 1.8.1. Firstly, the effect of NaOH on the peptides was tested to enhance their solubility thanks to its alkaline pH. The pretreatment of the synthetic peptide with sodium hydroxide is a proper method to achieve increased solubilization results [195], without using organic solvents or

sophisticated preparation. Then PBS and water were added to obtain the proper concentration of peptide and the transition in pH from about 11 to neutral. In addition, heavy water was tested (see Figure 9A). D_2O is frequently used in this technique due to the shift of the bending mode of OD groups in D_2O molecules to 1210 cm^{-1}. Whereas, in H_2O the bending mode of OH groups is about 1635 cm^{-1} and obscures Amide I signal. However, if we compare the results obtained from NaOH+PBS with D_2O, we can observe a shift to the lower wavenumbers of approximately 3 cm^{-1} in the Amide I' band. In addition, the overall shape of Amide I' is changed in comparison to Amide I. No shoulder band at 1690 cm^{-1}, in spectrum of peptides dissolved in heavy water (see Figure 9B), means the secondary structure changes. Generally, heavy water is known to stabilize proteins (especially larger ones such as actin, tubulin, β-lactoglobulin A), enhance aggregation process, make proteins more compact and less flexible than in H_2O, increase the solubility of polar amino acids and proline, change pK_A value, and shift the equilibrium between inter-residue hydrogen bonds [65].

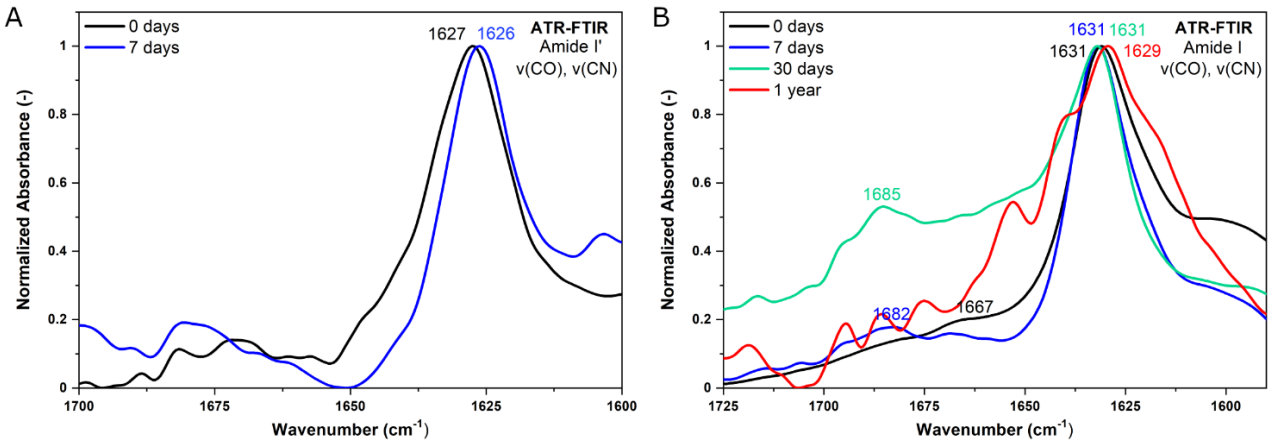


Figure 9 (A) Exemplary, normalized ATR-FTIR spectrum of ISFLIF in the Amide I' (1700–1600 cm^{-1}) region. Peptide dissolved in D_2O on the day of dissolving and after 7 days of incubation at 37°C. **(B)** Exemplary, normalized ATR- FTIR spectrum of ISFLIF in the Amide I (1700–1600 cm^{-1}) region. Peptide dissolved in NaOH+PBS at various incubation times. Cpep=500 µM.

On the contrary, increased aggregation in D_2O was observed for longer sequences up to 23 amino acids, particularly the fragments of CsgA protein of *S. enterica* (results showed in the further section 2.2.1.).

The next phase of the study involved determining an ***effective concentration***. In this research, the concentration was established at 500 µM. Lower concentrations were found to result in a significant decrease in the intensity of absorbance and poor signal-to-noise ratio (as shown in Figure 10). The intensity of the IR spectrum is directly proportional to the quantity of material present.

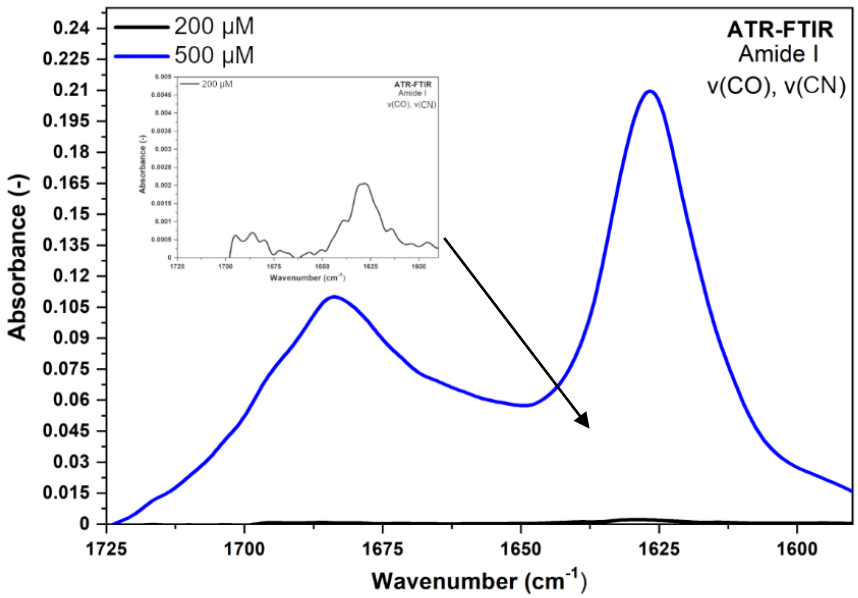


Figure 10 Exemplary ATR-FTIR spectrum of YTVIIE peptide in concentration: 200 µM and 500 µM in the wavenumber range of 1725–1590 cm^{-1} (Amide I). Additionally, rescaled spectrum of 200 µM .

The next step in establishing the aggregation protocol was to determine an appropriate ***incubation time*** for evaluating the aggregation propensity of the peptide. It was found that 30 days (incubation at 37°C) was the optimal time for such identification. However, spectra were also analyzed after longer periods of time, up to one year. During the aggregation process we can observe that the maximum of the Amide I band is shifted towards lower wavenumbers, meaning more ordered structures (see Figure 11). The time required for peptide/protein classification may vary and may be challenging to predict. As an example, peptides with low aggregation potential, such as the hexapeptide NPQGGY, may not form aggregates, even after extended periods of time (data not presented). Conversely, peptides with higher aggregation potential may form fibrils after a specific period of time, such as a few days at 37°C.

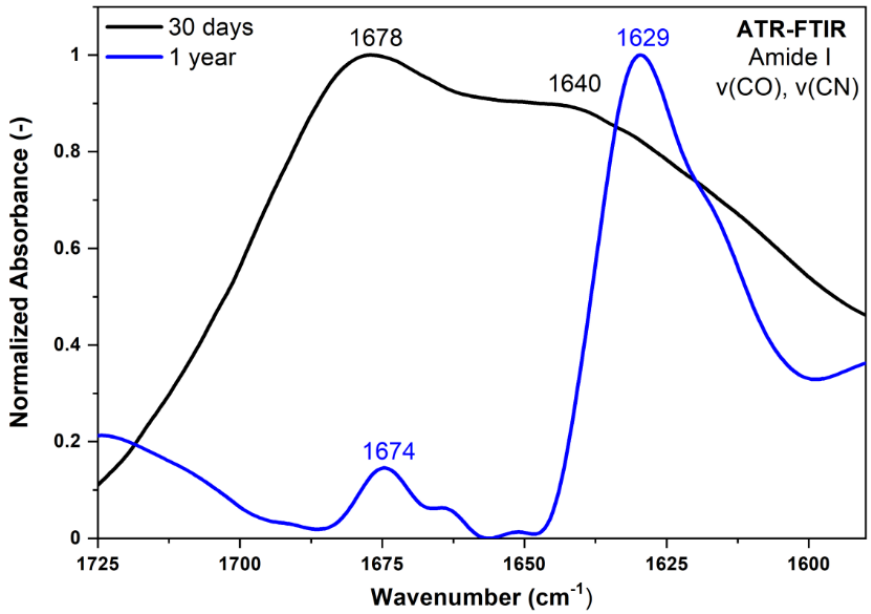


Figure 11 Exemplary, normalized ATR-FTIR spectrum of TKPAES in the Amide I (1700–1600 cm^{-1}) region. Peptide dissolved in NaOH+PBS after 30 days of incubation at 37°C and one year. Cpep = 500 µM.

The *test dataset* and *reference dataset* were analyzed using two IR spectroscopy ***techniques***, ATR-FTIR and µIR, as detailed in sections: 1.5.2.1 and 1.5.2.2. This technique is rapid and enables the differentiation of oligomers from fibrils. The results from two measurement modes were then compared. In ATR-FTIR spectra, the amide bands positions were shifted towards lower wavenumbers. It indicates the acceleration of the aggregation process, due to the additional interaction of the peptide with diamond surface. Consequently, this interaction causes structural modifications within the peptide/protein molecules [111], [196]. On the other hand, in µIR, the composition of dehydrated material revealed huge crystals of PBS buffer, which resulted in spectral signals that overlapped with those of the peptides (see Figure 12). Nevertheless, µIR allows for the selection of a specific ROI for analysis.

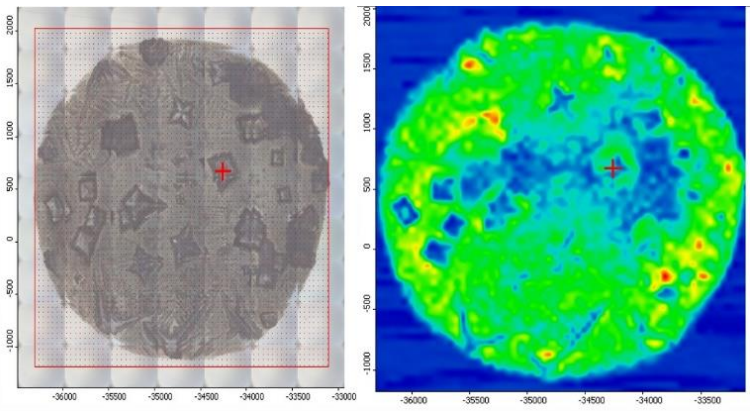


Figure 12 Microscopic image (x15) on the left, spectral map obtained by IR microscopy on the right. PBS Crystals are visible in square and star-like shapes on the microscopy image. Peptides aggregates are present in the remaining part of the dehydrated drop.

Considering the previously established parameters: solvent, peptide concentration, and incubation time, the results of the two IR modes were compared for both *reference data* and *test data*. Based on the spectral characteristics, three general classes were distinguished: amyloid, non-amyloid, and oligomer (see Table 10 and Table 11). None of the bioinformatics predictors takes into account the class of oligomer. It is clear in experimental analysis, and could be a challenge for computational tools to classify it correctly. In IR, amyloid oligomers possess two distinct stretching frequencies. One at about 1630 cm^{-1} and the second at about 1695 cm^{-1}. This second frequency is attributed to a strong dipolar coupling, which results in a frequency intensity that is significantly lower, roughly five times, compared to the primary frequency at 1630 cm^{-1} [115].

However, in the case of amyloids, variations in the position of the Amide I band in ATR-FTIR and μIR spectra can be observed. These variations may be caused by differences in the accessibility of the carbonyl group to water molecules, which in turn affects the hydration level. The peptides which were annotated as amyloids, possess a dominant conformation of intermolecular aggregates (see Table 2). The second classification group was created by examining the presence of oligomers. Oligomers have unique characteristics, as described in sections 1.2. and 1.3. They possess a more complex secondary structure composition. An example of this is the SFLIFL peptide as shown in Figure 15. All spectra from Table 10 and Table 11 are available in SI material of [124]. The last class comprised of non-amyloids, which in bioinformatics predictors are classified negatively. It should be highlighted that in amyloid predictions is the insufficient representation of sequences that have been negatively annotated. Researchers often do not publish negative data, which may cause difficulties for computational methods.

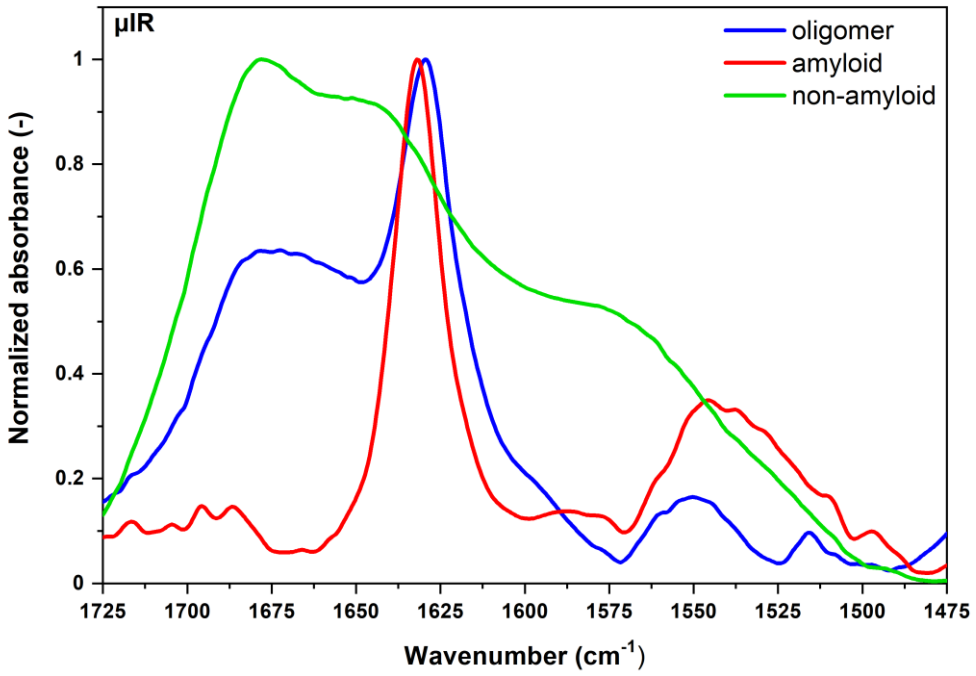


Figure 13 The representative μIR spectra of the selected classifications in the wavenumber range of 1725–1475 cm^{-1}. [124]

Table 10 Reference set of sequences and their amyloid propensity by different experimental methods (yes – identified as amyloid, no – non-amyloid, yes* – oligomer, s – strong band, m – medium band, w – weak band, br – broad band, sh – shoulder band, band maxima in bold). [124]

No	Sequence	Data-base	IR microscopy		ATR-FTIR		AFM	Con-sensus with data-base annota-tion
			Amide I [cm^{-1}]	Class	Amide I [cm^{-1}]	Class	Class	
1	FNPQGG	no	1679(m)/**1641(s)**	no	**1665(s,br)**	no	no	yes
2	FTFIQF	yes	1689(m,sh)/**1628(s)**	yes	1690(w)/**1622(s)**	yes	yes*	yes
3	ISFLIF	yes	1689(m,sh)/**1631(s)**	yes	1685(w)/**1631(s)**	yes	yes	yes
4	KPAESD	no	**1665(s,br)**	no	**1678(s,br)**/1640(m,sh)	no	no	yes
5	LVFYQQ	yes	**1631(s)**	yes	1683(w,sh)/**1629(s)**	yes*	yes	yes
6	NPQGGY	no	**1658(s,br)**	no	**1658(s,br)**	no	no	yes
7	SFLIFL	yes	1689(m)/**1633(s)**	yes*	**1632(s)**	yes	yes*	yes
8	TKPAES	no	**1652(s,br)**	no	**1678(s)**/1640(sh)	no	no	yes
9	YLLYYT	yes	1686(m,sh)/**1629(s)**	yes	1685(m)/**1630(s)**	yes	yes*	yes
10	YTVIIE	yes	1685(m)/**1627(s)**	yes	1684(m)/**1626(s)**	yes	yes	yes

Table 11 Test sequences and their amyloid propensities (yes – identified as amyloid, no – non-amyloid, yes* – oligomer, s – strong band, m – medium band, w – weak band, br – broad band, sh – shoulder band, band maxima in bold), compared with the original database annotation, which is in disagreement with AmyloGram classification. [124]

No.	Sequence	Database	IR microscopy		ATR-FTIR		Consensus with database annotation
			Amide I [cm^{-1}]	Class	Amide I [cm^{-1}]	Class	
1	ALEEYT	yes	**1655(s,br)**	no	**1654(s)**	no	no
2	ASSSNY	yes	**1649(m,sh)**	no	**1655(m,br)**	no	no
3	DETVIV	no	1685(w)/**1635(s)**	yes*	1685(m)/**1633(s)**	yes*	no
4	ELNIYQ	no	1661(w,sh)/**1635(s)**	no	1681(m,br)/1668(m,br)/**1635(s)**	no	yes
5	FGELFE	no	**1660(s)**/1650(w)	no	**1659(s)**	no	yes
6	FQKQQK	no	**1660(s,br)**	no	**1682(s,br)**	no	yes
7	FTPTEK	no	**1660(s,br)**	no	**1680(s,br)**	no	yes
8	HGFNQQ	yes	**1662(s,br)**	no	**1682(s,br)**	no	no
9	HLFNLT	yes	**1674(s,br)**	no	**1680(s,br)**/1633(m,br)	no	no
10	HSSNNF	yes	**1649(m,br)**	no	**1680(s)**/1646(m,sh)	no	no
11	MIENIQ	yes	**1656(s,br)**	no	**1655(s,br)**	no	no
12	MIHFGN	yes	**1677(s,br)**	no	**1680(s,br)**/1646(m,br)	no	no
13	MMHFGN	yes	**1675(s)**	no	**1676(s,br)**	no	no
14	NIFNIT	yes	**1657(s)**	no	**1663(s,br)**	no	no
15	NNSGPN	yes	1676(sh)/**1648(s,br)**	no	**1676(s,br)**/1654(m,br)	no	no
16	NTIFVQ	no	**1629(s)**	yes	1682(w)/**1631(s)**	yes*	no
17	QANKHI	yes	**1680(s,br)**	no	**1681(s)**/1653(sh)	no	no
18	QEMRHF	yes	**1679(s,br)**	no	**1676(s,br)**/1655(sh)	no	no
19	SHVIIE	no	1688(m)/**1630(s)**	yes	1684(m)/**1633(s)**	yes	no
20	STTIIE	no	**1657(s,br)**	no	1681(m)/**1630(s)**	yes*	yes
21	STVVIE	no	1685(w,br)/**1633(s)**	yes	1682(w,br)/**1630(s)**	yes*	yes ambiguous
22	SWVIIE	no	1682(w,sh)/**1631(s)**	yes	1684(w)/**1631(s)**	yes*	no
23	WSFYLL	no	**1658(s,br)**	no	1675(w,sh)/**1637(s)**	yes	no
24	YYTEFT	no	**1665(s,br)**	no	**1659(s,br)**	no	yes

The sequences from Table 10 and Table 11 were examined by hand. Additionally, the PCA (for more, see 1.5.4.) were applied on the spectroscopic data. Based on PCA plots, in the *reference dataset*, only non-amyloid group was separated from all examined peptides (see Figure 14A & B). On the other hand, in the *test dataset*, the PCA revealed that better data separation was observed for IR microscopy (see Figure 14D), than ATR-FTIR (see Figure 14C). These results showed how even small differences in the experimental technique influence the obtained amyloid classification. This could be one of these factors may lead to misannotation of the sequences.

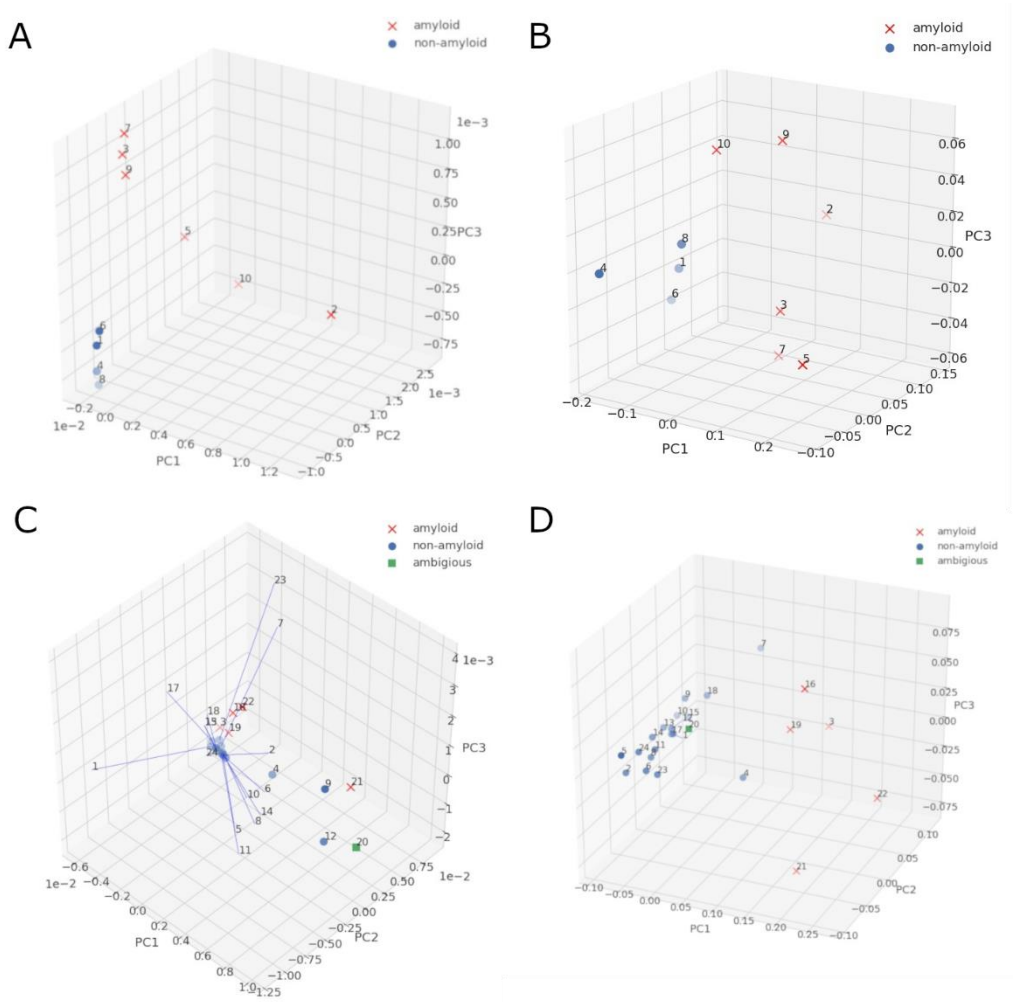


Figure 14 PCA plots of: **A** and **C**, ATR-FTIR spectra, **B** and **D**, µIR spectra. Amyloids marked with red cross, non-amyloid with blue dot, and ambiguous sequences with green square. Modified from [124].

It was found that a significant portion of the non-compatible sequences, 17 out of 24, had been inaccurately labelled in the initial databases (as demonstrated in Table 11). The AmyloGram system was able to identify and correct these errors, and was not impacted by overfitting. Comparisons were made between the predictions of the AmyloGram system and other bioinformatics predictors, such as FoldAmyloid, PASTA 2.0 and PATH, which had also been based on the same mislabeled data. These predictors were able to correctly classify the mislabeled data with an accuracy of 80% or more (as reported in [124]). However, the effect of errors in the input data on predictors utilizing statistical and machine learning methods has yet to be determined. It should be noted that the results may vary depending on the experimental conditions used for prediction.

Based on the results described above, a procedure for dissolving peptides was developed. Hexapeptides are dissolved in 0.1 M sodium chloride, then 50 mM PBS buffer is added along with water in order to achieve the appropriate concentration, of about 500 µM. The incubation temperature was set up to 37 ºC and the aggregation was assessed after one month of incubation time at the same

temperature. However, even minor modifications in experimental design can impact the obtained result. Additionally, the bioinformatics predictors can be fairly robust to certain level to insignificant annotation errors.

The results obtained in this study were inspiration for the first hypothesis: *minor deviations in experimental conditions strongly influence the validity of a protein classification as an amyloid.*

Chapter 2 Functional amyloids – CsgA fragments

The subsequent phase of the study involved a comprehensive examination of longer amino acid sequences, specifically those involving up to 23 amino acids, which originate from functional amyloids (see Table 6 and Table 7). Previously, the protocols used for dissolving and further peptide studies, were only applied to hexapeptides (see Table 5). This raises the question of whether similar conditions can be applied for longer sequences. Furthermore, the examination of the aggregation process of functional amyloids mimics more closely approximate physiological conditions. Additionally, functional amyloids allow for greater control over the aggregation process compared to studying pathological amyloids.

One of the most well-researched functional amyloids is CsgA, which is found e.g. in *E. coli*. Functional amyloids exhibit structural characteristics similar to pathological amyloids, however, they possess distinct stability and lifetimes [7], [15]. This is attributed to the presence of repetitive units, which play a similar role as APR [197]–[199], and contribute to the formation of curli in bacteria. The CsgA protein of *E. coli* is composed of five imperfect repeats known as R1–R5 fragments, which add up to a total of 151 amino acids. The fragments share a common pattern, which in *E. coli* follows as Ser-X5-Gln-X4-Asn-X5-Gln, where amino acids in X-positions may differ depending on the bacterial species. Among these fragments, R1 and R5 are considered as the most amyloidogenic and are critical for seeding and the curli formation in *E. coli* [161], [200]. Research has shown that peptides fibrillate in the order of R5 > R1 > R3, while R2 and R4 do not fibrillate at a concentration of 0.2 mg/mL. However, at a concentration of 2 mg/mL, fibrils were observed for R2 and R4 fragments through TEM microscopy [56].

The objective of this research was to investigate the structural determinants of the amyloidogenic potential of repeats and evaluate the impact of intrinsic and extrinsic factors on aggregation propensity, with the goal of gaining a deeper understanding of these processes.

To test my first hypothesis: *minor deviations in experimental conditions strongly influence the validity of a protein classification as an amyloid*, my approach focuses on how the use of deuterium oxide affects the aggregation process of *S. enterica* fragments. I have chosen this particular CsgA homologue from *S. enterica* because it has not been studied before, and this is an important area of study as *S. enterica* is a common food pathogen that can cause serious illness in humans. A second hypothesis was proposed due to the noted variations: *aggregation propensity of closely related homologous functional amyloids may be different*. Furthermore, influence of minor sequence differences on the behavior of CsgA R4 fragments in *E. coli* and *S. enterica* was studied. Despite the significant similarity between *E. coli* and *S. enterica*, I inquire whether the aggregation properties of these fragments are the same.

2.1. Fragments characterization

The CsgA proteins from *E. coli* and *S. enterica* are closely related, with 75% identity as determined by BLAST analysis. To investigate any differences between the corresponding fragments, pairwise alignments were conducted (see Figure 15).

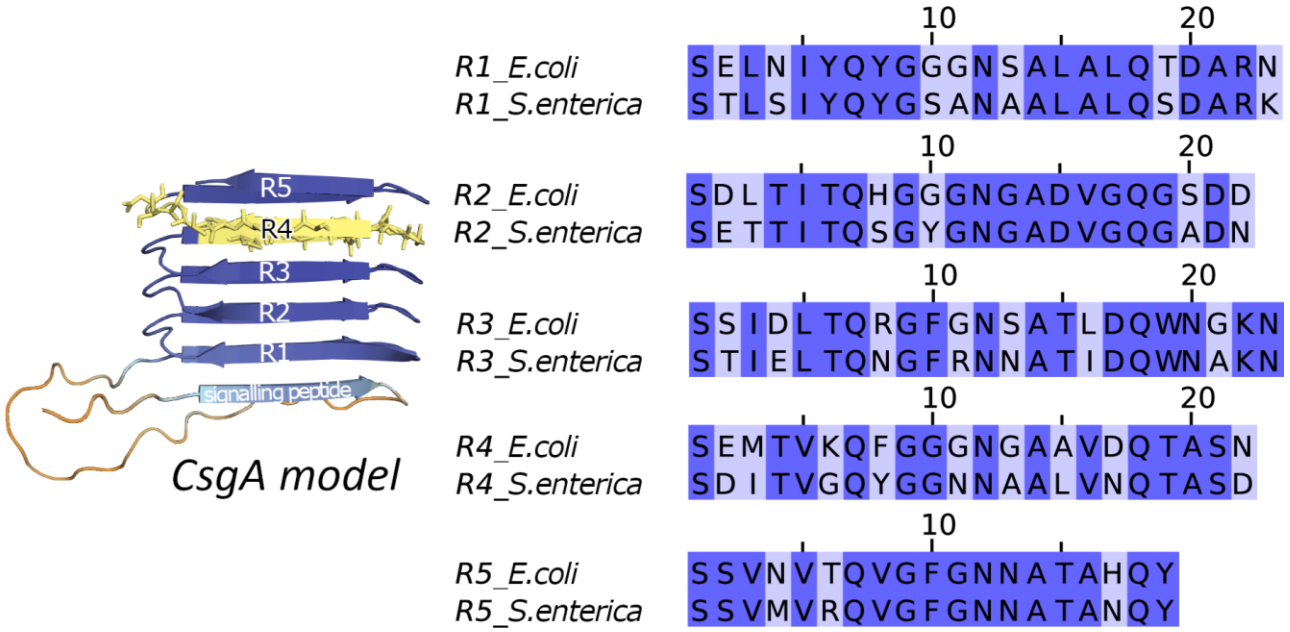


Figure 15 On the right: pairwise sequence comparison of the CsgA fragments from *E. coli* and *S. enterica* bacteria. The variations in amino acid composition were represented by light purple highlighting. Modified from [163]. On the left: structural model of CsgA protein from *E. coli* (strain K12) predicted by AlphaFold and available in the AlphaFold database.

These alignments confirmed that the sequences are highly similar, with almost all of the peptides sharing similar features such as hydrophilic properties, charged amino acids at the ends, and a flexible glycine-rich linker in the middle. This suggests a tendency towards β-arch structure, which is consistent with computational results from ArchCandy. However, their aggregation propensities may vary due to point mutations in the csgA genes, which can change properties such as hydrophobicity or pI. Mutations in aromatic amino acids, which stabilize amyloid peptide structures, can also play a significant role in the formation of aggregates. Similarly, charged residues can affect amyloid propensity depending on their location. Not only mutations of individual residues, but also minor changes in the sequence order, can influence peptide aggregation susceptibility. This was discovered through statistical analysis and has led to the development of various bioinformatics predictors. The sequence alignment of corresponding pairs from both bacterial species suggests that different amyloid propensities of the CsgA fragments cannot be ruled out. Not only theoretical analyses but also bioinformatics analysis were carried out [163]. Our study used two amyloid predictors, AmyloGram and PATH, which were trained on hexapeptide data from AmyLoad and Waltz 2.0 databases. Both methods identified R1

fragments as amyloidogenic, while R2 and R3 fragments as non-amyloidogenic for both bacterial species. However, R4 fragment from *S. enterica* was classified as amyloidogenic by AmyloGram, while R4 from *E. coli* was classified as non-amyloidogenic, indicating a potential difference in aggregation propensities. Additionally, several other bioinformatic predictors were applied for comparison, but their predictions were not consistent (see Table 12).

Table 12 Prediction results of amyloidogenicity. Where 0 denotes non-amyloid, 1 stands for amyloid. [163]

Organism	Fragment	AmyloGram	PATH	PASTA2.0	Waltz	AmylPred2	FoldAmyloid	MetAmyl	ArchCandy*
E. coli	R1	1	1	0	1	0	0	0	1
	R2	0	0	0	0	0	0	1	1
	R3	0	0	0	0	0	0	0	1
	R4	0	0	0	0	0	0	1	1
	R5	0	0	0	0	0	0	1	1
S. enterica	R1	1	1	0	1	1	1	1	1
	R2	0	0	0	0	0	0	1	1
	R3	0	0	0	0	0	0	0	0
	R4	1	0	0	0	0	0	1	1
	R5	0	0	0	0	0	0	1	1

* β-arch motif predictor

Based on the performed analyses of bioinformatics predictors, it can be observed that functional amyloids are problematic for them when used as their input data. Functional amyloids are underrepresented in the reference datasets used to train computational predictors of amyloids, leading to the hypothesis that the statistical sequence profile of functional amyloids may differ significantly from that of pathological amyloids. This is likely due to the structural differences, varying temporal properties, stability and susceptibility to environmental factors and interactions between the two classes of amyloids.

2.2. Experimental analysis of fragments from different bacteria

Firstly, the aggregation propensity of *E. coli* and *S. enterica* fragments were compared. The aggregation propensity of all *E. coli* and *S. enterica* fragments was investigated using spectroscopic techniques such as: CD, ATR-FTIR and μIR. Additionally, TEM was used to examine the morphology of the selected fragments.

CD spectroscopy was used to understand the general characteristics of the secondary structure of CsgA fragments. The CD spectra of *E. coli* fragments are shown in Figure 16A.

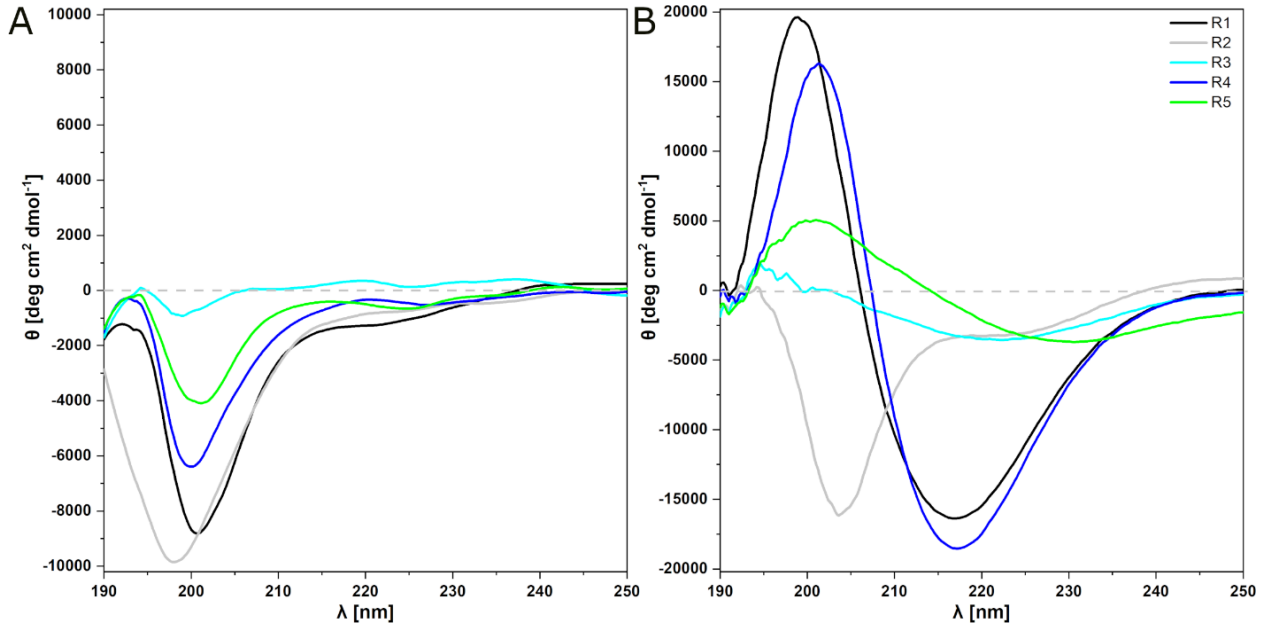


Figure 16 Far-UV CD spectra of CsgA fragments on the day of the dissolving. A Spectra of *E. coli* fragments. B Spectra of *S. enterica* fragments. Cpep = 500 µM. [163]

On the day of dissolving the samples, a minimum at around 200 nm was observed in all spectra taken from the *E. coli* fragments, which is a characteristic of a random coil conformation. The CD spectra of *S. enterica* fragments are in Figure 16B. The spectra of the R1, R3, R4 and R5 fragments shared the characteristics of β-sheet conformation. Only R2 fragment revealed a signature, at 203 nm, which is characteristic to a random coil structure. The results of this analysis indicate that the fragments from *S. enterica* exhibit a higher propensity to adopt a β-sheet conformation compared to their counterparts from *E. coli*. To assess more detailed information about secondary structure between bacteria, the ATR-FTIR and µIR were conducted.

A comparative analysis was conducted on the spectra of *E. coli* fragments obtained from samples dissolved initially (Figure 17A) and after incubation at 37°C for one month (Figure 17B). Results indicated that fragments R1, R3, and R5 displayed a main band located below 1630 cm^{-1} in the spectrum obtained on the day of dissolution, which corresponded to cross-β amyloid architecture and indicated the presence of amyloidogenic aggregates. Further analysis of the second derivative spectrum of fragment R3 revealed that it formed more ordered and rigid fibrils compared to R1 and R5. The Amide I band in the spectra of R1, R3, and R5 had an additional local maximum located at approximately 1665 cm^{-1}, which was attributed to parallel β-sheet structures and is characteristic of loop-turn conformations in β-helix structures. Conversely, fragment R2 displayed a complex spectral characteristic with a broad local maximum at 1678 cm^{-1} assigned to β-turns and a shoulder at 1648 cm^{-1} typical for random coil structures, indicating a lack of aggregation properties. Although fibers were observed in TEM

experiments with high concentrations of R2 or R4, and long incubation times, fragment R4 displayed low absorbance of β-sheet components in the secondary derivative range.

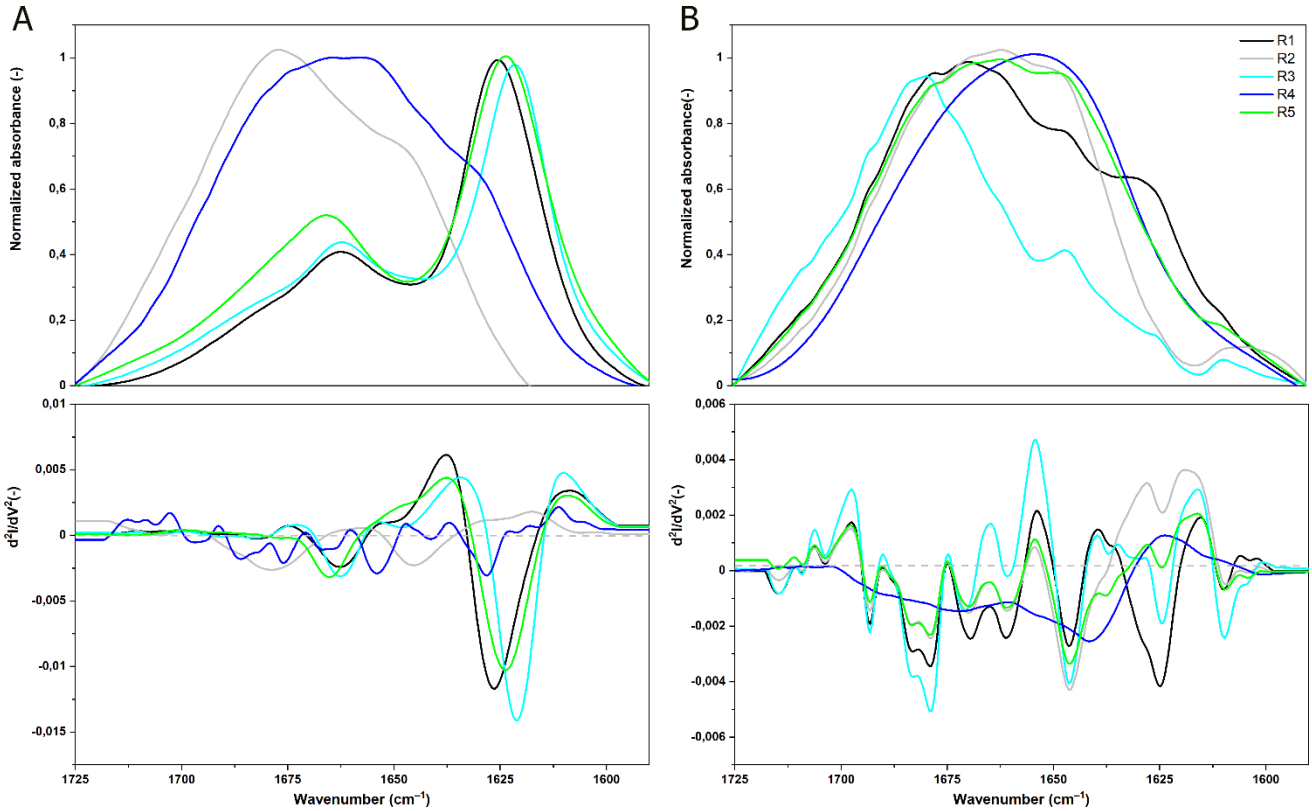


Figure 17 Normalized ATR-FTIR spectra of *E. coli* fragments, with the second derivatives spectra, in the wavenumber range of 1725–1590 cm^{-1} (Amide I). **(A)** on the day of the dissolving. **(B)** after one month of incubation at 37 °C. Cpep = 500 μM. [163]

After a month of incubation, the spectra of *E. coli* fragments showed that the Amide I bands in R1, R3, and R5 were broadened and lost their signatures of aggregates (see Figure 17B). However, these fragments still remained partially aggregated. R2, R4, and R5 showed random structures and lost stability over time due to deamination of amino acids. This process weakened the fibrils and led to their eventual disintegration. The instability of *E. coli's* CsgA fibrils is different from the stability of hexapeptides and pathological amyloids.

2.2.1. Solvent influence

In this study of *S. enterica* fragments of CsgA protein, the effects of two solvents, D_2O and NaOH+PBS, were evaluated in two IR modes, using R1–R5 fragments. On the day of dissolution in NaOH+PBS, R3, R4, R5 fragments displayed high intensity absorbance at about 1622 cm^{-1}, indicative of long and rigid amyloid fibrils. Additionally, R1 fragment also displayed characteristic typical for amyloids, however the Amide maximum was shifted towards lower wavenumbers at 1626 cm^{-1}, which indicates less rigid fibril. In turn, fragment R2 showed a broad band at 1645 cm^{-1}, characteristic of

disordered proteins. After 30 days of incubation at 37°C, no significant changes were observed in the ATR-FTIR spectra [163]. *S. enterica* fragments were found to be more stable than those of *E. coli*, but both disintegrated after 3 months of incubation. The locations of the most characteristic spectral components are presented in Table 13, highlighting the propensity of each peptide for amyloid aggregation and its secondary structure.

Table 13 The assignments of secondary structures in the range of Amide I (ν(CO) 76%, ν(CN) 11%, δ(CCN) 8%, β(NH) 5%) of CsgA fragments of *S. enterica* studied by ATR-FTIR and μIR in D_2O and NaOH+PBS solvents. Where: sh – shoulder band, m – medium band, the most band intense local minimum in bold. Cpep = 500 μM.

After dissolving									
Mode	**ATR-FTIR**					**μIR**			
Solvent	**D_2O**		**NaOH+PBS**			**D_2O**		**NaOH+PBS**	
Structure	**aggregates**	**β-turns**	**aggregates**	**β-turns**	**random**	**aggregates**	**β-turns**	**aggregates**	**β-turns**
R1	**1624**	1665(sh)	**1626**	1660(sh)		**1627**	1665(sh)	**1624**	1663(sh)
R2	**1624**	1666(sh)		**1667**	1645(m)	**1625**	1666(sh)	1627(sh)	**1664**
R3	**1618**	1671(sh)	**1622**	1661(sh)		**1629**	1671(sh)	**1625**	1665(sh)
R4	**1622**	1667(sh)	**1622**	1660(sh)		**1622**	1667(sh)	**1625**	1666(sh)
R5	**1619**	1666(sh)	**1622**	1666(sh)		**1625**	1666(sh)	**1625**	1666(sh)

Meanwhile, the studies in D_2O, in ATR-FTIR mode, confirmed the presence of fibrils for all *S. enterica* fragments (Table 13). In contrast to the results in NaOH+PBS, where fragment R2 had a spectral characteristic for a β-turns and random structure. The aggregation of R2 fragment of CsgA protein of *S. enterica* was accelerated due to D_2O. This fragment formed fibrils directly after dissolving in heavy water, based on the TEM micrographs. In turn, in NaOH+PBS, the fibrils were not observed, see Figure 18. In case of other fragments of CsgA protein, the morphology of fibrils in heavy water is different than in NaOH+PBS solution. This confirms the influence of the solvent on morphology of studied fragments.

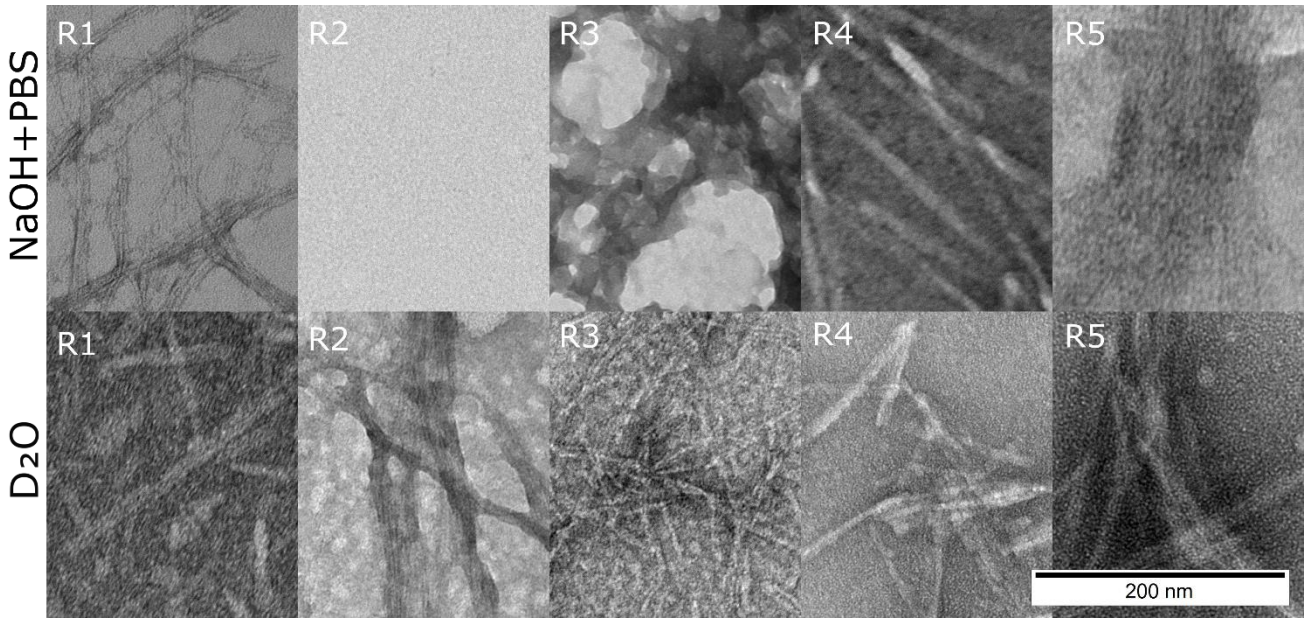


Figure 18 Electron micrographs of CsgA fragments after dissolving (magnification of 40 000) First row dissolved in NaOH+PBS, second row dissolved in heavy water. Cpep = 500 μM.

The IR spectroscopy also revealed differences in the secondary structure of studied fragments of CsgA, see Table 13. These results are in agreement with TEM. ATR-FTIR spectra of R1, R3, R4 and R5 showed a dominant absorbance about 1622 cm^{-1} (see Figure 19). Based on that, we can observe that the most intense local minimum in R2 fragment, in D_2O, in both IR modes is shifted to the lower wavenumbers (see Figure 19). This feature at about 1624 cm^{-1} is an indicator of fibrils. In NaOH+PBS, R2 fragment is showed to adopt predominantly β-turn structure and random. The secondary and tertiary structures of studied fragments were different in D_2O and NaOH+PBS, as observed with IR spectroscopy and TEM techniques. The acceleration of aggregation can be caused by hydrogen-deuterium exchange in the peptide sequence. No general expected effect of D_2O on acceleration of aggregation of amyloid proteins in case of short fragments (hexapeptides) was observed. This effect may be dependent on the amino acid composition, as well as length of the amino acid sequence. However, this issue requires further systematic research.

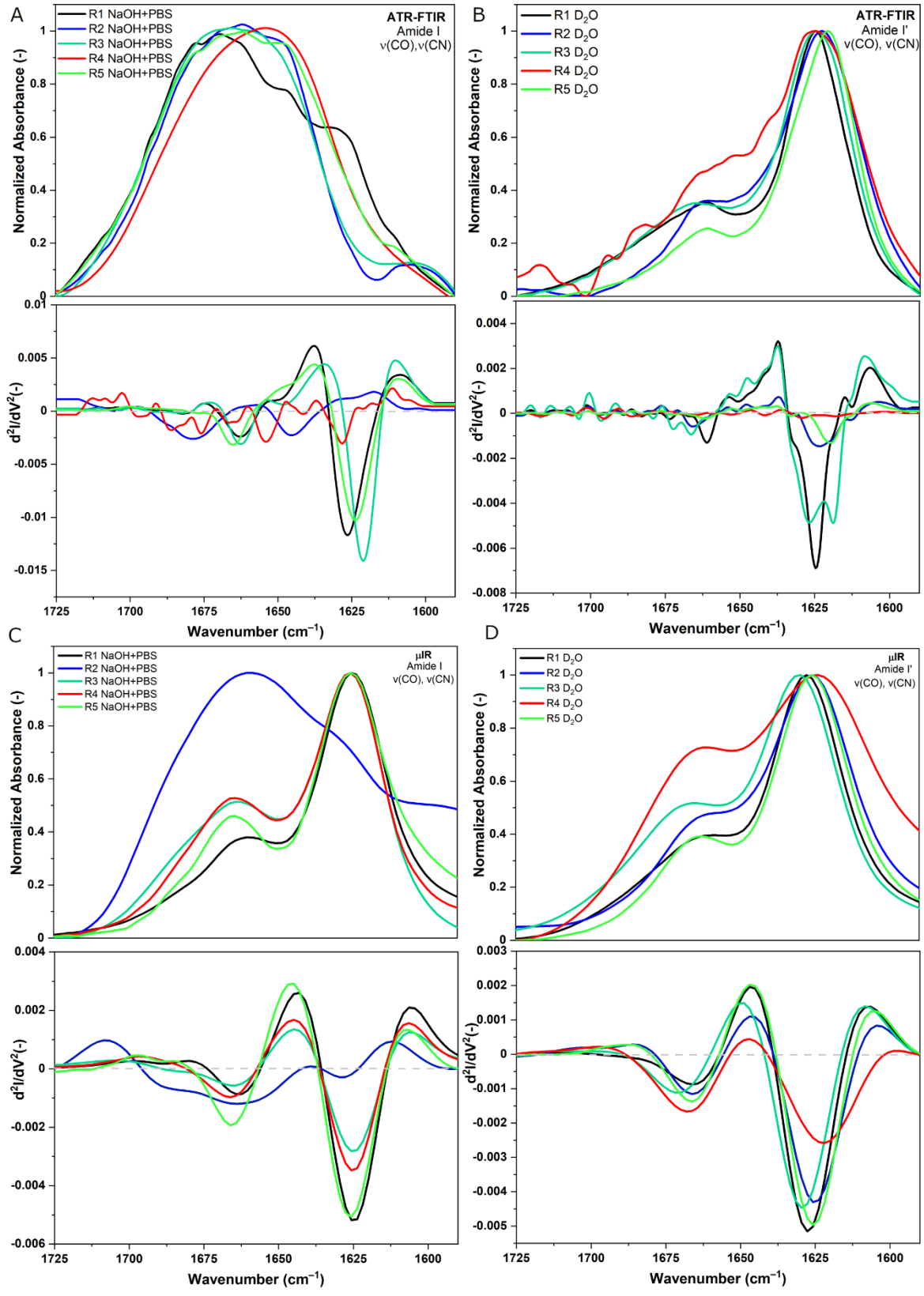


Figure 19 Normalized ATR-FTIR and μIR spectra of CsgA fragments of *S. enterica* in the Amide I (1700–1600 cm^{-1}) region. Peptide dissolved in NaOH+PBS and D_2O directly after dissolving. Cpep = 500 μM.

Based on the spectral assignments presented in Table 13, the interaction of peptides with the hydrophobic surface of the ATR accessory diamond may have accelerated the formation of amyloid fibrils in the ATR-FTIR experiments in comparison to transmission mode.

To sum up, the ER4 possess spectral characteristic which is not typical for amyloid structures (Figure 17A). In turn, SR4 had subcomponent of intermolecular aggregates, characteristic for amyloids. Additionally, this study was supported also with TEM and ThT assay [163]. The fluorescence of R4 of *S. enterica* was roughly 9 times higher compared to that of the *E. coli* fragment (see Figure 20). No lag phase was observed, indicating rapid aggregation. The fibrillation stages of R4 of *S. enterica* reveal an immediate elongation phase followed by a saturation phase.

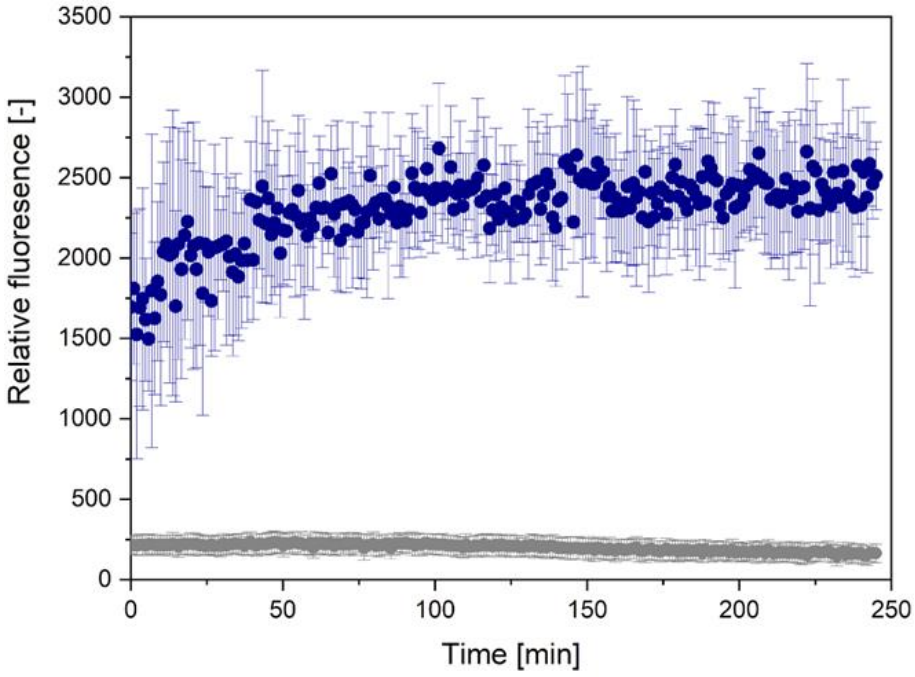


Figure 20 ThT fluorescence curves over time for R4 fragments, with gray dots representing *E. coli* and blue dots representing *S. enterica*. Cpep = 500 µM.

Additionally, it was observed that the micrographs of the R4 fragment from *E. coli*, taken on the day of dissolution, did not exhibit fibrils (see in [163] Figure 9). This observation contrasted with the micrographs of *S. enterica*, which displayed fibrils. However, after seven days of incubation of R4 of *E. coli* at 37°C, fibrils were also observed.

This highlights the presence of amyloid propensity in both bacterial species, however, the aggregation process of isolated R4 fragment of *S. enterica* was found to be more rapid compared to that of R4 of *E. coli*.

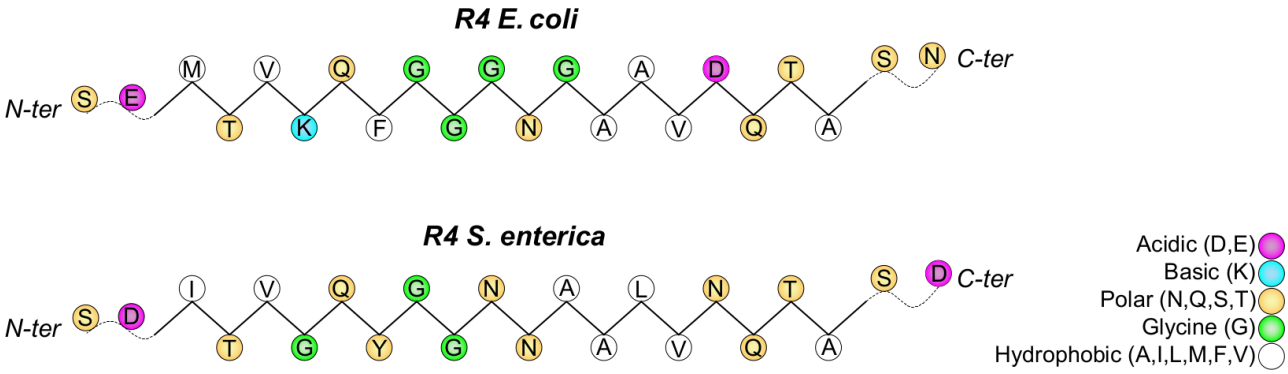


Figure 21 Comparison of the physicochemical properties of the amino acids in the R4 fragments of *E. coli* and *S. enterica*.

In support to my 1st and 2nd hypothesis, two observations can be made. First, the solvent used in the experimental protocol can change the result and amyloid classification. Secondly, even minor differences in ER4 and SR4 influence the self-assembly process. For example, R2 fragment is classified as amyloid in deuterium oxide and as non-amyloid in NaOH+PBS, because it did not possess spectral signature characteristic for amyloids, in this solvent. The IR results were supported with TEM analysis which did not reveal fibrils. On the other hand, the effect of deuterium oxide was not so strong for hexapeptides (see Figure 9), which could be possibly caused by much shorter amino acid sequence. My study proved the larger aggregation tendency of SR4 in comparison to ER4 fragment. Based on the theoretical sequence analysis, the ER4 has lower hydrophobicity (the grand average of hydropathicity index, GRAVY: -0.5, net charge: -1) and lower electric charge value (-1) than SR4 (GRAVY: -0.359, net charge: -2), see Figure 21. Dobson et al. [45] showed that mutations that increase hydrophobicity and reduce surface charge of a protein result in significant pro-aggregation effects. This is reflected in a higher probability of the protein conformation converting from α to β. What is more, the R4 fragment exhibits a more significant variation in gatekeeper residues in *E. coli* and *S. enterica* species. Gatekeeper residues are essential amino acid residues present in functional amyloid proteins, which have a crucial function in regulating the formation of amyloids. These residues are generally found in or near the β-strand region of the protein and have the ability to stabilize or destabilize the protein structure, which can impact the amyloid formation process. Consequently, the presence or absence of gatekeeper residues can significantly affect the propensity of the protein to form amyloid structures [51]. In the case of the ER4, the presence of glycine (G) at position 13 and aspartic acid (D) at position 17 has been found to slow down aggregation. However, in SR4, these residues are replaced by alanine (A) and glutamic acid (N) (see Figure 21), respectively, which could potentially increase the amyloidogenicity. These findings suggest that amino acid substitutions at specific positions can have a significant impact on the propensity of functional amyloid proteins to form amyloid structures, which can have important functional consequences.

In this study, I investigated the problem of accurately predicting the functional amyloids using bioinformatics methods. Experimental studies demonstrated that fragments from *S. enterica* were more robust and stable than those from *E. coli*, even after a month of incubation at 37°C. The use of experimental methods allowed for the comprehensive characterization of fragments from both *E. coli* and *S. enterica* species. Overall, the findings of the study indicated that not only the physicochemical properties of the sequence but also environmental factors are crucial in influencing the propensity for aggregation.

Chapter 3 Mutants of R4 fragments

Another important factor which influences misfolding of proteins and peptide sequences, are mutations. Mutations can have a destabilizing effect on the native structures of proteins and may lead to aggregation, as seen in the case of familial amyloidosis. Gaining insight into how even minor changes such as single substitutions, amino acid position swapping, or mutations in the peptide/protein sequence can accelerate or decelerate the aggregation process would be valuable. This knowledge could have benefits beyond the study of pathological amyloids. Functional amyloids are a special group, because their function is evolutionary conserved in the sequence. Based on the sequence homology, they reveal high similarity of the sequence within one organism.

The previous results showed that amyloid structures of the SR4 are much more easily formed and more stable than those in ER4 [163]. Taking into account the differences in both R4 fragments, the point mutations designed involved replacing amino acids in the sequences to make ER4 more similar to SR4 and vice versa. The number of possible substitutions was very high. All of them were tested using the AmyloGram tool, after which I selected substitutions with the same or lower amyloid score value. Point mutations were introduced into fragment R4 of *S. enterica* and *E. coli* to study the self-assemble properties of these fragments.

In this section I will present the results of the effects of mutations on the aggregation propensity of peptide sequence, in support of the hypothesis: *aggregation propensity of closely related homologous functional amyloids may be different*. The aim of this examination was to investigate the effects of specific mutations on the amyloid propensities of the peptides, the morphology and structural properties of the fibrils.

Additionally, in the study, a new experimental approach, based on measurements of CPD and WF, using a SKP, was employed. The impact of ions on amyloid formation was investigated through the use of MD, by comparing the effects of K_3PO_4 and NaCl. This part of the work is bound with hypothesis: *minor deviations in experimental conditions strongly influence the validity of a protein classification as an amyloid.*

3.1. Theoretical characterization of designed peptides

The designed and synthesized R4 fragments from the two species (ER4 and SR4) and their mutants are the following: M1 (D11N, N22D) based on ER4 and M2 (L15A), M3 (G6K, N17D, D22N) and M4 (G6K) based on SR4 (see Figure 22).

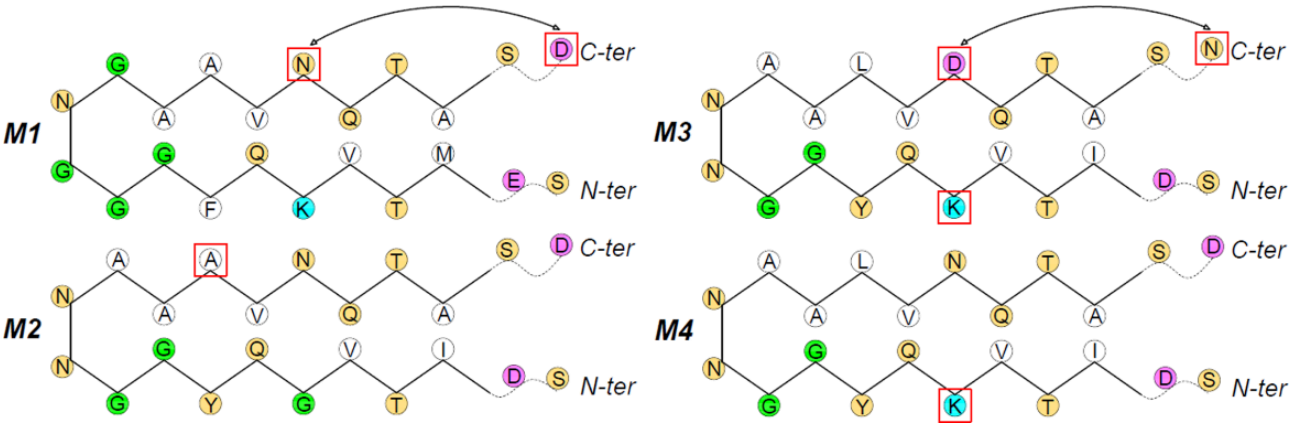


Figure 22 Schematic representation of mutant sequences with introduced mutations, marked with red squares.

Analysis showed that the amino acid substitutions did not cause substantial changes in the biochemical characteristics of each peptide regarding parent sequences, as reported in Table 14. All peptides had hydropathy scores below zero, which is characteristic of hydrophilic proteins [200]. Their pI values are mostly acidic (Table 14) and no significant changes in the net charge and the predicted solubility were found (Table 19). Nevertheless, the bioinformatics predictor AmyloGram showed emergence of two distinct classes: amyloids and non-amyloids (Table 15).

Table 14 Biochemical characteristics of the studied peptides (the introduced mutations are bolded and underlined): the net charge, pI, the grand average of hydropathicity index (GRAVY), the solubility as measured by CamSol and results of AmyloGram predictions.

Name	**Source organism**	**Sequence**	**Mutation**	**Net charge**	**pI**	**GRAVY**	**CamSol**	**AmyloGram (amyloid score) [0; 1] threshold: 0.5**
ER4	*E. coli*	SEMTVKQFGGGNGAAVDQTASN	WT	-1	4.07	-0.500	1.64	0 (0.27)
M1	*E. coli*	SEMTVKQFGGGNGAAV**N**QTAS**D**	D17N N22D	-1	4.07	-0.500	1.64	0 (0.27)
SR4	*S. enterica*	SDITVGQYGGNNAALVNQTASD	WT	-2	2.93	-0.359	1.57	1 (0.71)
M2	*S. enterica*	SDITVGQYGGNNAA**A**VNQTASD	L15A	-2	2.93	-0.450	1.65	0 (0.47)
M3	*S. enterica*	SDITV**K**QYGGNNAALV**D**QTAS**N**	N17D G6K D22N	-1	3.88	-0.518	1.60	1 (0.71)
M4	*S. enterica*	SDITV**K**QYGGNNAALVNQTASD	G6K	-1	3.88	-0.518	1.50	1 (0.71)

Furthermore, the aggregation propensity of the peptides was assessed using nine bioinformatics tools (see 1.13.1.3) that predict amyloidogenicity. Most of them identically classified each of the peptides (as amyloidogenic or non-amyloidogenic, Table 15).

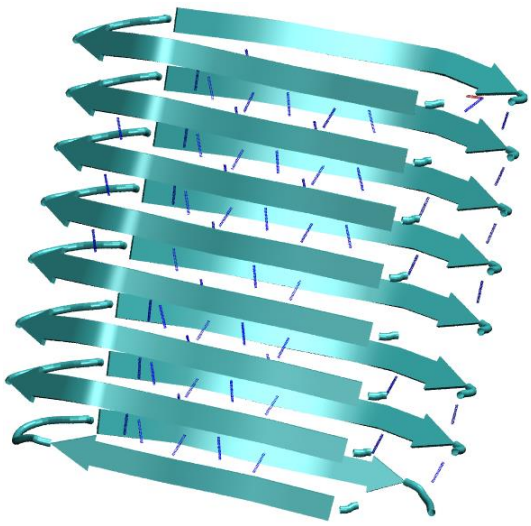
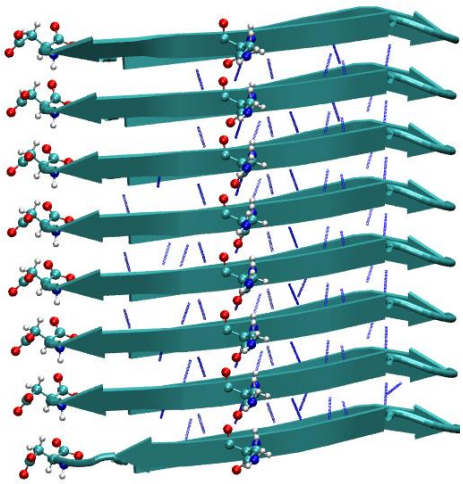
Table 15 Results predicted by various bioinformatic tools. Where * denotes to amorphic aggregate, and ** to slower aggregation. Peptides classified as non-amyloidogenic are marked in yellow, score 0. Amyloidogenic peptides are marked in green, score 1. Light green represents peptides with score close to the classification threshold.

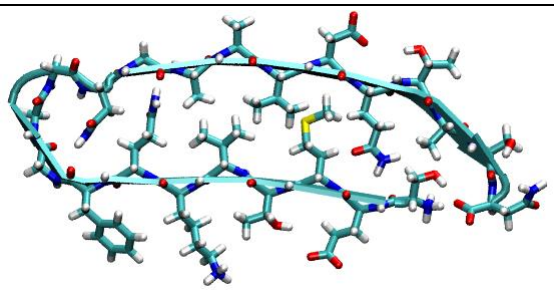
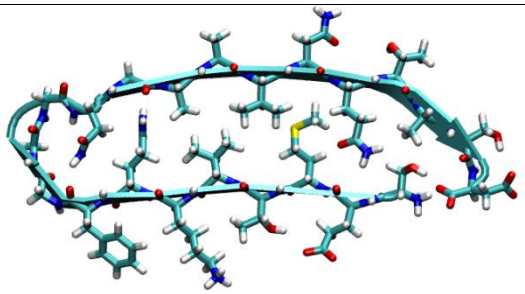
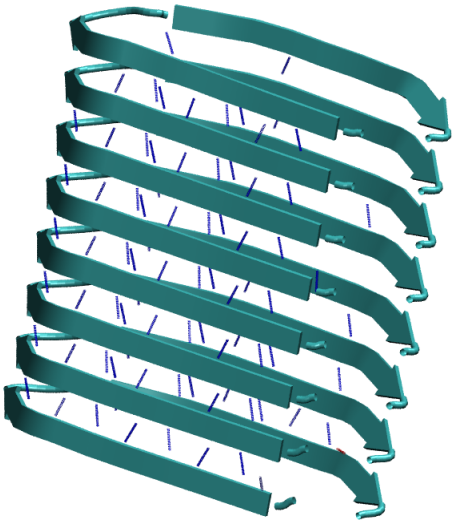
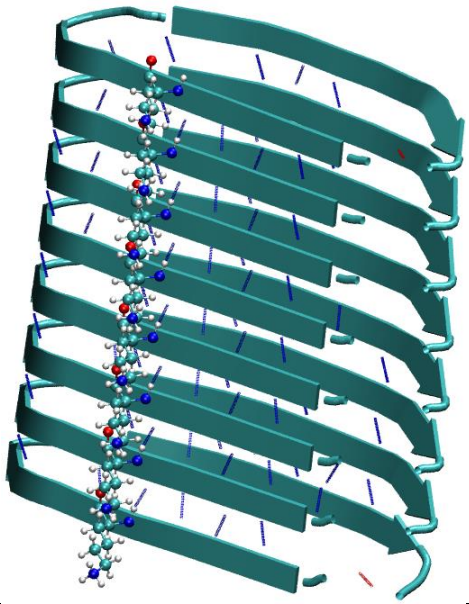
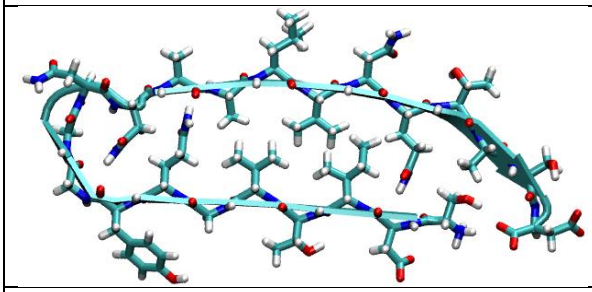
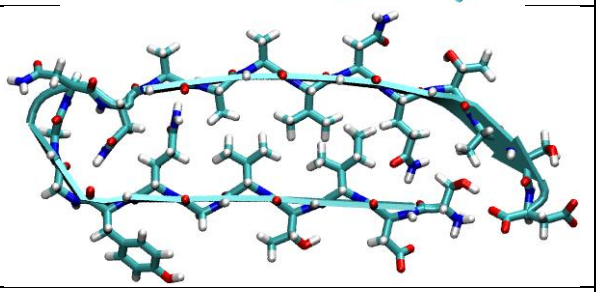
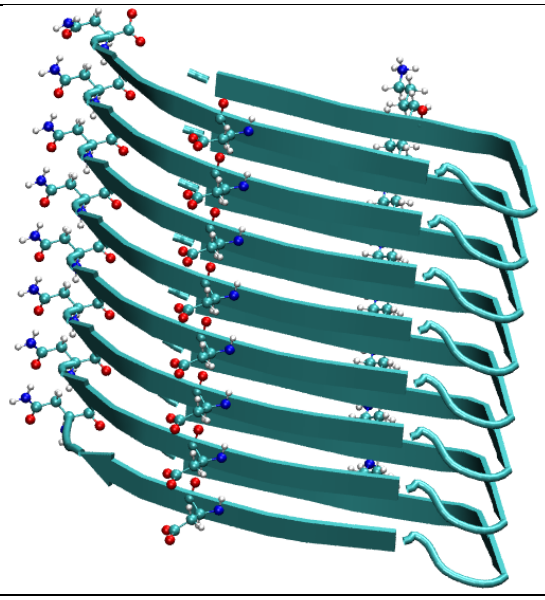
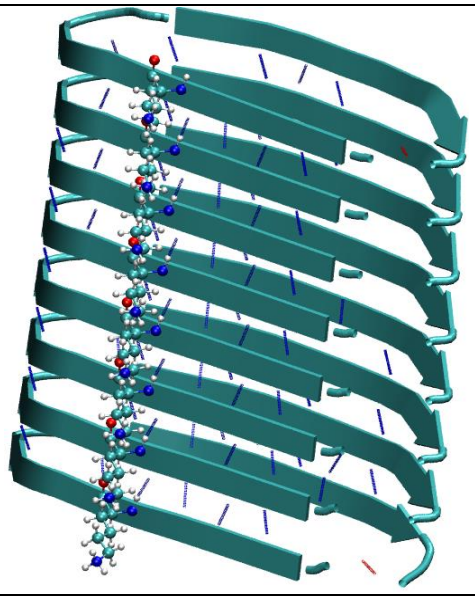
Sequence	Experimental	AmyloGram	PATH	PASTA 2.0	FoldAmyloid	Aggrescan	Waltz	MetAmyl	ArchCandy
ER4	1**	0 (0.27)	0 (0.16)	0	0	1	0	1	1
M1	0*	0 (0.27)	0 (0.16)	0	0	1	0	1	1
SR4	1	1 (0.71)	0 (0.32)	0	0	1	0	1	1
M2	1	0 (0.47)	0 (0.31)	0	0	1	0	1	1
M3	1	1 (0.71)	0 (0.42)	0	0	1	0	1	1
M4	1	1 (0.71)	0 (0.42)	0	0	1	0	1	1

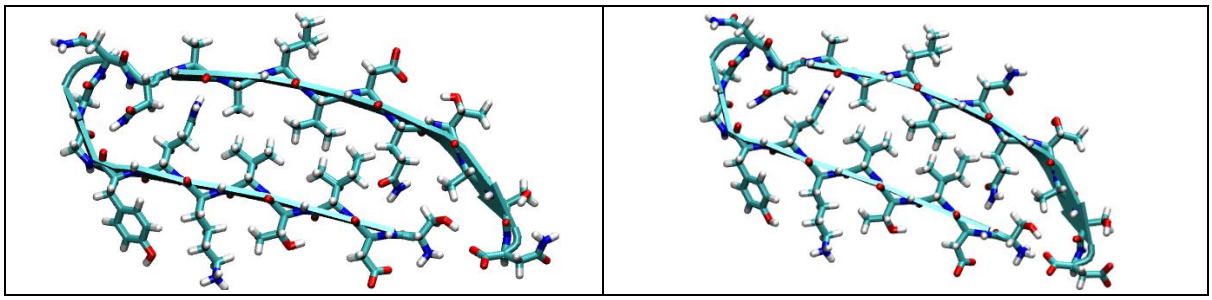
However, the output of AmyloGram was different. The predictor classified ER4, M1 and M2 as non-amyloidogenic, while SR4, M3 and M4 as amyloidogenic (Table 15). Nevertheless, the amyloid score (0.47) of peptide M2 was close to the classification threshold (0.5) and its amyloidogenic propensity was very close to positive. As we will show later, these results were in line with our experimental analyses.

The amyloid propensities of the peptides were also investigated based on their structure. Structural models of octamers of R4 peptides were assessed using AlphaFold 2 (Table 16).

Table 16 The results predicted by AlphaFold 2. Each model showed in the cartoon representation, the mutations drawn as Corey–Pauling–Koltun coloring (CPK), hydrogen bonds colored in dark blue. The protein is drawn in the cartoon representation and the mutated amino acid is presented in the CPK representation.

ER4	M1
	

	
SR4	M2
	
	
M3	M4
	



All predicted models represented cross-β conformation with the presence of a highly flexible, glycine rich, loop region. Similar arrangement could be observed in the CsgA model derived from the AlphaFold database (see Figure 15). Noticeably, no significant structural differences resulting from mutations could be observed in the models predicted by AlphaFold 2 (Table 16), which would imply identical structural and amyloid propensities of all modified peptides. This could be attributed to the fact that AlphaFold is often insensitive to single point mutations. The ArchCandy modeling method specifically devoted to amyloids, confirmed susceptibility of all studied peptides to form β-arch structural motifs, which indicates their potential to form amyloid conformations (Table 15).

3.2. Experimental characterization of mutants

Experimental analysis included classical methods widely used in amyloid studies and a novel quantitative approach using the Kelvin probe. The general characteristics of the peptides' secondary structure were first studied by CD. A random coil conformation with a minimum of ca. 200 nm was identified for M1 and ER4 peptides (Figure 23).

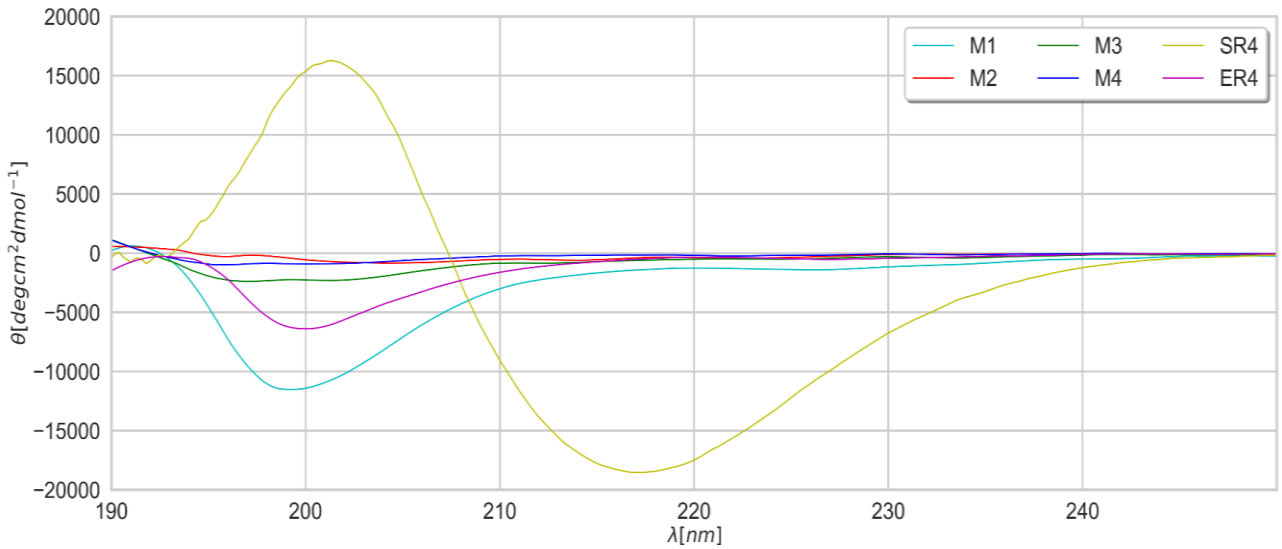


Figure 23 Far-UV CD spectra of mutant fragments with initial R4 fragments on the day of the dissolving (final peptide concentration 500 μM).

The CD spectrum of SR4 showed a maximum at 201 nm and a single minimum at 216 nm, which can be assigned to the β-sheet conformation. Peptides M2, M3 and M4 showed hallmarks of the presence/formation of beta sheets. To obtain more detailed information about the peptide secondary structures, ATR-FTIR spectra were collected and analyzed in the Amide I and II region (1750 to 1490 cm^{-1}), based on the performed deconvolution (Figure 24 and Figure 60). It was observed that about six subcomponents of the Amide I band were assigned to: β-antiparallel (1695 cm^{-1}), beta turn (1685 cm^{-1} and 1675 cm^{-1}), β-helix (1665 cm^{-1}), β-sheet (1635 to 1625 cm^{-1}), aggregates (1624 cm^{-1} to 1610 cm^{-1}) and other structures (i.e. extended unordered structures and random coil, as well as residual water). The percentage area of the sub-band corresponding to aggregates clearly distinguishes amyloidogenic peptides from those which are not able to form fibrils under introduced conditions. For SR4 the value of this parameter is 11%, while ER4 does not exhibit this band at all. This is in line with the fact that R4 from E. coli's CsgA is generally considered to not have self-assembly properties, and as such was classified in previous study (see 2.2.). In the case of mutant peptides, the highest value was calculated for M2 (25%) and lowest for M1 (0%).

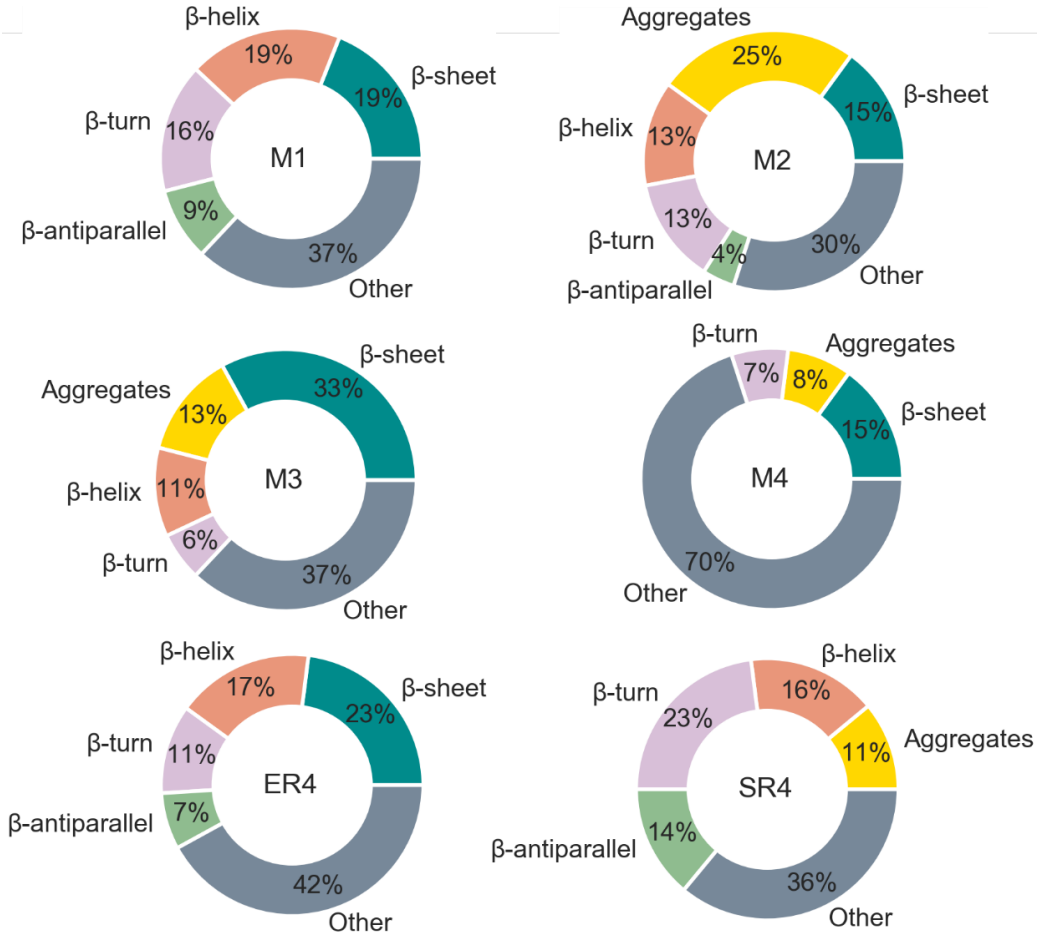


Figure 24 The percentage contribution of the Amide I band subcomponents to the total area (from 1750 to 1490 cm^{-1}). Data based on deconvolution of ATR-FTIR spectra of M1 to M4, SR4 and ER4 peptide samples on the day of dissolving. Cpep = 500 µM.

The analysis of FT-Raman spectra corroborated the previously reported results (Figure 25). In the FT-Raman spectra of R4 fragments of CsgA, from both studied species, the Amide I band is located at ca. 1670 cm^{-1}, which is characteristic of beta structures [201].

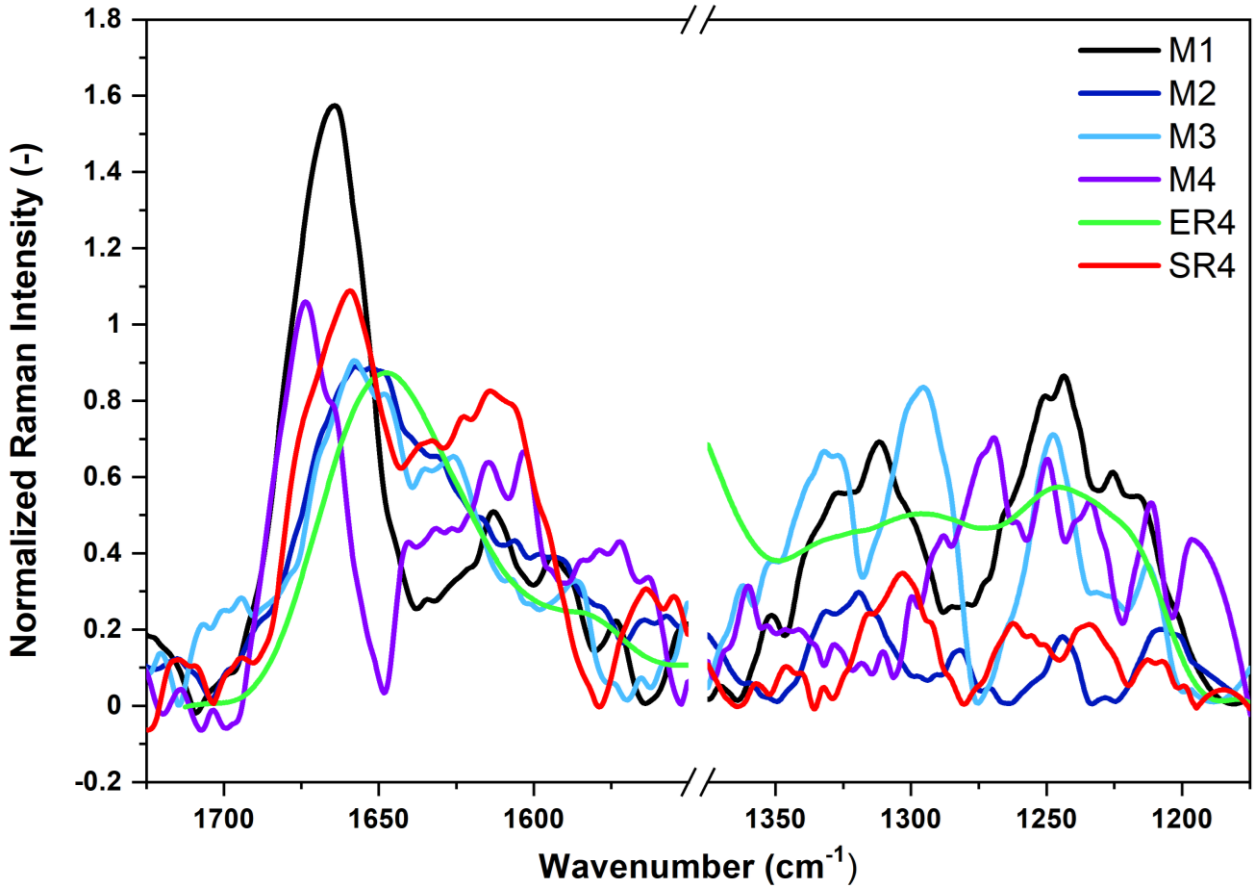


Figure 25 Normalized FT-Raman spectra of studied fragments, smoothed with SG 35, in the wavenumber range of 1725–1185 cm^{-1} (Amide I & III).

However, in case of SR4 this spectral feature is more intensive and narrower, indicating fibril formation [130]. A similar FT-Raman signature was observed for M2 peptide, which proved its amyloidogenic propensity. The overall shape of the Amide I bands in other peptides differed significantly, suggesting more complex structures. Based on the spectroscopic techniques, all studied sequences adopted β-hairpin conformations, in accordance with our structural modeling results.

The ThT fluorescence assay was utilized to evaluate the kinetics of the aggregation process (see Figure 26). The fastest aggregation was observed for the M1 peptide, with the shortest lag phase and a duration of 15 hours.

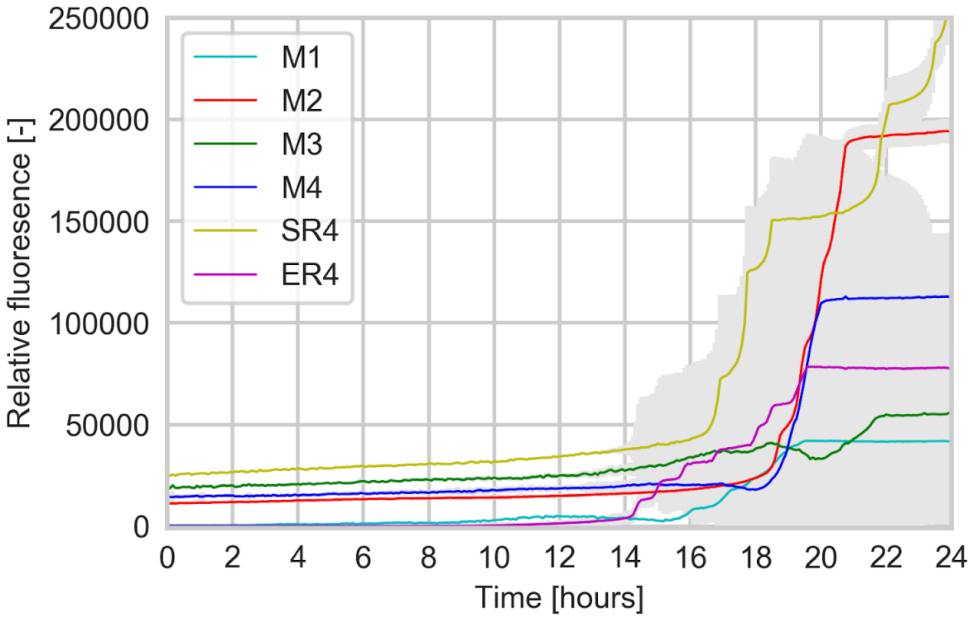


Figure 26 ThT curves for mutant sequences following the aggregation process.

The ER4, M2, M3, and M4 fragments had a longer lag phase of approximately one hour, and the SR4 fragment had a lag phase of about two hours, which was longer than that of the M1 peptide. The rising value of the ThT end point intensity was observed in the following order: SR4 (250×10^3 RFU)→M2 (194×10^3 RFU)→M4 (112×10^3 RFU)→ER4 (75×10^3 RFU)→M3 (55×10^3 RFU)→M1 (42×10^3 RFU), this intensity is proportional to the amount of formed amyloid [38].

TEM was employed to study the morphology of the peptides and the impact of the mutations on the self-assembly process. TEM images (Figure 27) reveal amorphous aggregates in the case of the M1 peptide. For M3, TEM images revealed the presence of several interconnected clusters of oligomers with an average size/length of around 7 µm (Figure 61). In contrast, the M2 mutant was found to form long, rigid fibers, measuring 160–175 nm in width (Figure 61) and several micrometers in length. These observations are consistent with the typical characteristics of amyloids [5], which are known to have an overall length of many micrometers and consist of unbranched fibrils that are 5–15 nm wide.

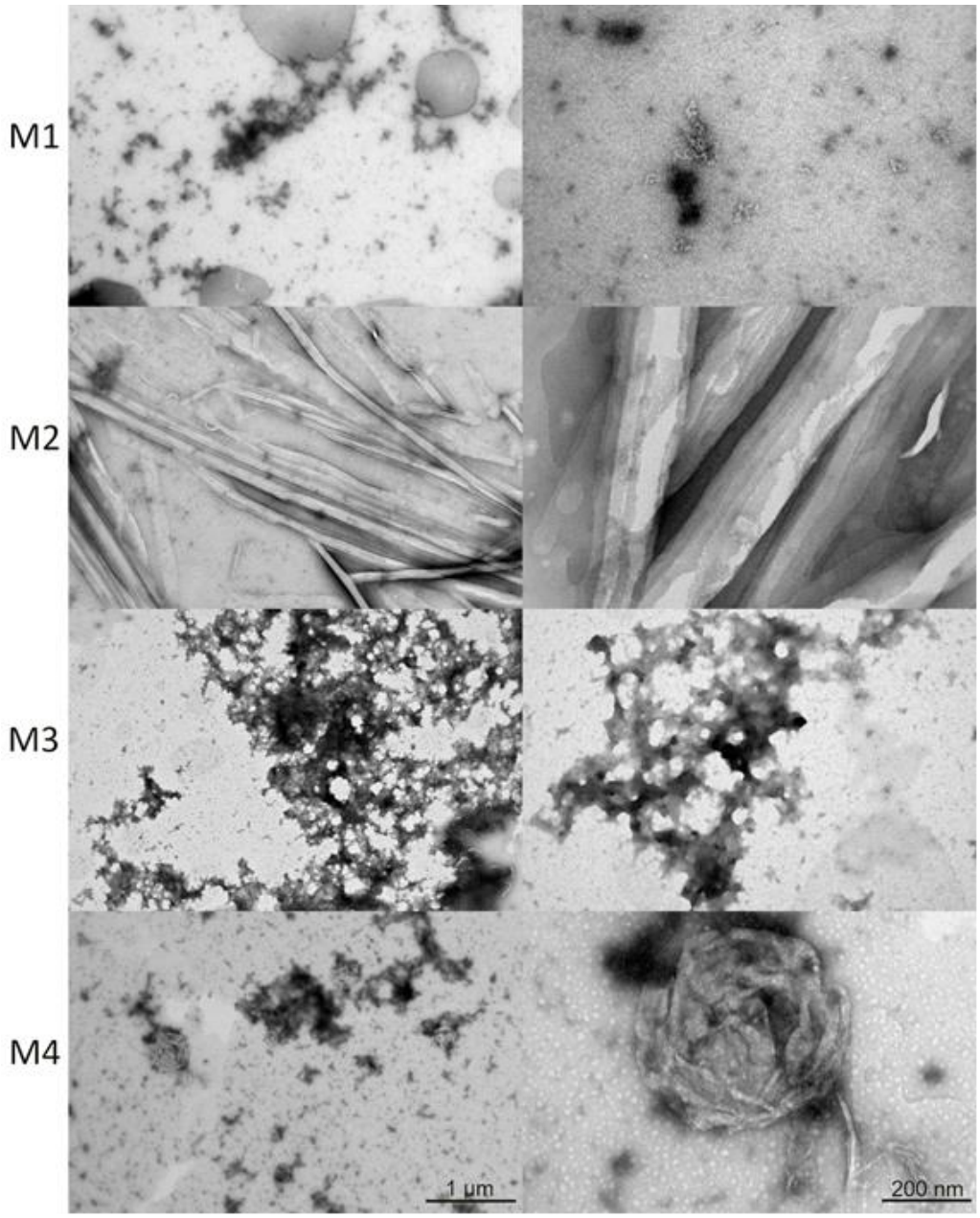


Figure 27 Electron micrographs of mutants on the day of dissolving (left column: magnification of 10 000, right column: magnification of 40 000). Cpep = 500 µM.

The most noteworthy structures were observed for peptide M4, which displayed snail-like, circular structures. As depicted in Figure 27, the M4 mutant formed a unique cellular structure characterized by a single fibril curling up into a spherical shape. The snail-like shape observed in the M4 peptides is not a common characteristic of amyloid structures. However, previous studies have reported the presence of spherical structures in various forms such as the β-amyloid (amylospheroid) [202], prolactin-producing pituitary adenoma [203], [204], cytochrome c [205], and α-synuclein [206]. These reported structures have only had spherical shapes. First study reported the presence of spherical structures in various forms, such as highly toxic amylospheroids of Aß40 peptide, with dimensions in the

range of 10–15 nm [202]. Another example is a single sphere from cytochrome c, which was formed through laser trapping [205], with sizes of 4–26 μm. Additionally, spherical oligomers of α-synuclein with structures sized 2–6 nm were observed directly after dissolving. The incubation of these oligomers resulted in the observation of ring-like structures, which were characterized into two general classes, circular rings with diameters ranging from 35 to 55 nm and elliptical rings with widths between 35 and 55 nm and lengths between 65 and 130 nm [206].

The M4 mutant fibrils were observed to form dense spherical clusters with an average diameter between 185 nm and 500 nm, indicating a unique morphology distinct from previously reported structures.

The M1 fragment carries two mutations that interchange the positions of asparagine and aspartic acid residues (D17N/N22D). The mutations in the M1 fragment do not alter the overall amino acid composition of the peptide, thus the global properties remain unchanged. However, the mutations induce the charge location in the β-harpin region of the peptide. The negative charge is shifted from a highly ordered fragment of the β-arch structure into the likely unstructured C-terminal region (see Figure 15). This change in charge distribution is expected to have an impact on the ionic environment of the peptide, potentially impeding the assembly process. In contrast, in the parent sequence (ER4), the negatively-charged aspartic acid residue is located in the main chain, leading to increased potential for electrostatic interactions with cations such as Na^+ and K^+ present in the buffer. In the previous study [163], it was reported that ER4 formed rigid, unbranched fibrils, although this occurred over a longer time scale. In contrast, the M1 fragment only formed amorphous structures. The absence of negatively charged aspartic acid residues in the M1 sequence may be the reason for the observed change in the aggregation process.

In the case of M2, whose sequence is based on SE, the substitution of leucine with alanine (L15A) led to a more compact arrangement and faster aggregation (see Figure 26). This suggests that an aggregate form is more energetically favorable in comparison to the WT structure. This mutation (L15A) was designed based on the position of alanine in the ER4 sequence. The AmyloGram prediction score for this mutation was very close to the classification threshold (see Table 15). Based on all experimental studies, M2 appears to be a fibril-forming peptide.

In the case of M3, a substitution of glycine to lysine (G6K) was introduced in order to investigate the effect of adding a positive charge on the self-assembly properties of the peptide. Additionally, an asparagine was replaced with a negatively charged aspartate residue (N17D) to introduce electrostatic interactions between the lysine and the aspartic acid. An aspartate in the C-terminal region was also replaced with asparagine (D22N) to remove the charge on the terminal portion of the β-arch. As a result of these modifications, the peptide regained a positive charge on one arm of the core region of

the β-arch. The aggregation of M3 was found to be faster than that of the WT, SR4 sequence (see Figure 26). It was noted that the lag phase was approximately one hour shorter in the case of M3. It is well established that the introduction of charged amino acids in a sequence can accelerate the aggregation of amyloids [163]. This is supported by a study on phenol-soluble modulin α3 (PSMα3), which is also a functional amyloid, that highlighted the crucial role of positive charges, particularly lysine amino acid in PSMα3. This amino acid is essential for its cytotoxicity, enables its interaction with the lipid membrane, and stabilizes the structure of cross-α fibrils [207]. In this case, the M3 peptide formed unstructured assemblies. Additionally, the ThT assay showed that the aggregation pathway was altered.

In M4, glycine was replaced with lysine (G6K), which is positively charged at physiological pH. Lysine residues are known to play an important role in protein stability by forming electrostatic interactions, as they are mostly exposed on the surface of the protein [208]. Lysine residues are also known to electrostatically interact with negatively charged buffer molecules such as phosphate anions [209]. In this case, the charge was introduced on one of the arms of the β-arch structure, which forms the core of amyloid fibrils. As a result of this modification, unique snail-like structures were observed for the M4 peptide (see Figure 27). It is hypothesized that these structures may have formed due to interactions with ions. However, a more detailed examination on the atomic level is required to fully understand this process. In order to gain a deeper understanding of the structural effects of the selected mutations, the influence of Hofmeister salts on the morphologies of the studied peptides was investigated.

3.3. MD simulations of ions interactions with M3 and M4 mutants

In order to better understand the morphological diversity of the mutated peptides observed in experimental studies conducted in the presence of phosphate buffer, MD simulations were carried out. Each simulation system comprised of three M3 or M4 peptides. The simulations were performed in the presence of 100 mM phosphate ions (K_3PO_4) or 100 mM sodium chloride (NaCl). However, it was observed that at a concentration of 100 mM K_3PO_4, a crystallization process of ions occurred (data not shown), thus the concentration of phosphate ions was decreased to 50 mM.

The results of MD at the atomic scale revealed that PO_4^{3-} molecules approached the lysine amino acids in close proximity within 1 μs of the simulation, as illustrated in Figure 28B & D. The RDF analysis between the phosphate atoms of PO_4^{3-} molecules and the nitrogen atoms of lysine residues in the K_3PO_4 system revealed a prominent peak at 0.35 nm for the M3 and M4 trimers (Table 20), suggesting the formation of ion pairs, or strong electrostatic interactions between the two species. In contrast, the NaCl systems showed no significant peaks or clustering of PO_4^{3-} molecules near lysine residues (Figure 28F & H). The observed clustering of PO_4^{3-} molecules near lysine residues in the

K_3PO_4 system led to changes in the conformation of M4 trimers, while no such effect was seen in the NaCl system. This supports the hypothesis that this phenomenon may contribute to the snail-like conformation.

Phosphate ion is a strong kosmotropic agent, one from the Hofmeister anion series [210]. The stabilizing effect of phosphate anions on the native structure of proteins is well known. Phosphate ions affect noncovalent water-protein interactions and have salting-out effects, promoting aggregation processes and accelerating fibril formation, especially for proteins with intrinsic amyloidogenic propensity [73], [211], [212]. What is more, the salt ions, such as for example Na_2SO_4, Na_2HPO_4, NaCl, $NaNO_3$, $NaClO_4$, have a significant effect on the self-assembly process of proteins [72].

Based on the MD simulations, we showed that the phosphate anions, naturally present in the buffers in our experiments, interact with positively charged amino acid residues on the protein surfaces [213]. Creating a solvation shell around the peptide core, which could distinctly affect its aggregation process, may have resulted in the spherical structures. Conversely, the shielding effect was not observed for NaCl ions (Figure 28).

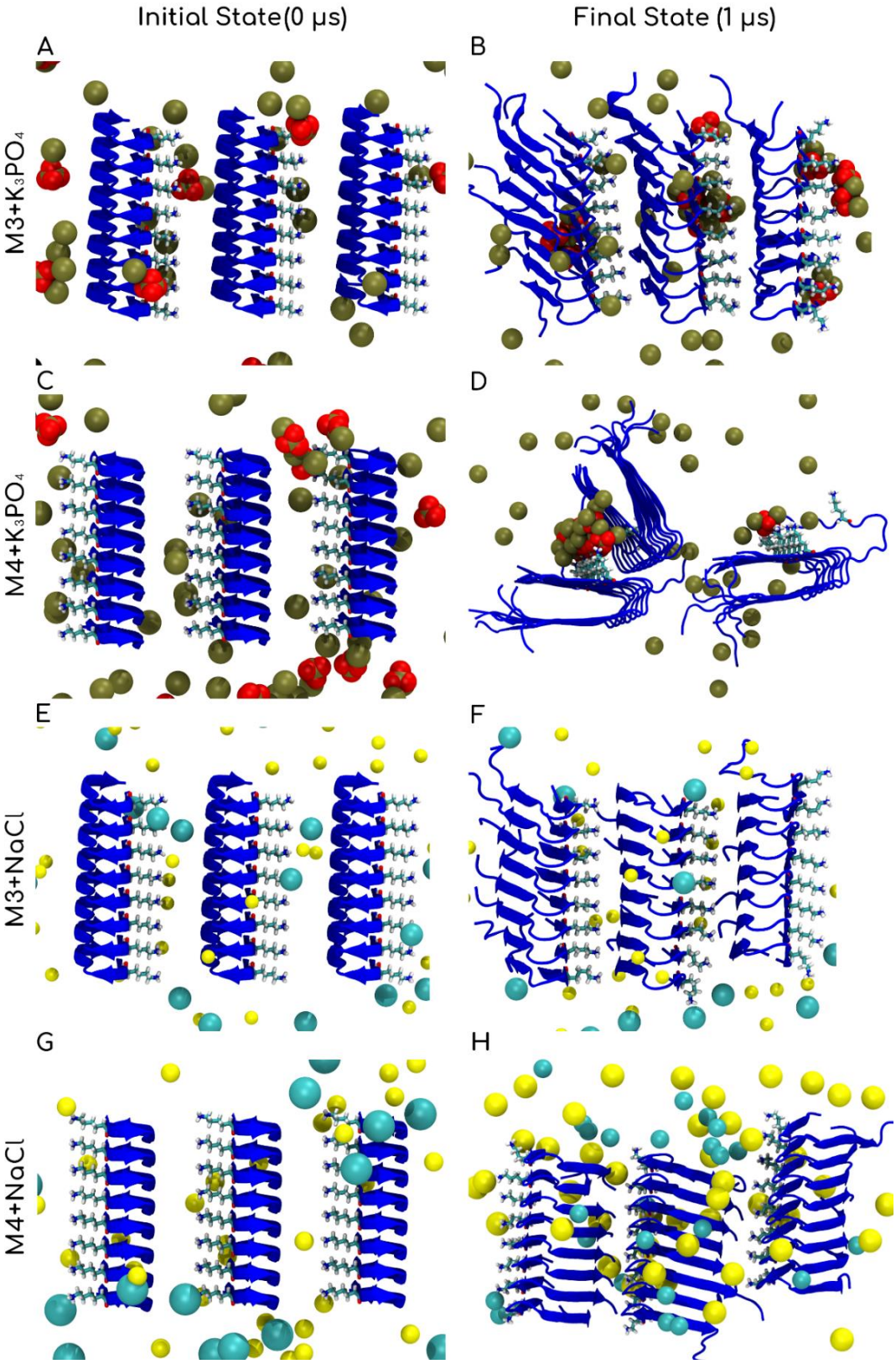


Figure 28 MD simulations of M4 and M3 trimers solvated in water with presence of phosphate ions (K_3PO_4) and sodium chloride (NaCl). The rows in the table represent the M3+K_3PO_4, M4+K_3PO_4, M3+NaCl, M4+NaCl systems, respectively. First column (A,C,E,G) shows initial state of simulation for all combinations. Final states of the simulations of all systems are shown on second column (B,D,F,H). The trimer in the cartoon representation (blue), the potassium (gold), sodium (yellow), chloride (blue) are drawn as small spheres. PO_4^{3-} molecules are in VMD representation with the oxygen in red and phosphate in gold. Lysine is drawn in the licorice representation. Water is omitted for clarity. The analyses were performed over the entire length of a corresponding trajectory (1 μs).

Based on the performed analyses of bioinformatics predictors (Table 15), it can be observed that functional amyloids used as their input data, are problematic. Even tough, functional amyloids share structural characteristic similar to pathological amyloids, however, have various stability and life-time [7], what proved also this study. This is caused, for example, by the presence of repetitive units, which are known to play significant role in them. A lot of amino acids are conserved due to evolutionary function. Generally, in amyloid research relying on homology is not advisable, especially in the pathological amyloids. However it can be observed that also in the functional amyloids homology analyses could me misleading. Because even small mutation, such as single mutations may change the aggregation tendency of the peptide.

3.4. Characterization of aggregates with scanning Kelvin probe

Furthermore, the potential of Kelvin's probe technique was evaluated for quantitative studies of amyloid fibril ordering. Figure 29A illustrates a representative image obtained using optical microscopy for a dehydrated ER4 peptide deposit on Au. Figure 29B shows the corresponding CPD surface distribution. It should be noted that the gradual change of the CPD at the rim of the deposit is attributed to the limitations of the spatial resolution of the SKP tip utilized in the experiments. Despite this, the CPD spatial distribution within the area of the peptide deposit is relatively homogeneous, with a relative standard deviation of less than 10%. Therefore, the mean CPD value calculated over the deposit area was employed to quantify the mean WF.

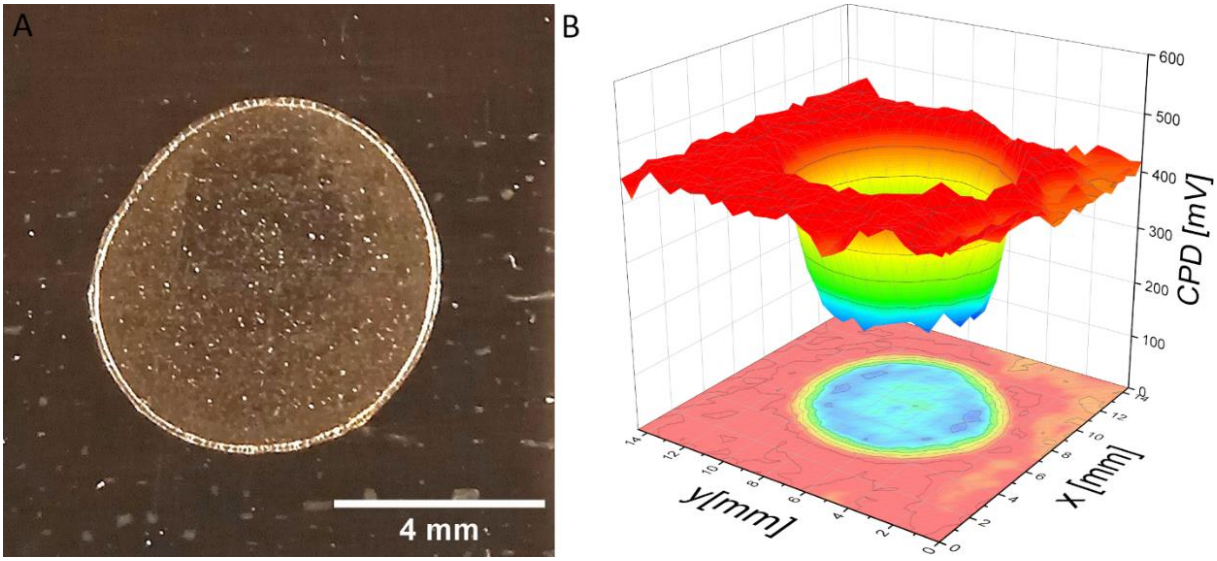


Figure 29 Exemplary dehydrated ER4 deposit on Au-coated glass slide: **A** optical microscopic image. **B** CPD surface distribution.

Figure 30A presents the mean WF values of the peptides, deposited on Au specimens. These values varied from 5.07 ± 0.35 eV (measured for M2) to 4.64 ± 0.33 eV obtained for ER4. The WF

gradually decreased in the following sequence M2→M3→M4→SR4→M1→ER4. However, it is worth noting that the difference in WF between M4 and SR4 is not statistically significant. As a result, the sequence of the results can also be arranged as M2→M3→SR4→M4→M1→ER4, reflecting the size order of peptide assemblies determined by TEM, as well as the complexity of their morphology.

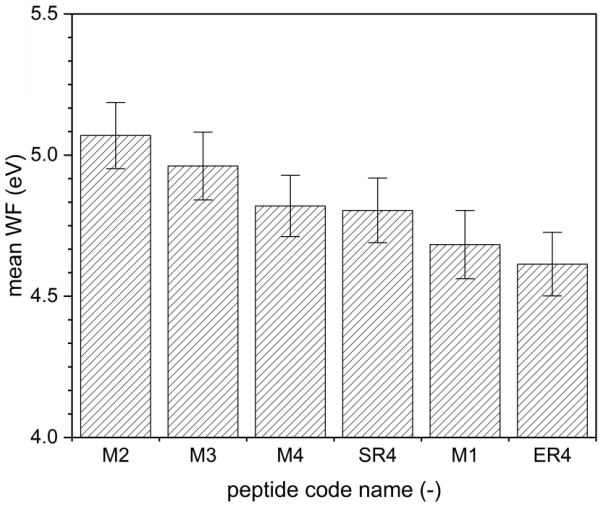


Figure 30 Mean WF with standard deviation error of dehydrated peptide deposits.

The results of this study show a strong correlation between the findings of ATR-FTIR and the WF values obtained from Kelvin's probe technique. This correlation can be seen in the varied contents of secondary structures of the aggregates, which implies the following order: M2 (26%)→M3 (13%)→SR4 (11%)→M4 (8%)→M1 (0%)→ER4 (0%). The WF values were found to be highest for fibrillar structures such as M2, SR4, and M3, followed by snail-like structures and finally amorphous structures (M1). The last position in this sequence was occupied by ER4, which had the slowest aggregation rate. Additionally, it should be noted that while no numerical value exists to correlate the results of TEM and CPD, it can be inferred that the highest WF values are associated with fibrillar structures.

In this study, the influence of charged amino acids on the essential aggregation properties of peptides was demonstrated, including aggregation rate, general structure, and morphology. Certain mutations were found to affect sub-molecular properties despite conservation of the CsgA protein motif. A new type of assembly in the form of spherically packed globular aggregates was discovered among the studied peptides, which was attributed to the effects of phosphate buffer ions on the local electrostatic interaction network of polypeptide chains. MD simulations revealed that interactions between positively charged amino acids and negatively charged phosphate moieties in the buffer contributed to the observed morphology and snail-like conformation. The addition of phosphate anions may also affect the aggregation pathways.

Amyloid material properties can be studied using various techniques, including bioinformatics tools like AmyloGram. In addition to traditional experimental methods, the Kelvin probe was tested for its ability to study amyloids, and the results showed that CPD obtained with the probe can provide quantitative information on the molecular characteristics of amyloid aggregates, which correlates well with other methods. Therefore, the Kelvin probe can be utilized for characterizing amyloid-based nanomaterials.

Chapter 4 Interactions of Aβ42 and hIAPP

Thanks to all experimental experience gained during previous studies of amyloidogenic peptides, the ability to conduct a study on pathological amyloids in more complex model systems could be considered. The focus of the carried out investigation was on the Aβ42 and hIAPP peptides. To mimic a more physiological environment, experiments and simulations were conducted in the presence of a lipid membrane as an extrinsic factor. The peptides were first studied separately, then together in a subsequent phase, to address somehow the cross-seeding phenomenon. Whenever possible, the results of the MD simulations were validated with experimental AFM analysis.

Regarding my 3rd hypothesis: the *presence of a lipid membrane affects aggregation of the native Aβ42 peptide*, this examination was aimed to check the stability of Aβ42 peptide in the lipid membrane and its tendency to self-assemble in the lipid environment. Afterwards, we considered amylin in various environments. Amylin consists of 37 amino acid sequences and is a hormone peptide. The expression of amylin, which is a part of 89 amino acid preprotein (22 amino acid signal peptide and two flanking peptides), occurs at the 12th chromosome. The signal peptide is cleaved in the endoplasmic reticulum (ER) and finally the conversion from proIAPP to IAPP takes place in the secretory vesicles [214]. Afterwards, hIAPP is stored in the secretory vesicles together with insulin [193]. However, the amylin aggregates have been found not only in post-mortem brains and pancreatic cells of patients, but also in the cerebrospinal fluid (CSF) [215], [216]. Several studies reported the coexistence of both Aβ and hIAPP in blood serum and CSF at nanomolar concentrations [217]. Additionally, it was shown that APP protein, tau and hIAPP, were altogether found in pancreas cells, by using RT-PCR and Western blot. Aβ was also colocalized with amylin in islet amyloid deposits [218]. Andreetto et al. investigated interactions between short hot spot regions of Aβ40/42 and hIAPP, and showed that the binding affinities between them were within the nanomolar to low micromolar range [217]. The cross-seeding can occur, due to high sequence similarity, around 50% and 25% of identity. However, the question in which environment favors the interaction of amylin with Aβ42 remains unsolved. Herein, the final conformations from selected simulations of the single component systems were merged to study cross-interaction between Aβ42 and hIAPP in the presence of the lipid membrane, in order to test my 4th hypothesis: *the presence of another peptide/protein impacts the general fibrilization process.*

4.1. Is a single Aβ42 peptide stable in the lipid bilayer?

In this section we consider Aβ42 peptide embedded in the lipid environment, especially its extracellular part and its secondary structure changes within the time. A novel aspect of proposed mechanism of possible cross-talk between Aβ and amylin, starts from the prerequisite knowledge of

the Aβ peptide location. The main hypothesis of the amyloid cascade is schematically drawn in Figure 6. It assumes that the "toxic aggregates" are formed by assembly of the Aβ peptides that left the lipid membrane, following the γ-secretase cleavage of APP. The mechanism how Aβ peptide leaves the neuronal membrane and enters the extracellular space remains elusive. As a result, it raised the question what is the driving force in the proteolysis chain, which triggers this process.

Here, the hosting team in Nancy considers a novel hypothesis of aggregation of Aβ42 [219], that the oligomerization process occurs in the lipid membrane, is proposed (see Figure 31). This hypothesis is an alternative to the currently studied models of AD patho-mechanisms.

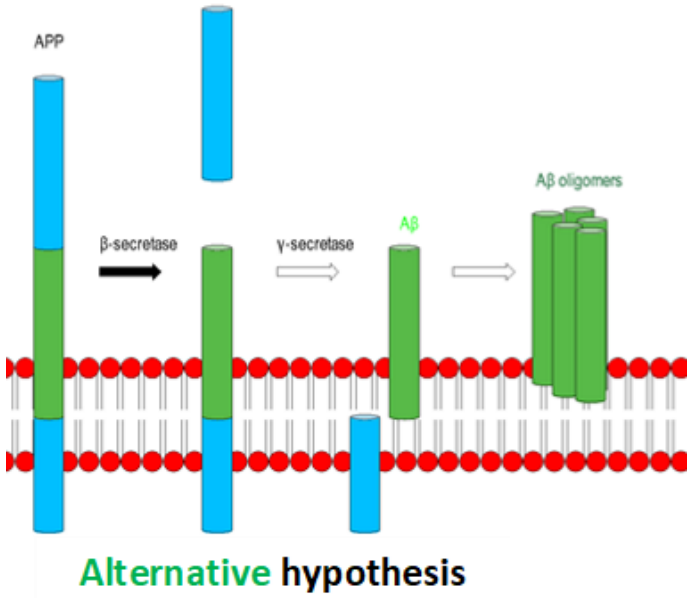


Figure 31 Alternative hypothesis: nascent Aβ peptide remaining inserted in the membrane and Aβ peptide oligomerization occurring in the membrane.

Results obtained from MD simulations carried out in this work showed that Aβ42 peptide embedded in DOPC is stable during time scales exceeding 3.75 μs (see Figure 32).

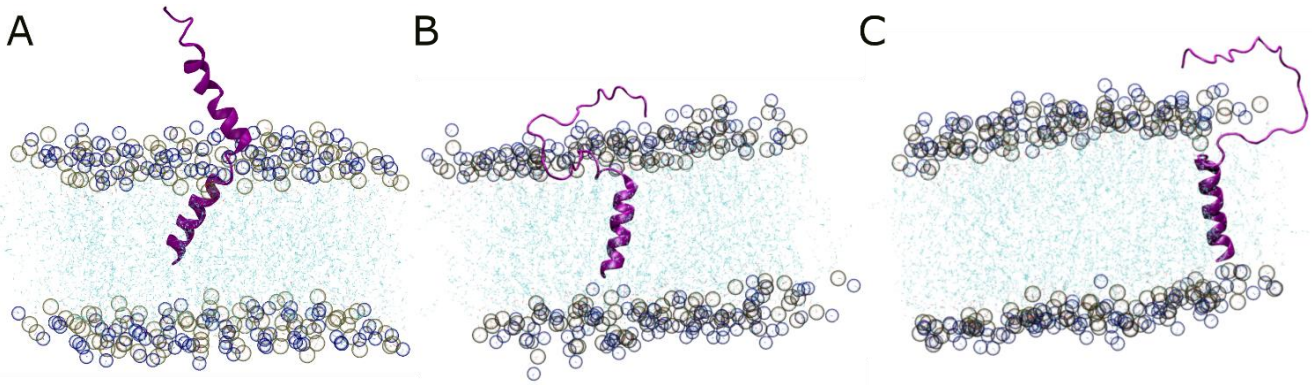


Figure 32 MD simulation of an Aβ42 peptide in a DOPC bilayer, at 0 **(A)**, 2 **(B)**, and **(C)** 3.75 μs. The lipid head groups phosphorus P (gold) and nitrogen N (blue) are drawn as small spheres. The acyl chains of the lipids are in blue, lines representation. Water is omitted for clarity.

Though after ~700 ns, the extracellular part (1–27 amino acid long) of Aβ42, loses its helical structure, the intramembrane part remains in its helical form embedded within the membrane. The secondary structures changes during the whole simulation time are showed on Figure 33.

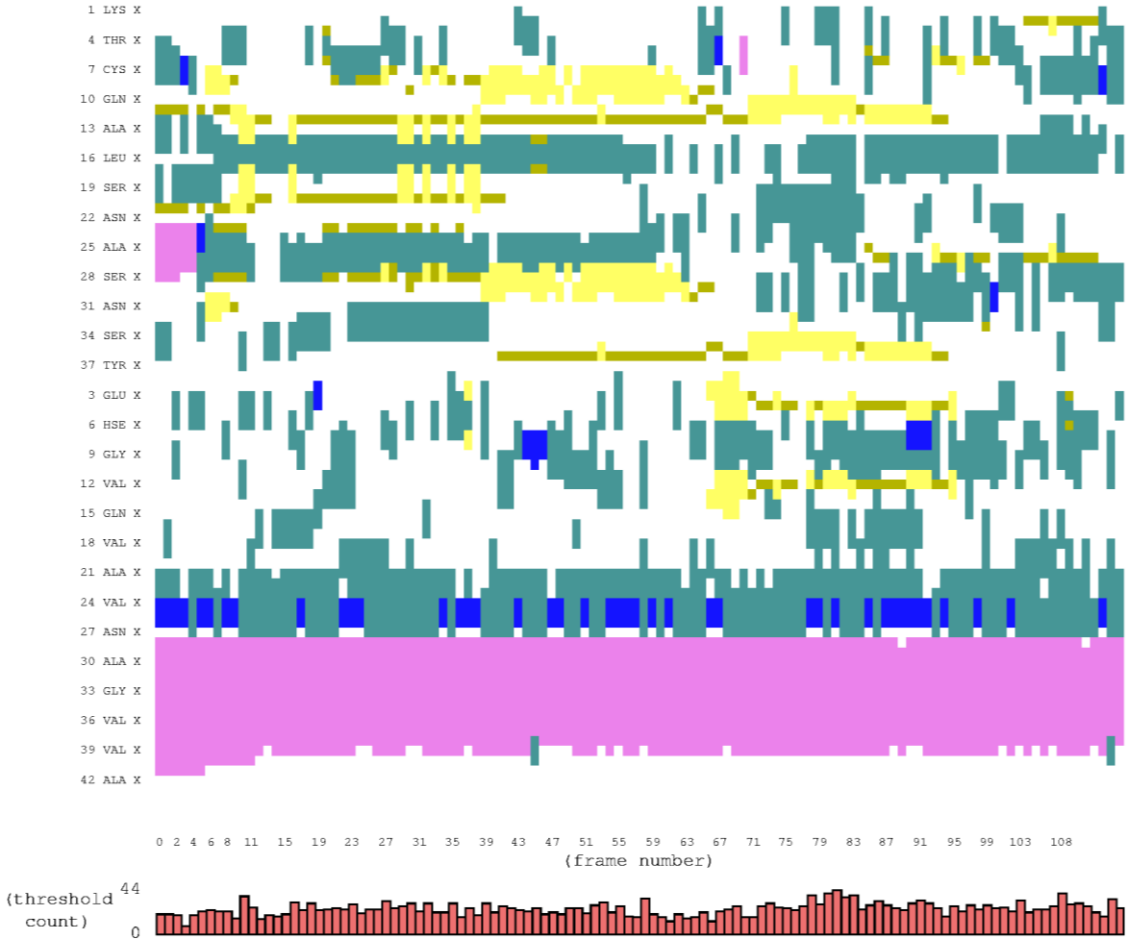


Figure 33 Plot displaying of Aβ42 secondary structure trajectory in DOPC. Where colors denote to secondary structure: green – turn, yellow – extended configuration, olive – isolated bridge, pink – α-helix, blue – 3_{10} helix, red – π-helix, white – random coil.

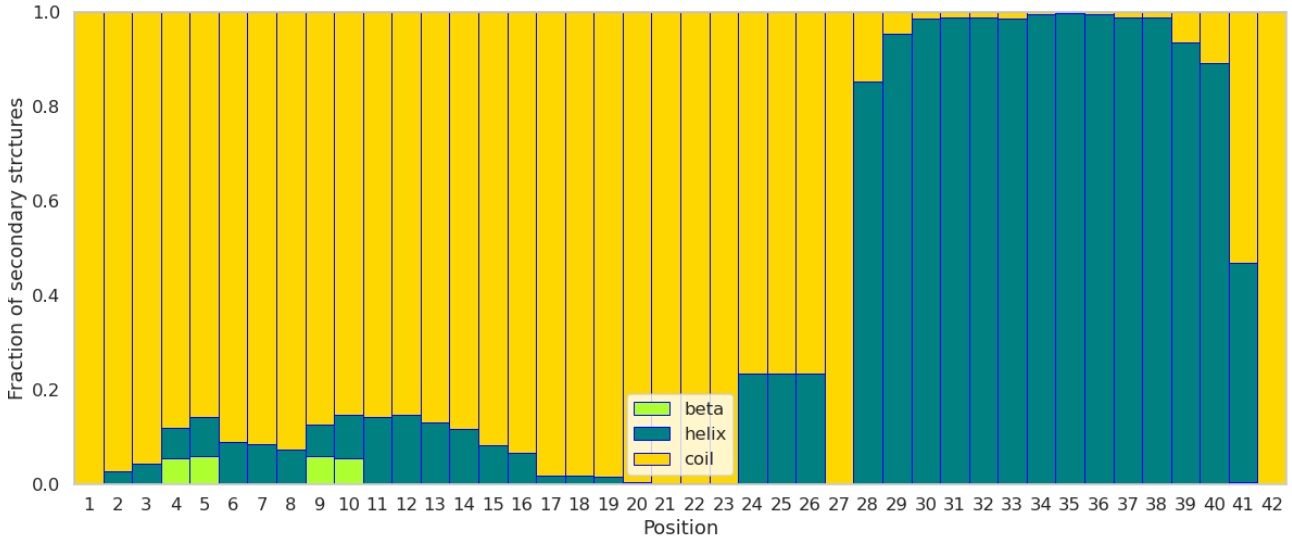


Figure 34 Plot of the fraction of the secondary structure of Aβ42 in each position of amino acid sequence during the simulation.

In Figure 34, we can observe that the transmembrane part (residues 32 to 41) possesses the dominant helical structure of Aβ42 (about 100%). While, the extracellular part (1–27) has rather random coil conformation. The KLVFFAE region is pointed out as the most essential for the self-assembly of Aβ peptide, called also as a core mutation region [220]. This mutation region (16–22 amino acid), did not exhibit high order (β-sheet) conformation, during 3.75 μs simulation time. These simulation results provide therefore valuable insights to consider not only for the study of the self-aggregation of the peptide, but also in cross-seeding, which involves the interaction of an approaching hIAPP peptide. The described results of MD simulations partially proved 3rd hypothesis, that Aβ42 is stable in the DOPC lipid membrane. As a result, we can infer that it is possible that the peptide oligomerization could start in lipid environment (see Figure 35). The proposed structure of the oligomers according to our modeling consists of the helical transmembrane part that holds a similar fold and oligomerization as, that of Aβ42 obtained from the solution NMR [221].

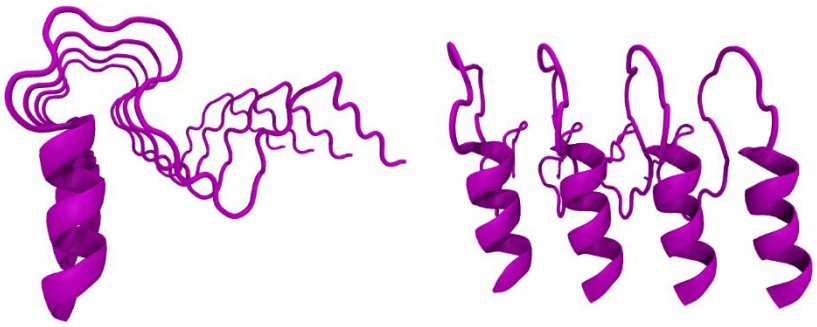


Figure 35 The proposed oligomerization hypothesis, the helical part inserted in the lipid membrane the S–shaped fibrillar part above the lipid head groups.

4.2. Amylin stability in different environments

The accessible conformation of native amylin originates from NMR structure, in which hIAPP was solvated in an SDS micelle [193]. This considers the presence of a heterogeneous media including a hydrophobic phase (SDS surfactant tails) and a hydrophilic one (SDS heads group) in contact with a solvent. Based on that, we carried out simulations in solution and in various conformations/position of amylin with respect to the lipid membrane: the hIAPP peptide was placed either on top of or within a DOPC bilayer in the cases studied. Additionally, the influence of ions was tested for the first instance especially NaCl versus $CaCl_2$ in concentration 0.15 M.

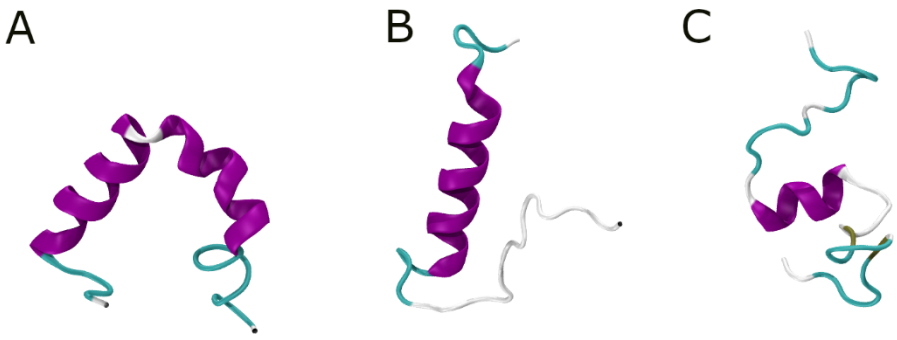


Figure 36 MD simulation of native amylin peptide in a 0.15 M NaCl solution: evolution of the fold as function of time **(A)** 0, **(B)** 300 ns, and **(C)** 8.12 µs. Water and ions are omitted for clarity.

The MD simulation of the peptide in solution showed that the kinked conformation of amylin was quickly lost after removing position constrains, leading to a rather straight helical conformation (Figure 36B), that was stable for ~ 300 ns. Later, the N-terminal of the peptide lost its helicity to form a random coil conformation. Hence the results indicate that the amylin conformation is rather disordered in solution (Figure 36C). It has also been reported by others [222], that amylin in solution lacks well-defined secondary and tertiary structures.

The N and C terminal regions of the amylin peptide transition from a helical to a disordered structure. Interestingly, after 8.12 µs of simulation, the external fragment of Aβ42 bares similarities in both sequence and structure to the kinked hIAPP peptide (reported in Figure 36). The latest publication of Cao et al. showed in agreement with our finding that hIAPP 19–29 and Aβ 24–34 have 50% of the similarity in the structure [223].

Note however, that the sequence comprising residues 13 to 19 had a helical conformation throughout the whole simulation (Figure 37 & Figure 38). These amino acid positions are close to the core mutation region of amylin, which is at residues 20 to 29 (SNNFGAILSS). Interestingly, region 14 to 20 is often linked to the cytotoxicity of hIAPP [224]. Additionally, Jaikaren et al. [225] pointed out that residues 8 to 20 and 30 to 37 are essential for amylin's aggregation process. The last fragment forms fibers in aqueous solutions [225].

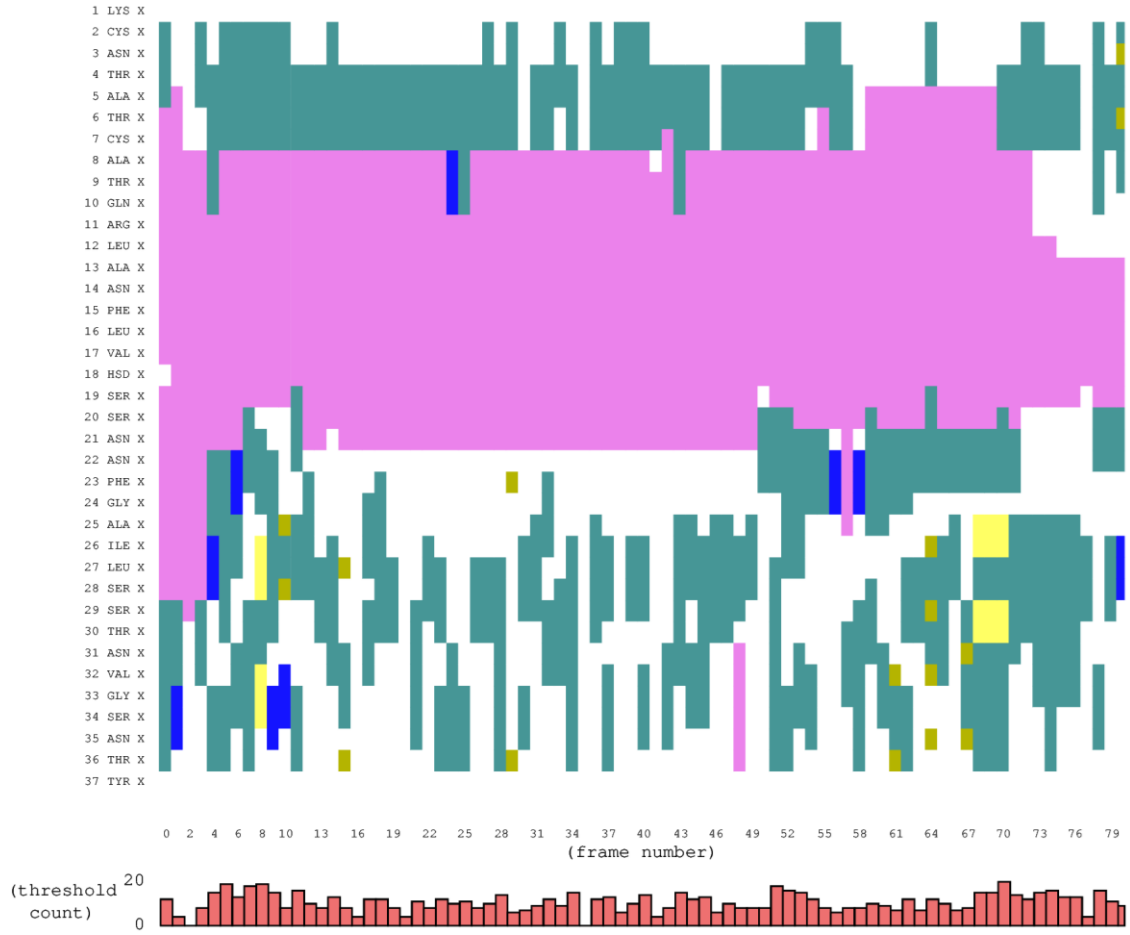


Figure 37 A plot depicting the changes in secondary structure of amylin during simulation in solution. Where colors denote to secondary structure: green – turn, yellow – extended configuration, olive – isolated bridge, pink – α-helix, blue – 3_{10} helix, red – π-helix, white – random coil.

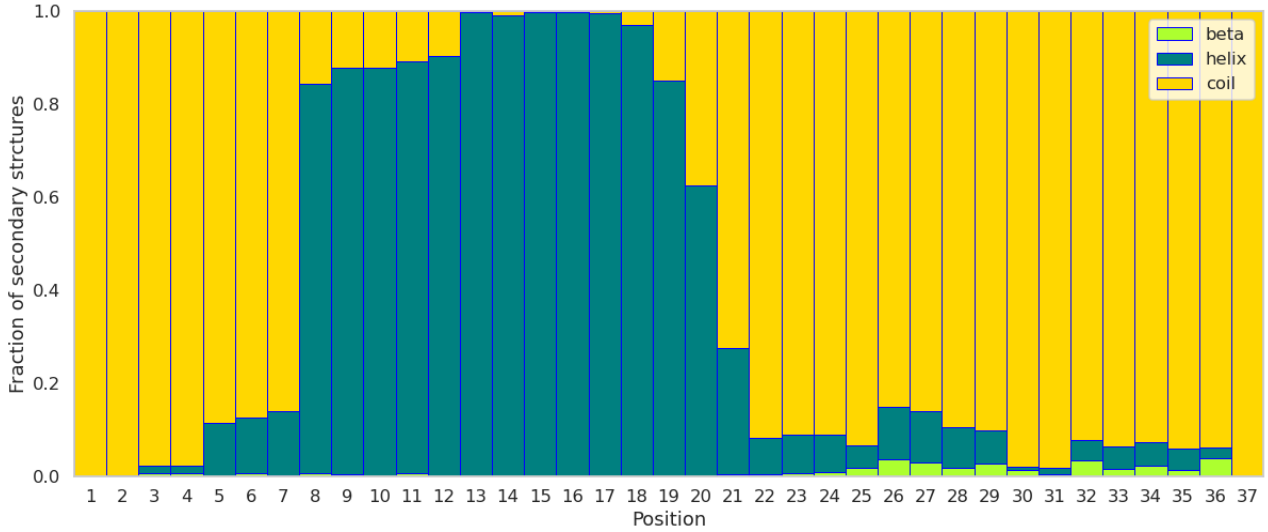


Figure 38 Plot of the fraction of the secondary structure of amylin in each position of amino acid sequence during the simulation.

We carried out a second set of simulations to analyze the potential interaction of the peptide with a membrane. The scenarios evaluated included placement of the hIAPP peptide embedded within the bilayer and above a DOPC bilayer with two types of ions (NaCl and $CaCl_2$). Firstly, the amylin was inserted in DOPC lipid membrane, the simulations were conducted with the presence of NaCl ions at 0.15M concentration. This environmental surrounding was similar to the NMR state of native amylin in SDS micelle [193]. Based on the obtained configuration we can observe that amylin lost it helicity at both N and C terminus (see Figure 39).

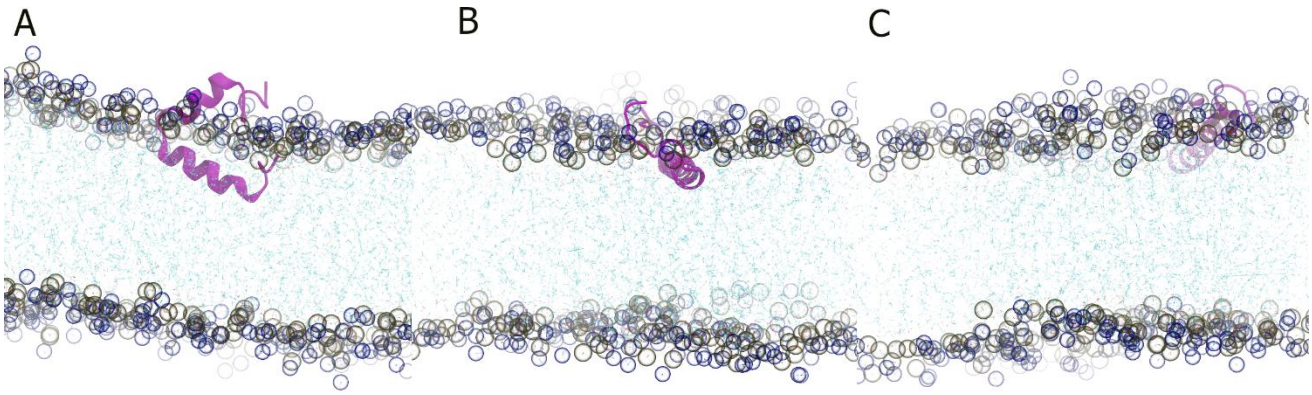


Figure 39 MD simulation of native amylin embedded in a DOPC lipid bilayer, at **(A)** 0, **(B)** 1, and **(C)** 5.1 μs. The lipid head groups phosphorus P (gold), and nitrogen N (blue) are drawn as small spheres. The acyl chains of the lipids are in blue, lines representation. Water is omitted for clarity.

However, in comparison to the behavior in solution (13 to 19), in the DOPC bilayer, helical structure was preserved between residues 8 and 20 (see Figure 36). This indicates that the peptide is more stable in an ordered fold in the lipid environment than in solution. Furthermore, after 200 ns, its conformation was changed from kinked to a straight helical. This study indicated that what is most stabilizing conformation of the peptide in its helical form is the instance where it interacts with a membrane lipid head groups.

Based on the analysis of the position of the peptide with respect to the lipid bilayer, we observed that amylin remained at the interfacial phase of the DOPC solution, undergoing fluctuations within the hydrophobic environment and close to the positively-charged choline moiety of the headgroups (see Figure 41). Compared to Aβ42, embedded amylin was found to be less stable in the lipid membrane.

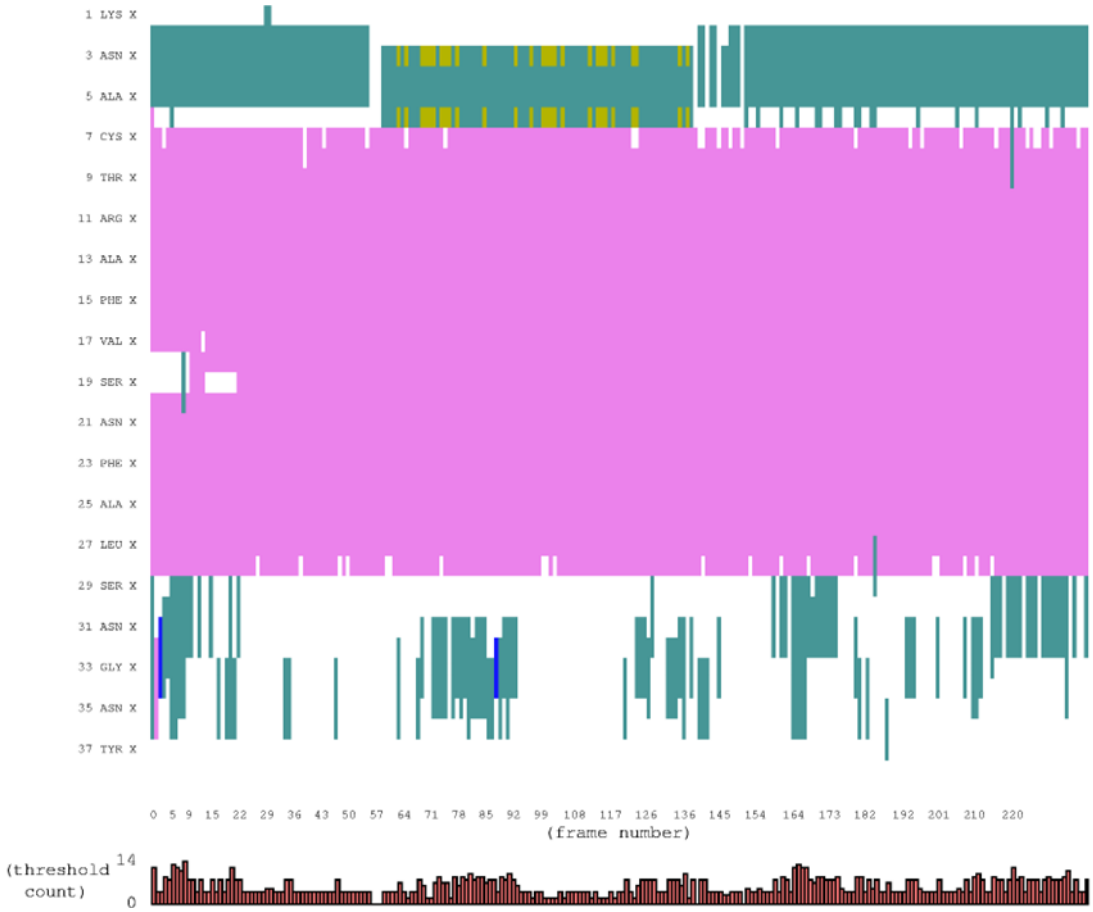


Figure 40 A plot displaying the changes in secondary structure during the simulation of amylin embedded within a DOPC lipid bilayer. Where colors denote to secondary structure: green – turn, yellow – extended configuration, olive – isolated bridge, pink – α-helix, blue – 3_{10} helix, red – π-helix, white – random coil.

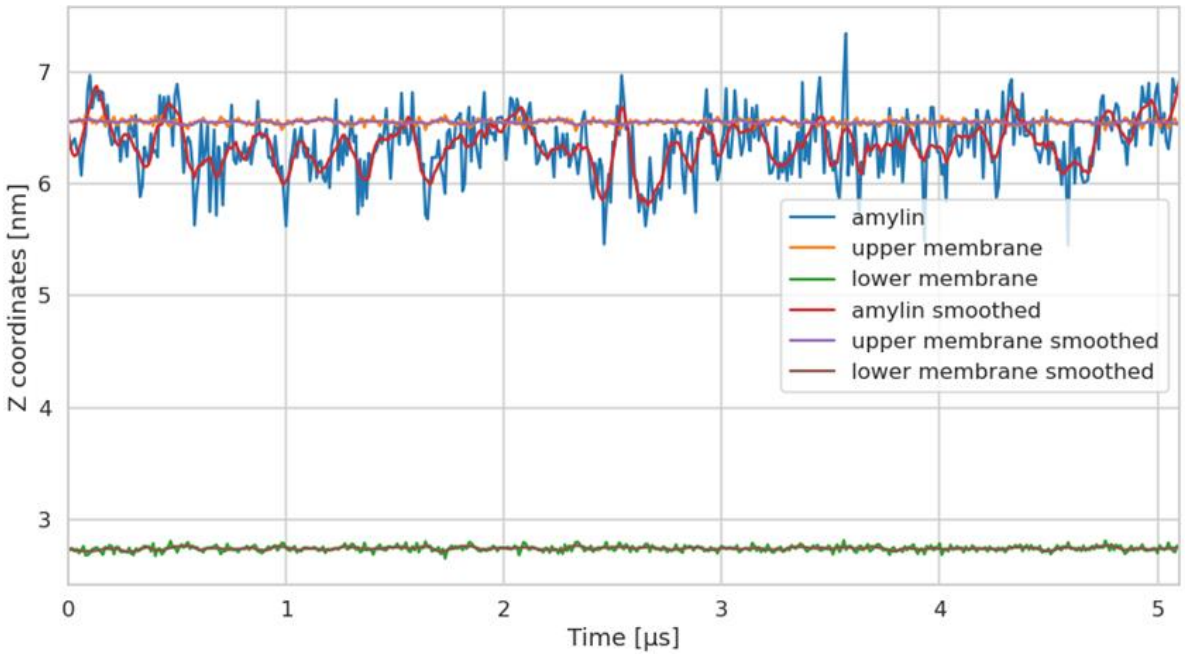


Figure 41 Values of Z coordinates of amylin, lower and upper membrane, in the time of simulation. Graph shows the real and smoothed values.

Based on the analysis of the secondary structure fraction during a 5.1 μs simulation period, a higher proportion of β-sheet structures were observed in the lipid environment compared to solution (see Figure 42). This suggests that conformational changes occur more rapidly in the lipid environment due to the effects of crowding and interaction with the lipid membrane.

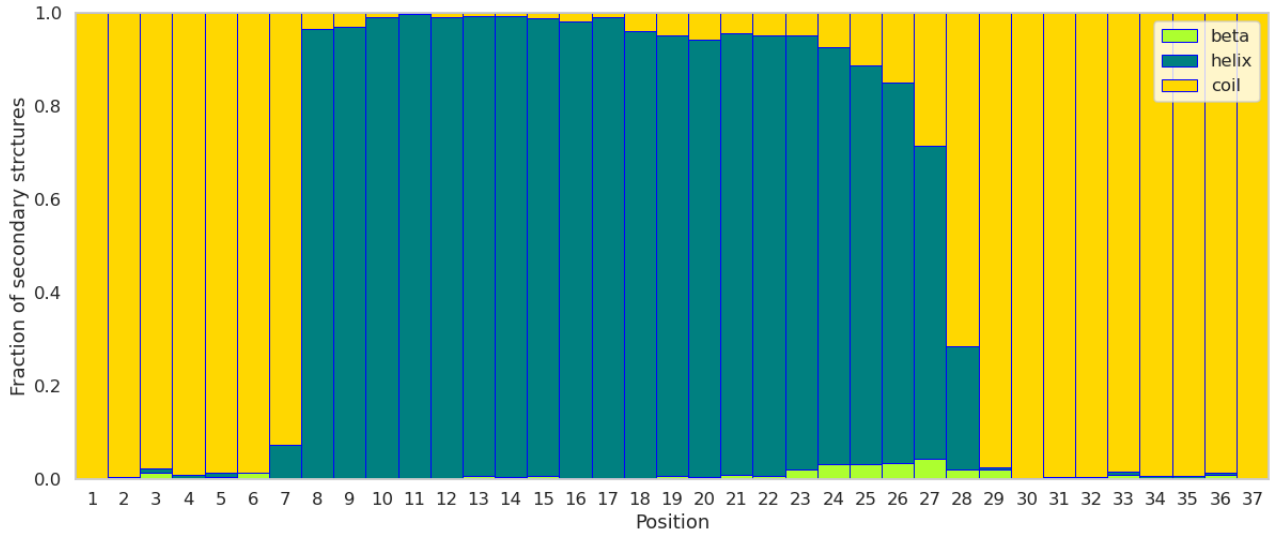


Figure 42 Plot of the fraction of the secondary structure of amylin embedded in a DOPC, in each position of amino acid sequence during the simulation.

We studied furthermore the behavior of amylin when placed just above a DOPC lipid membrane (see Figure 43). After approximately 100 ns, amylin that was initially placed above the lipid bilayer, entered into it and became stabilized in the interfacial phase of the DOPC lipid bilayer. Similar to the previous system, the conformation of amylin changed from a kinked to a straight helical shape, as illustrated in Figure 43B.

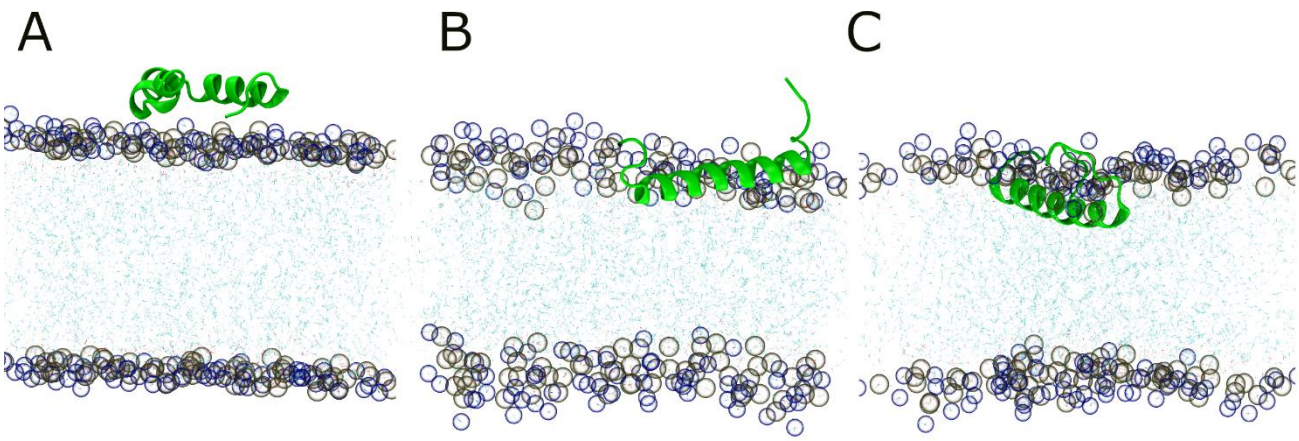


Figure 43 MD simulation of native amylin above the DOPC lipid bilayer with NaCl ions, at **(A)** 0, **(B)** 2 μs, and **(C)** 5.24 μs. The lipid head groups phosphorus P (gold) and nitrogen N (blue) are drawn as small spheres. The acyl chains of the lipids are in blue, lines representation. Water is omitted for clarity.

Figure 44 reveals that amylin is located in the interfacial phase near the lipid head groups. It shows greater stability in the hydrophobic (fatty acid) tail compared to its stability in solution. Similarly to the embeded amylin system, the straight amylin is located near the hydrophilic (phosphate) head.

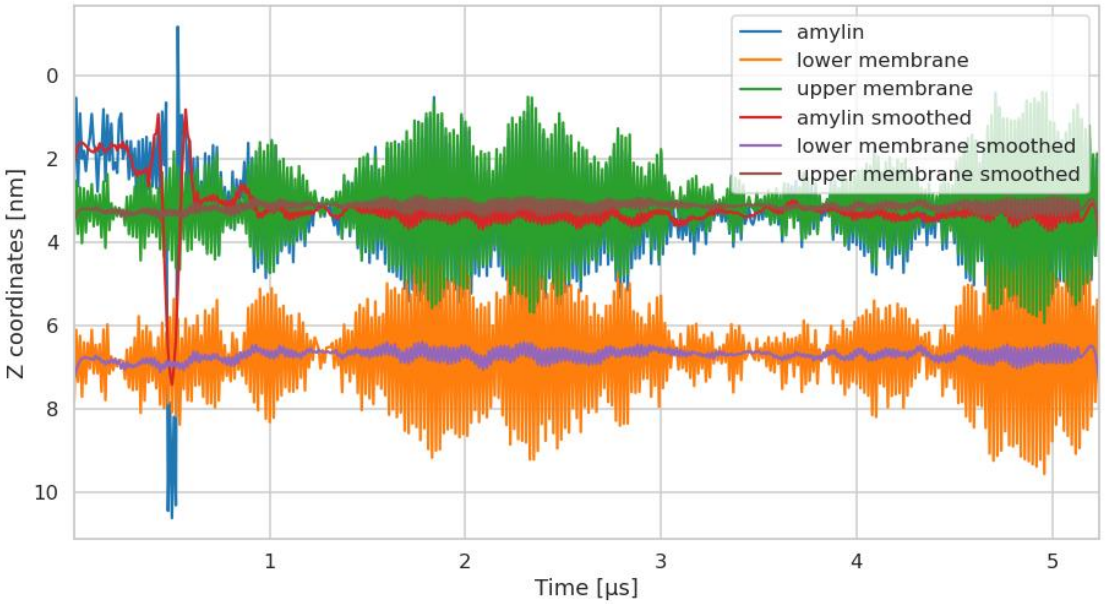


Figure 44 Values of Z coordinates of amylin placed above a DOPC lipid bilayer with NaCl ions, lower and upper membrane in the time of simulation. Graph shows the real and smoothed values.

The secondary structure fraction plot of amylin located above the DOPC, shown in Figure 45, reveals a predominant helical conformation (residues 8 to 25). The presence of amylin at the bilayer lipid/solvent interface can be attributed to the interaction of the positively charged histidine, which favors this peptides' location.

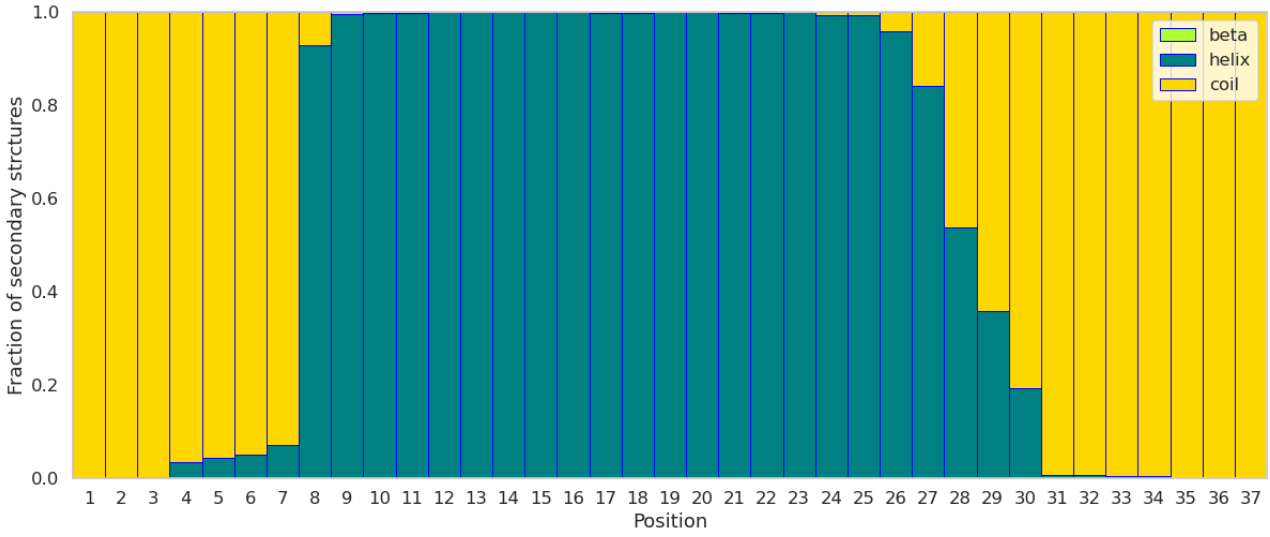


Figure 45 Plot of the fraction of the secondary structure of amylin placed above a DOPC lipid bilayer with NaCl ions, in each position of amino acid sequence during the simulation.

Additional MD simulations were performed with 0.15 M $CaCl_2$ solution and amylin was positioned just above the DOPC bilayer (Figure 46A), as a complement to the concurrent experimental study carried out in $CaCl_2$ buffer. Compared to NaCl ions, amylin appeared to is inserted deeper in the lipid bilayer (see Figure 47).

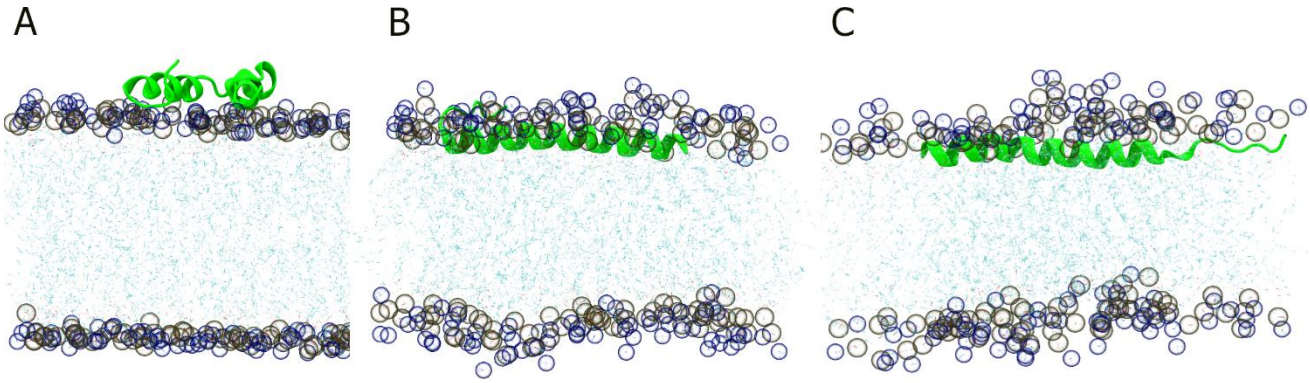


Figure 46 MD simulation of native amylin above the DOPC lipid bilayer with $CaCl_2$ ions, at **(A)** 0, **(B)** 4, and **(C)** 8.52 µs. The lipid head groups phosphorus P (gold) and nitrogen N (blue) are drawn as small spheres. The acyl chains of the lipids are in blue, lines representation. Water is omitted for clarity.

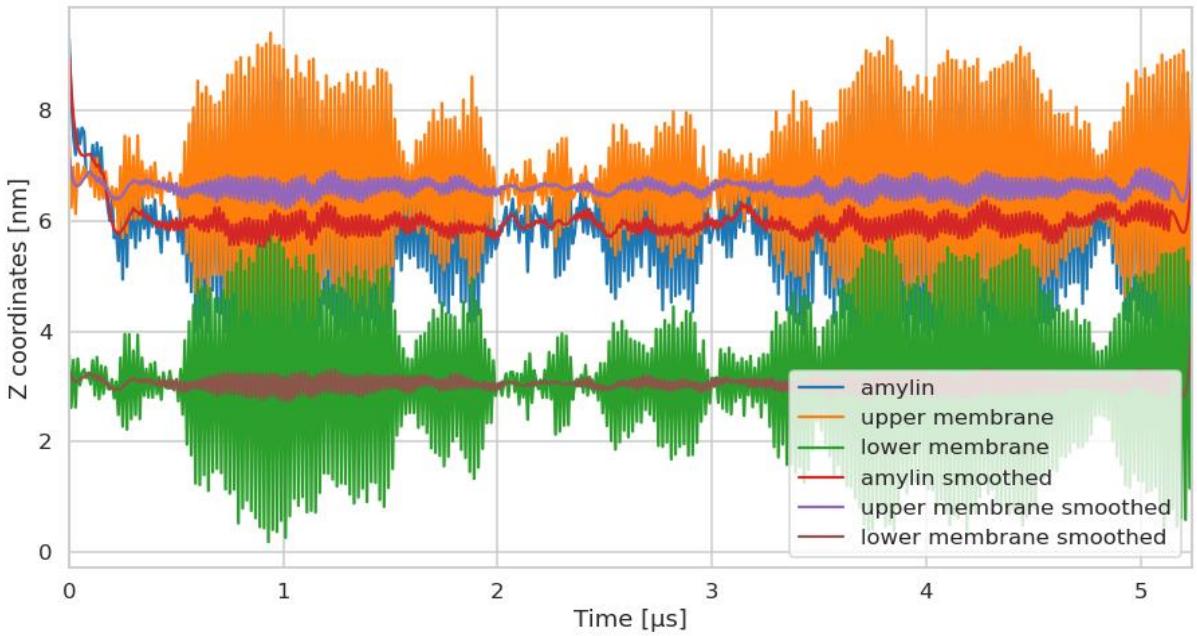


Figure 47 Values of Z coordinates of amylin placed above a DOPC lipid bilayer with $CaCl_2$ ions, lower and upper membrane in the time of simulation. Graph shows the real and smoothed values.

Based on the secondary fraction analysis, the helical confirmation of amylin was conserved for kept residues 8 to 30. Only the few residues from N and C terminus adopted random coil conformations. Within 300 ns, amylin penetrated deeper into lipid`s hydrophobic tails in case of $CaCl_2$ compared to NaCl buffer while it remained for microseconds at the interfacial region closer to the solution (see Figure 43).

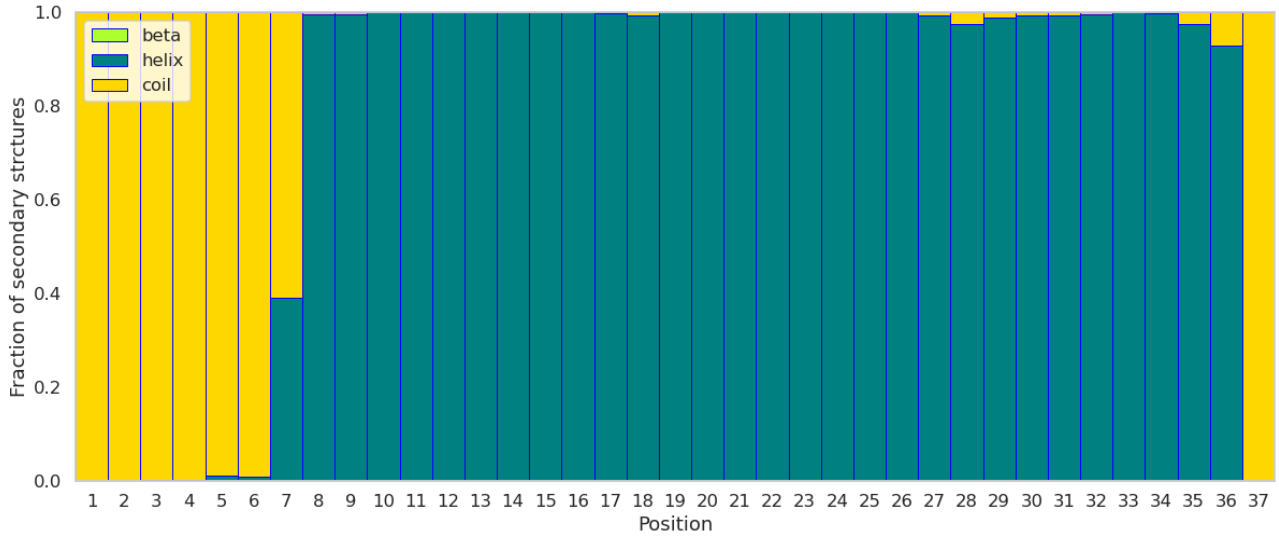


Figure 48 Plot of the fraction of the secondary structure of amylin placed above a DOPC lipid bilayer with $CaCl_2$ ions, in each position of amino acid sequence during the simulation.

Altogether, the data obtained from the simulations carried out, provide important clues for considering potential scenarios for the study of cross-interactions between Aβ42 and hIAPP.

The final conformations derived from the simulations performed were based on the simulation of Aβ42 embedded within DOPC with a disordered extracellular part, and amylin in the solution. The decision to consider these conformations was driven by the assumption that cross-interactions in the lipid bilayer may be restricted. One potential scenario is the type of interaction of Aβ42 with hIAPP in the CSF [217].

4.3. Cross-interaction Aβ42 and amylin

First, we performed an extensive examination of sequence and structural similarities between the two peptides. The native forms of Aβ42 and amylin were compared with their fibrillar structures. The regions between Aβ40/42, which potentially could interact with hIAPP, are: Aβ(19–22), Aβ (27–32) and Aβ(35–40) (see Figure 49) [217]. What is more, it is proved that Aβ(27–32) and Aβ(35–40/42) are core regions of Aβ-tau hetero-assembly [226]. The regions in the extracellular part, considering the proposed here, novel aspect of amyloid cascade hypothesis, are more possible to interact with amylin.

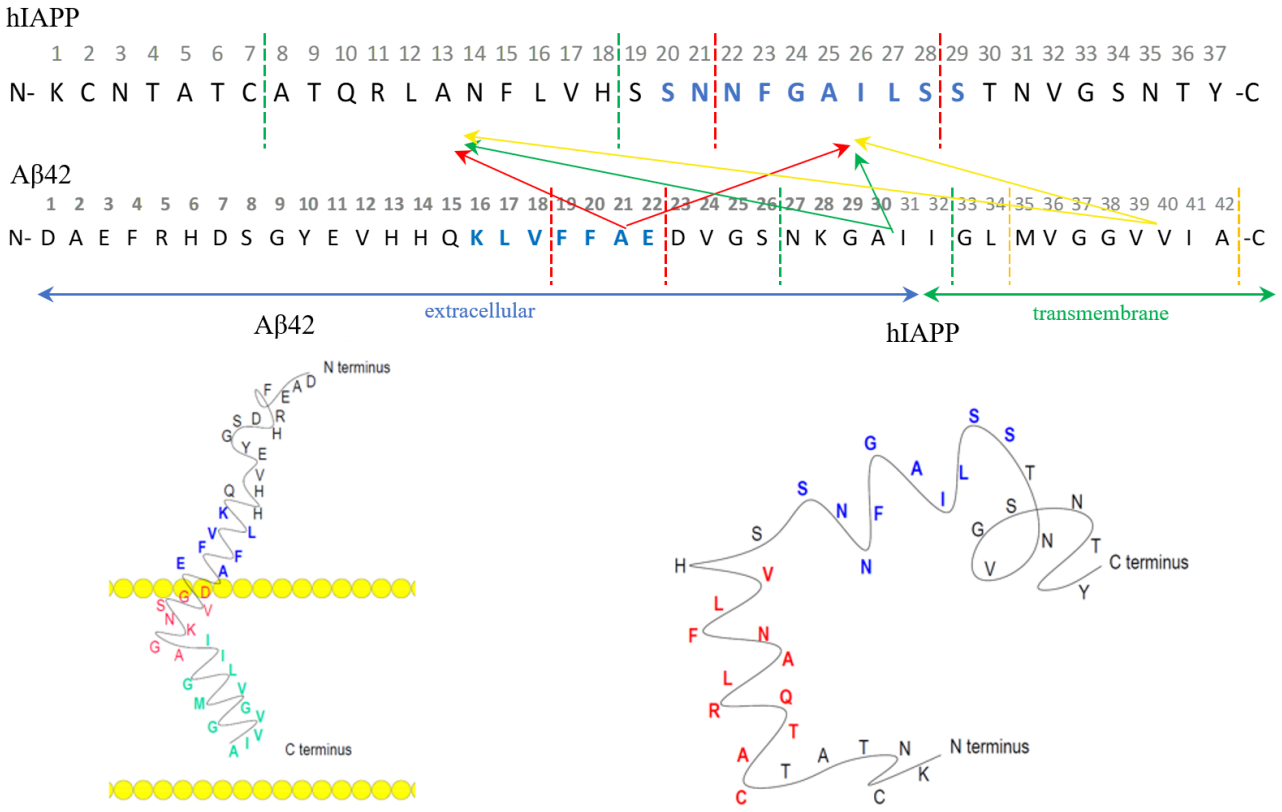


Figure 49 Possible interactions between Aβ42 peptide and hIAPP, where: the core mutation region is blue, the transmembrane part of Aβ42 – green, the amino acid in hIAPP strongly interacting with micelles – red (based on: [217]). Model native structures of Aβ42 peptide and hIAPP.

We found that high sequence similarity being one of the factors needed for cross-interaction is confirmed. However, we still wonder which structures should be considered in this process, the fibrillar or native ones. If we look at the structural similarities of fibrillar forms of these peptides, they possess a typical S-shape fibrillar pattern (see Figure 50), characteristic for amyloid fibrils [227].

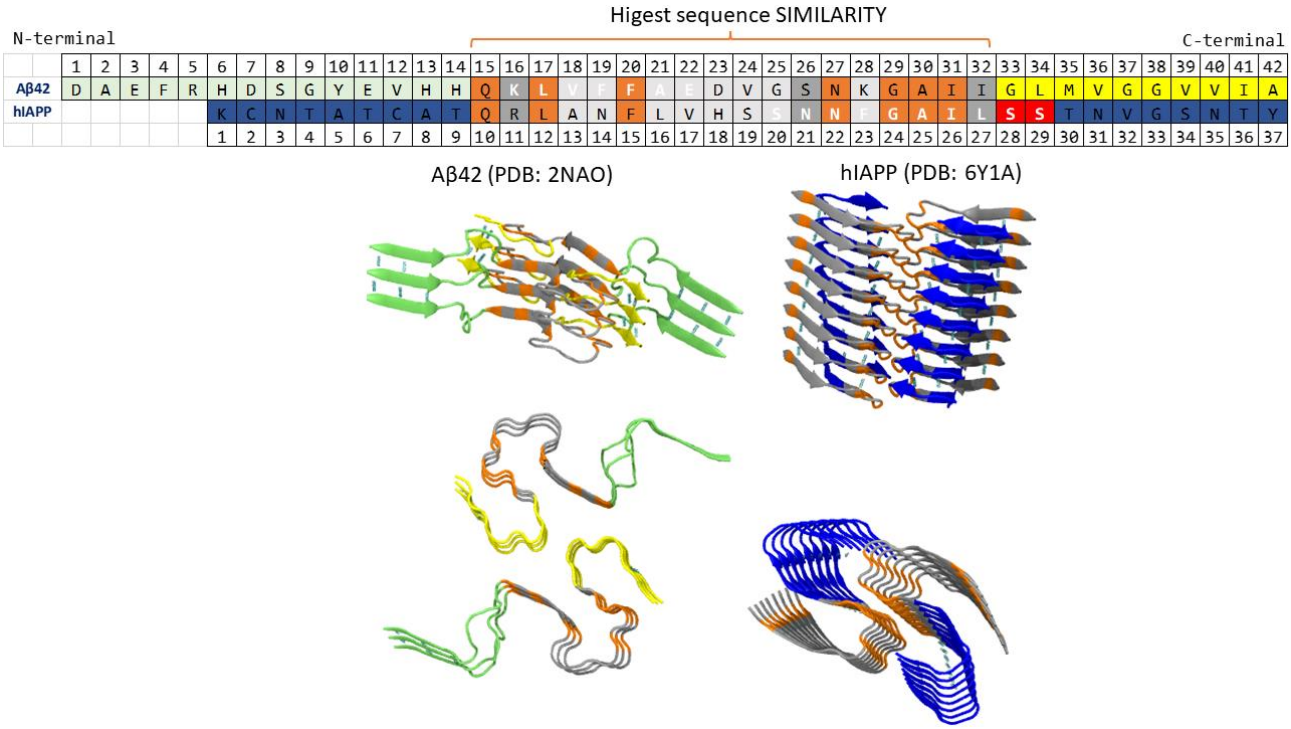


Figure 50 Sequence pair alignment of Aβ42 and hIAPP peptide, where: region with the highest sequence similarity is in grey-orange color, the transmembrane part of Aβ42 – yellow, the extracellular part of Aβ42 – green, the amino acid in hIAPP strongly interacting with micelles – blue. The solution NMR structure of Aβ42 dimers and the cryo-EM structures of hIAPP dimers. Colored according to the above description. Frontal and top view.

In the case of amylin, the fibril structure was obtained through the use of cryo-EM, which was performed at pH 6.0 [194]. Aβ42 structure was acquired by means of solution NMR at pH 7.4 [220].

MD simulations (> 5.41 μs) were carried out to gain further insight into the interaction of Aβ42 with amylin in the presence of lipid membrane. Initially, we used the CHARMM36 force field. Figure 51 shows that amylin changes conformation to adopt β-sheet folds while interacting with the extra cellular segment of Aβ42. We recall that amylin when considered alone and placed above the DOPC lipid bilayer entered the interfacial hydrophobic part of the membrane and remained helical. In simulation presented herein which studies the cross-interactions. Amylin did fluctuate near the lipid head groups (see Figure 52), yet the presence of an embedded Aβ42 changed its behavior.

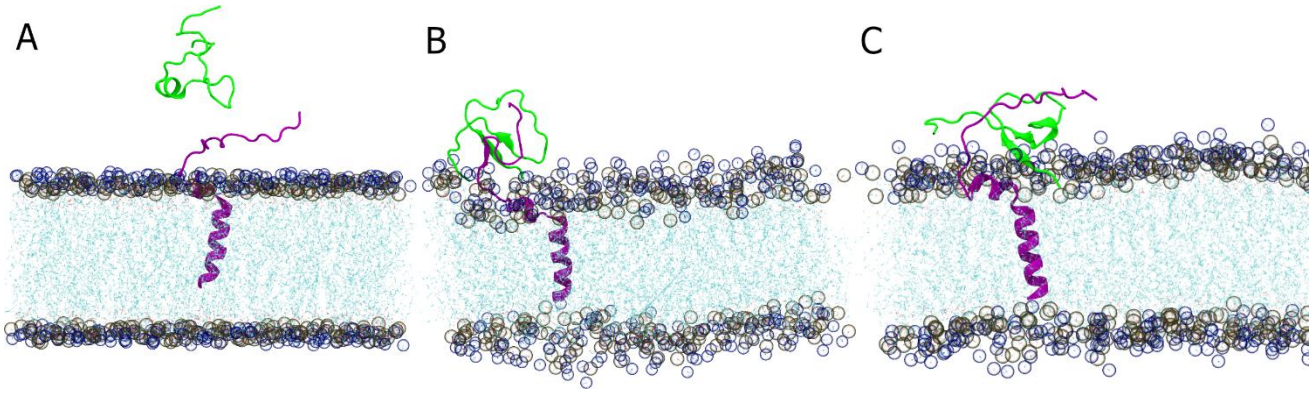


Figure 51 MD simulation of amylin (green) and Aβ42 (purple) in the presence of DOPC lipid bilayer, at **(A)** 0, **(B)** 2.5, and **(C)** 5.41 μs (CHARMM-36). The lipid head groups phosphorus P (gold) and nitrogen N (blue) are drawn as small spheres. The acyl chains of the lipids are in blue, lines representation. Water is omitted for clarity.

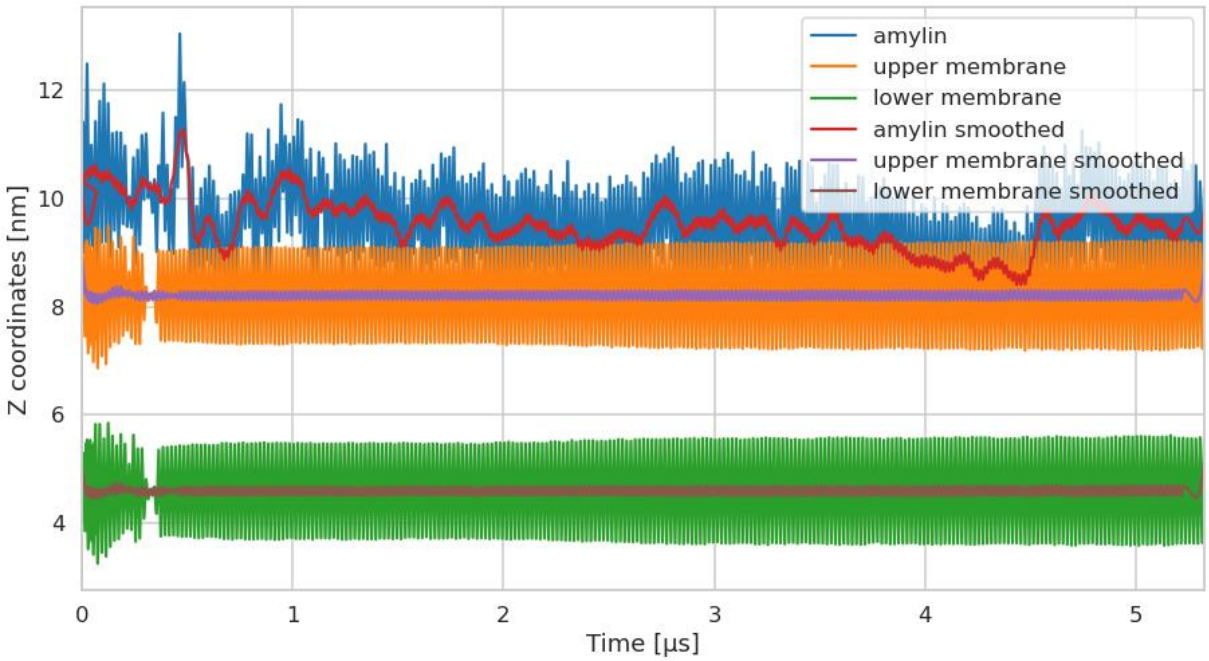


Figure 52 Values of Z coordinates of amylin placed above a DOPC lipid bilayer in the cross-interaction system, lower and upper membrane in the time of simulation (CHARMM-36). Graph shows the real and smoothed values.

The results of fraction of secondary structure changes showed that transmembrane part of Aβ42 possessed helical conformation, and its extracellular part had random coil characteristics with a minor proportion of β-sheets (see Figure 53A). On the other hand, the amylin lost it helicity and possessed random coil conformation with a fair high proportion of β-sheets at 0.6% (see Figure 53B), especially at these positions: 25 to 29. Those residues are a part of core mutation region, and the β-sheet conformation is characteristic for amyloid proteins.

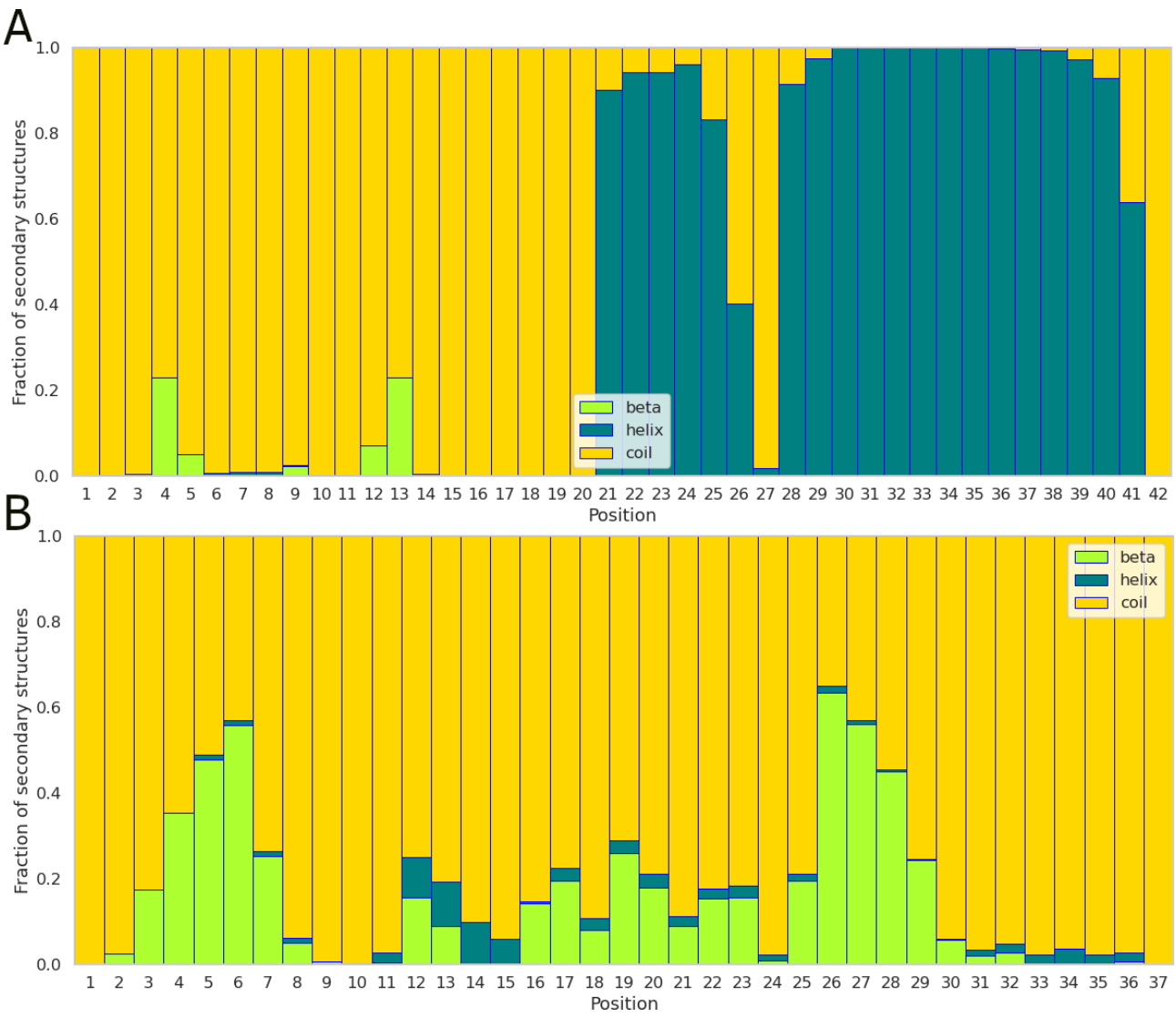


Figure 53 Plot of the fraction of the secondary structure of Aβ42 (**A**) and amylin placed above (**B**) a DOPC lipid bilayer in the cross-interaction system (CHARMM-36), in each position of amino acid sequence during the simulation.

Based on the contact maps, which take into account distances between alpha-carbons in amylin and Aβ42, the shortest distances were assessed for about 30 Å. The minimal distance was obtained for residues 11–26 of Aβ42 and 21–31 residues of amylin. This area contains the core mutation and highest sequence similarity region. Without this region, peptide loses its aggregation propensity. The alpha-carbon distance between amylin and Aβ42 is too large for them to form a hetero-aggregate. To form hydrogen bonds, the distance needs to be around 6 Å.

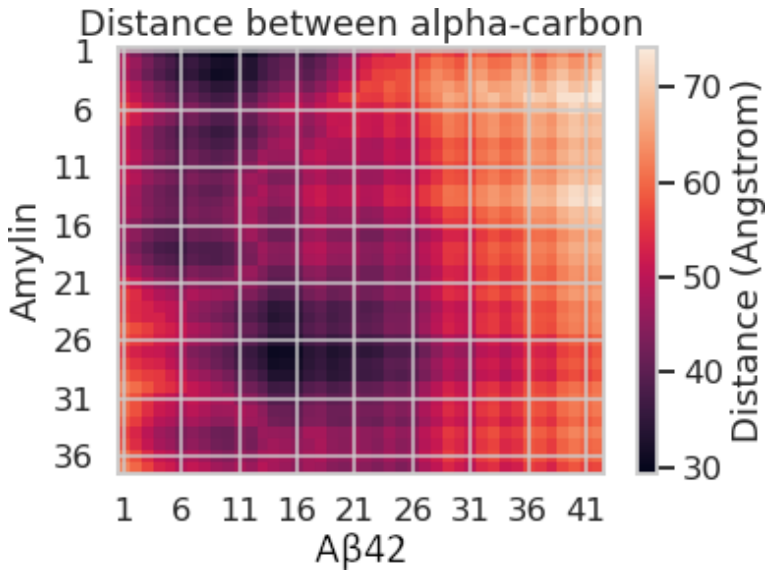


Figure 54 Amylin and Aβ42 contact map in simulation using CHARMM-36 force field.

The performed simulation of cross-interaction between Aβ42 and amylin in CHARMM-36 force field, proved that the amylin`s tendency to form β-sheet conformation is higher in the presence of Aβ42 peptide. In contrast, in isolated amylin or Aβ42, formation of β-sheet structures is less probable.

We further modelled the same systems, using the CHARMM-36m force field, designed to examine specifically disordered proteins. This force field has indeed been recently highly recommended to study amyloid proteins [228]. Figure 55, reports the results of this investigation. Amylin adopted indeed a β-sheet structure and was periodically in close proximity to extracellular part of Aβ42.

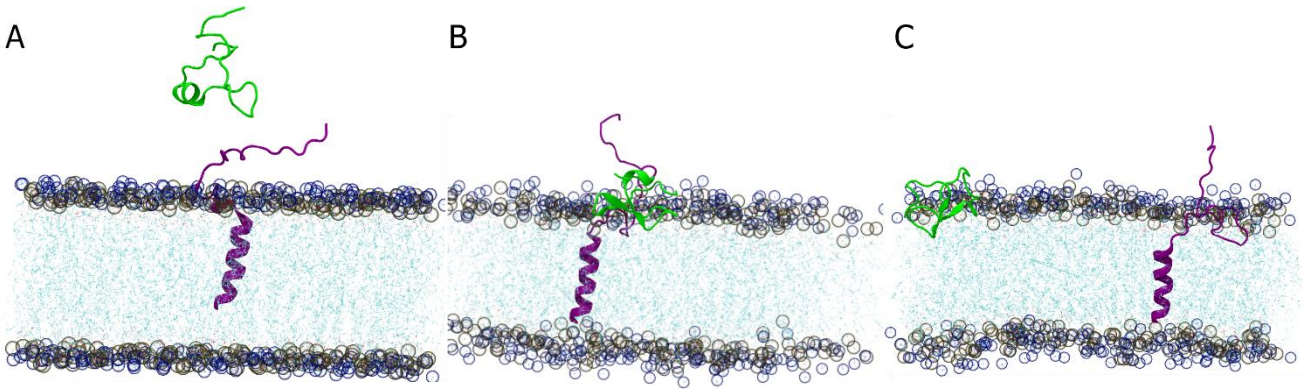


Figure 55 MD simulation of amylin (green) and Aβ42 (purple) in the presence of DOPC lipid bilayer, at **(A)** 0, **(B)** 1.5, and **(C)** 3.59 μs (CHARMM-36m). The lipid head groups phosphorus P (gold) and nitrogen N (blue) are drawn as small spheres. The acyl chains of the lipids are in blue, lines representation. Water is omitted for clarity.

Amylin was partially inside lipid membrane until 800 ns (Figure 56), then it fluctuated on lipid-water interface. Incorporation of Aβ42 resulted in a modification of this behavior.

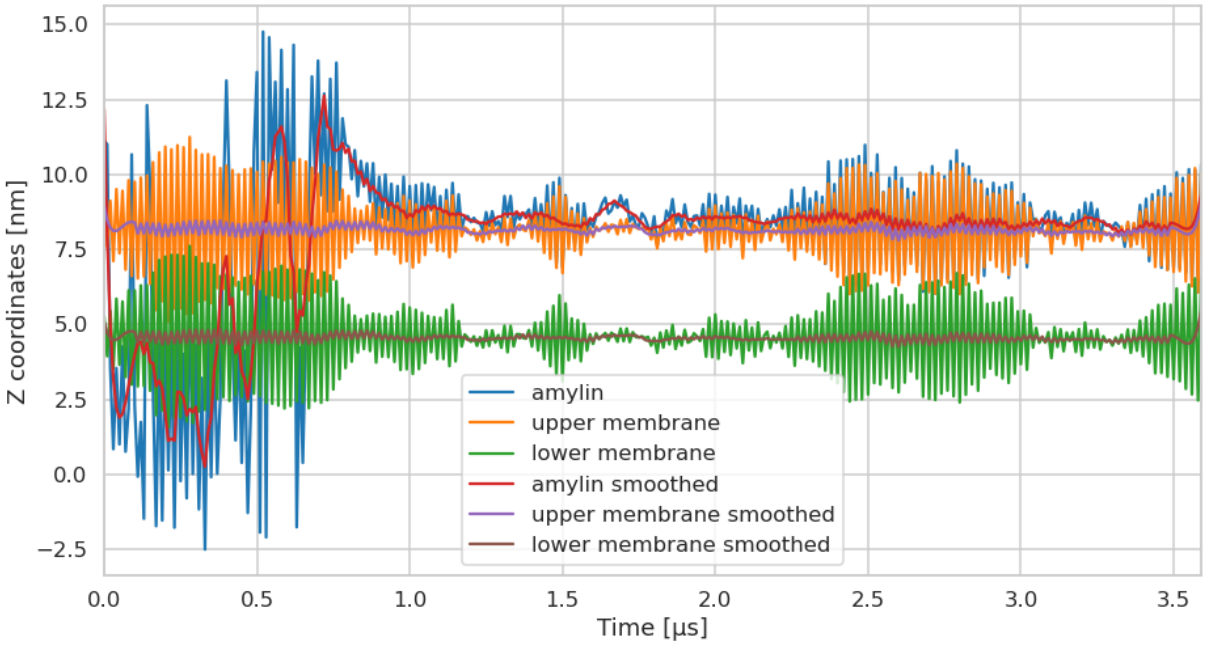


Figure 56 Values of Z coordinates of amylin placed above a DOPC lipid bilayer in the cross-interaction system, lower and upper membrane in the time of simulation (CHARMM-36m). Graph shows the real and smoothed values.

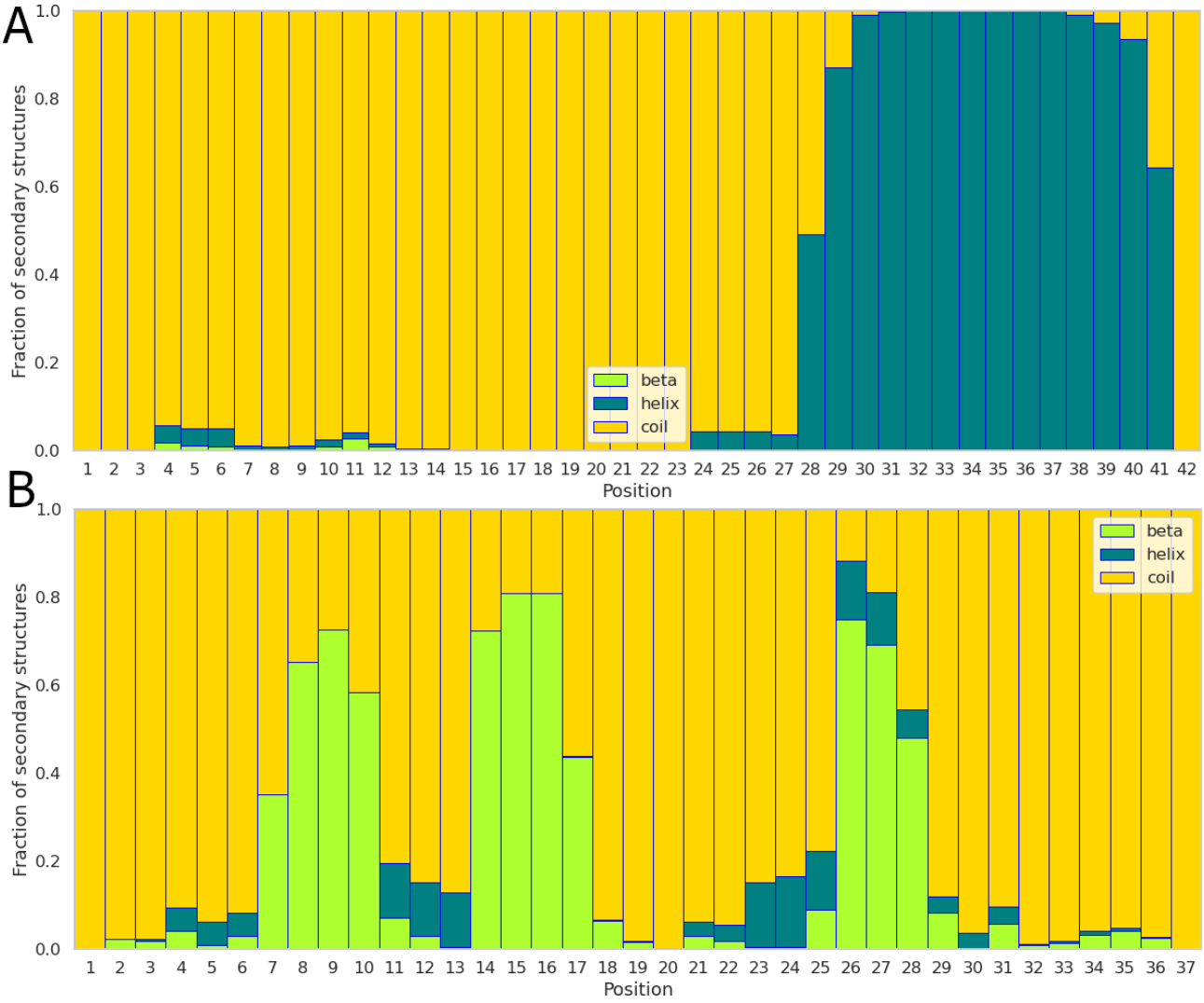


Figure 57 Plot of the fraction of the secondary structure of Aβ42 (**A**) and amylin placed above (**B**) a DOPC lipid bilayer in the cross-interaction system (CHARMM-36m), in each position of amino acid sequence during the simulation.

Based on the contact map, which shows distances between alpha-carbons in amylin and Aβ42, the shortest distances were about 20 Å (see Figure 58). The minimal distance was obtained for residues 6–11 of Aβ42 and 11–16 residues of amylin. This region includes the extracellular part of Aβ42 and the region with highest sequence similarity between amylin and Aβ42 (see Figure 50). The results of this investigation indicated that in the CHARMM-36m force field, the alpha-carbon atoms of amylin and Aβ42 were found to be in proximity to each other to a greater extent compared to the CHARMM-36 force field. The mean separation between β-sheet structures in amyloid fibrils is quantified to be approximately 10 Å .

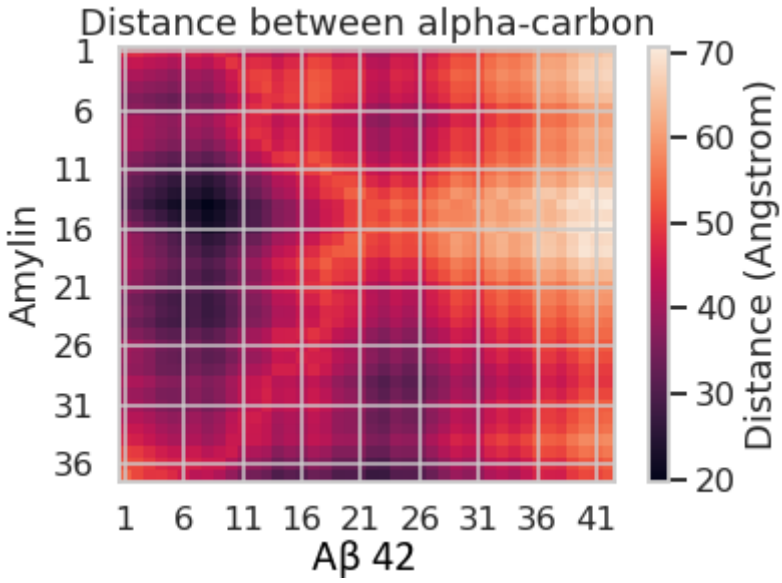


Figure 58 Amylin and Aβ42 contact map in simulation using CHARMM-36m force field.

To conclude, both MD studies of cross-interaction between Aβ42 and amylin tend to support the 4th hypothesis stated in this thesis: *the presence of another peptide/protein impacts the general fibrilization process.* The analyses of MD simulations results demonstrated that the integration of native Aβ42 into a lipid bilayer environment enhances amylin's tendency towards forming β-sheet structures, compared to previous evaluations of both peptides individually.

Hence, the presence of Aβ42 may be treated as environmental and physiological factor that contributes to a peptide's/protein's tendency to aggregate.

4.4. Experimental analyses

The proposed 3rd and 4th hypotheses were additionally validated experimentally by AFM. First, we investigated the interaction of Aβ42 with simple POPC/DPPC phospholipid model membrane. The results obtained from MD simulations showed that Aβ42 is stable within the membrane suggesting that its oligomerization process may occur in the membrane. This process could lead to formation of amyloid fibrils and lead to a destabilization or destruction of the lipid membrane. Evidently, due to difficulties in sampling the long-time scales of conformational transitions in amyloids, MD investigations are of limited use. Accordingly, we sought experimental validation of the hypotheses these simulations led us to consider. We have in particular used the AFM technique as a reporter of morphological changes of supported lipid bilayers (SLBs), consisting of POPC (liquid phase) and DOPC (gel phase) lipids (see Figure 59), in which Aβ42 peptides were embedded according to the protocol detailed in section 1.11. Overall our results indicated that the 20 µM of Aβ42, incorporated into lipid membrane, is sufficient to cause significant damage to the membrane's structure and organization (see Table 17).

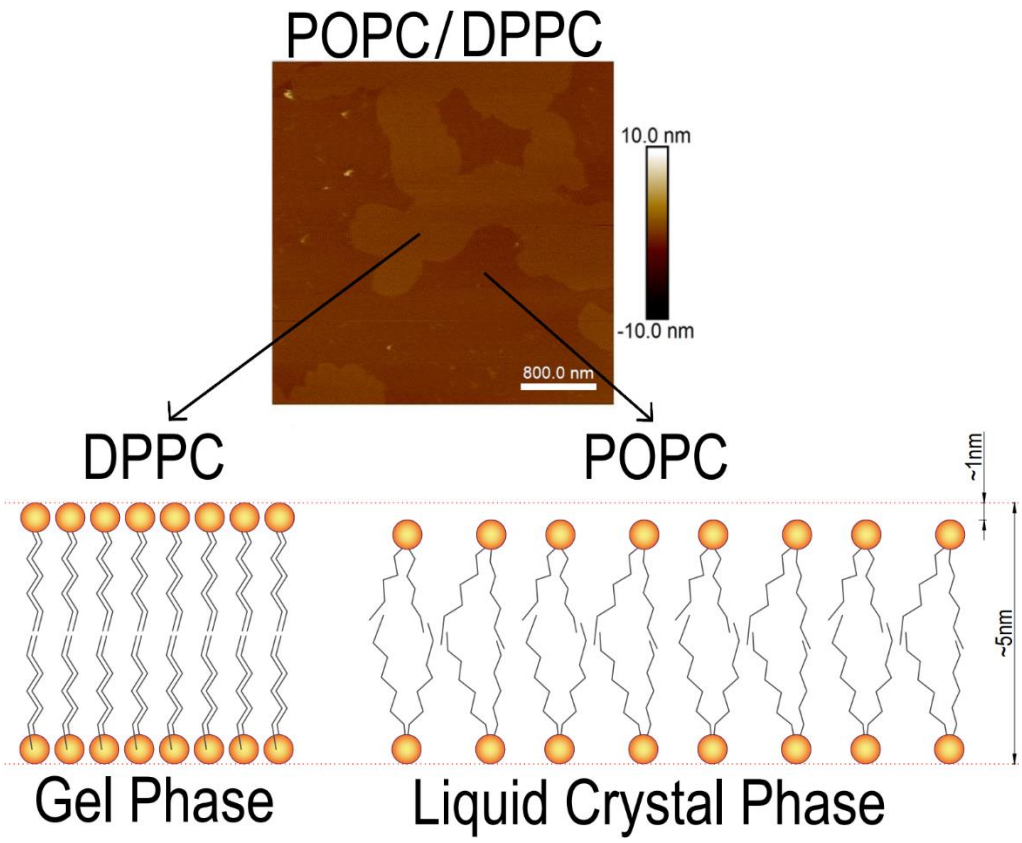
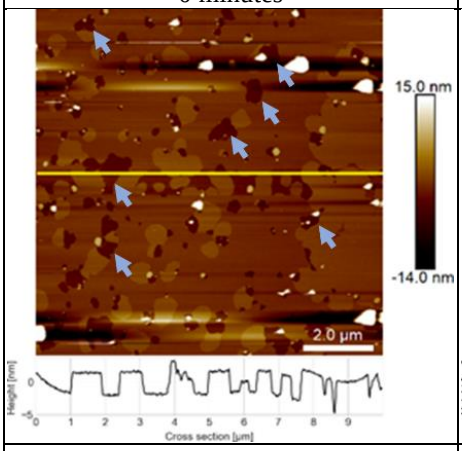
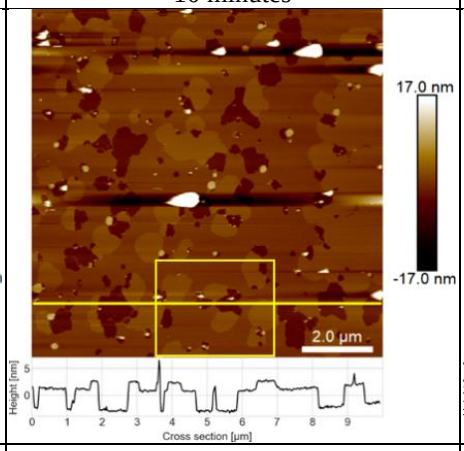
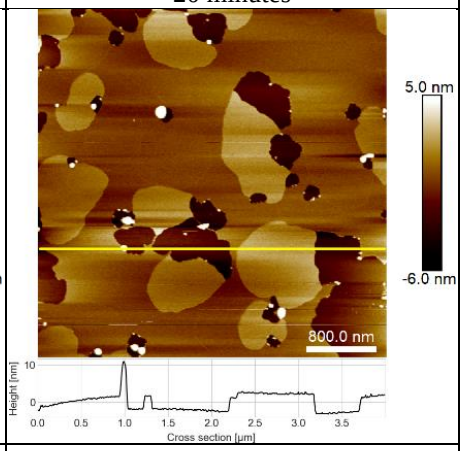
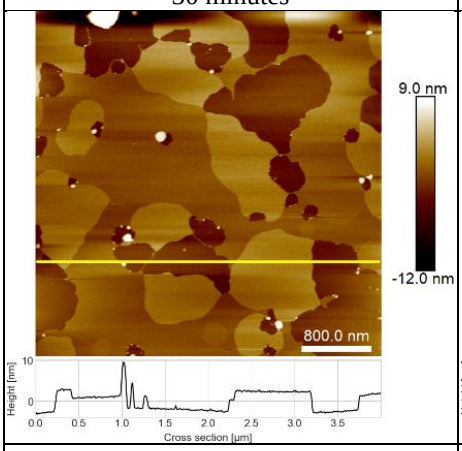
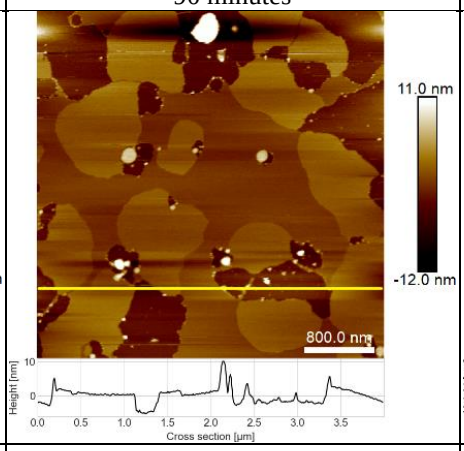
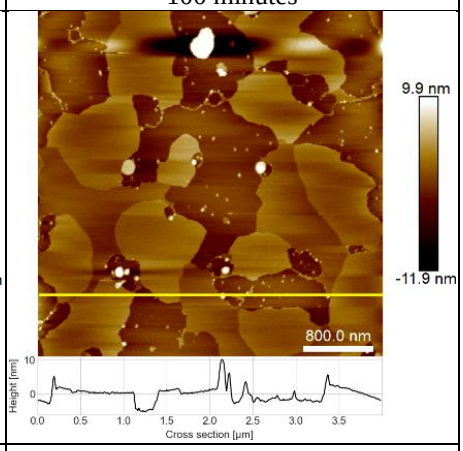
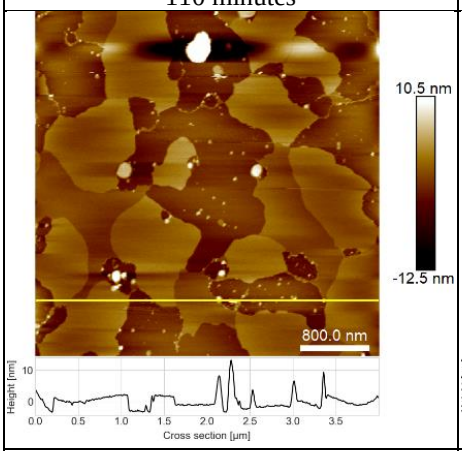
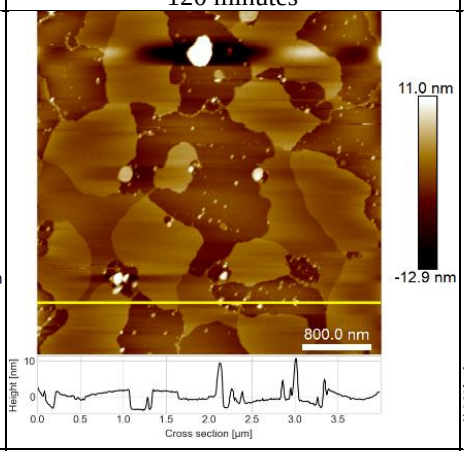
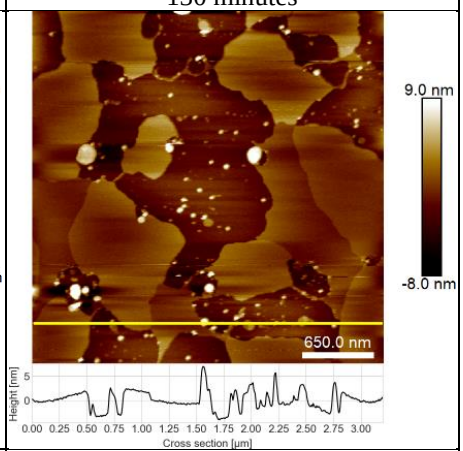
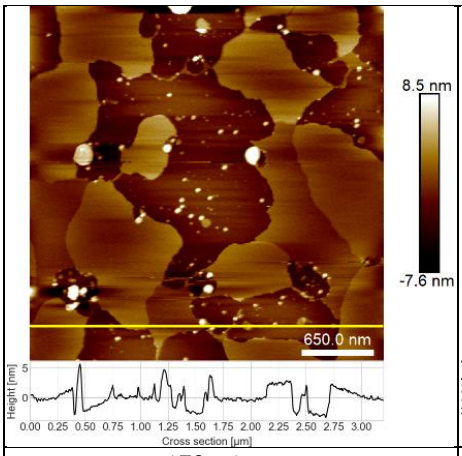


Figure 59 Topography image of SLB formed from POPC/DPPC on the top, model representation of two lipid domains on the bottom.

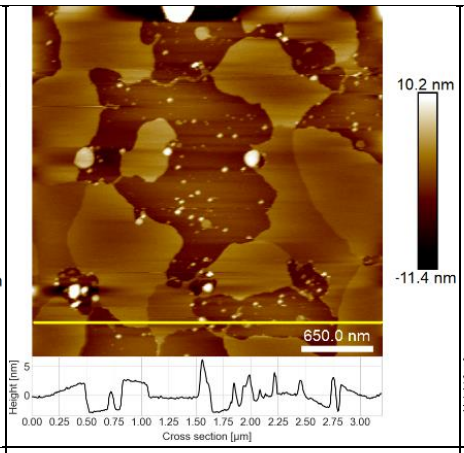
To fully understand the behavior of Aβ42 embedded within the SLB, and possibly the peptides self-assembly process within lipids, the evolution of monomeric form of Aβ42 was studied (see Table 17). The dynamics of lipid domains and impact of 20 μM Aβ42 were followed during 290 minutes. Based on the registered topographical images, we can observe that progression of the membrane damage occurred within minutes (compare the same region of SLB scan after 20 versus 100 minutes, Table 17). Directly after cooling down SLB and starting the visualization process, we can see holes in the lipid membrane (see image at 0 minutes, blue arrows). The rearrangements of the liquid and gel phases can be also observed. What is more, a lot of damages (holes) can be noticed within the lipid membrane, (dark domains ~10–100 nm, corresponding to the mica substrate) mainly located within the liquid phase, e.g. the POPC domains as well as cracks in the DPPC gel phase domains (bright areas). Assumably, as shown in a recent publication [141]. On the topographical images of SLB with embedded Aβ42, remaining peptides adsorbed on the surface of the SLB appear as bright dots more or less oligomerizing (large bright spot). These are believed to be deposits (few nanometer high) of free peptides.

Table 17 Evolution of monomeric form of Aβ42 peptides embedded in POPC/DPPC SLB at room temperature. Topography images from AFM with cross-section plots corresponding to the yellow lines of AFM images. Regions marked with yellow rectangle are showed in higher magnification on subsequent pictures. The peptides Aβ42 indicated with yellow arrows and holes in blue color. Cpep = 20 µM.

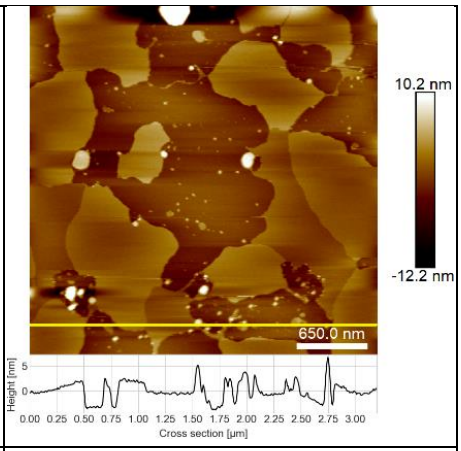
0 minutes	10 minutes	20 minutes
		
30 minutes	**90 minutes**	**100 minutes**
		
110 minutes	**120 minutes**	**130 minutes**
		
140 minutes	**150 minutes**	**160 minutes**



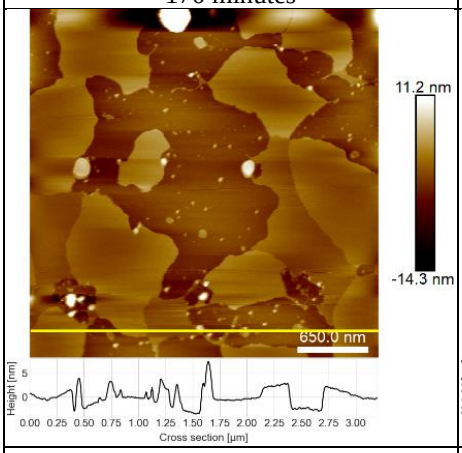
170 minutes



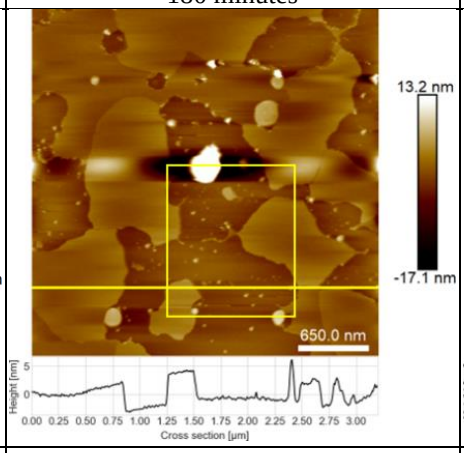
180 minutes



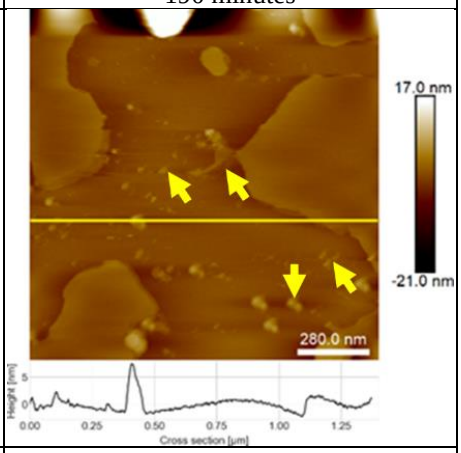
190 minutes



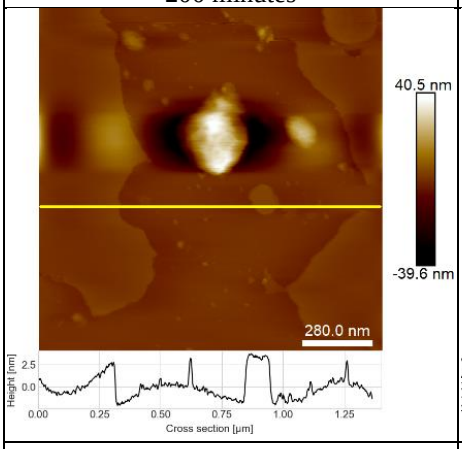
200 minutes



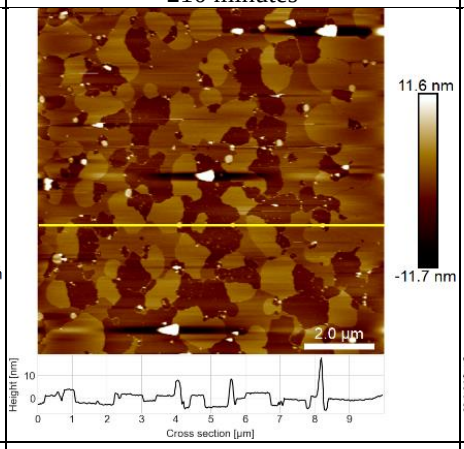
210 minutes



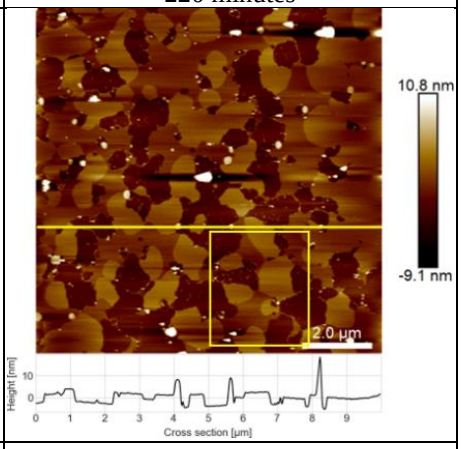
220 minutes



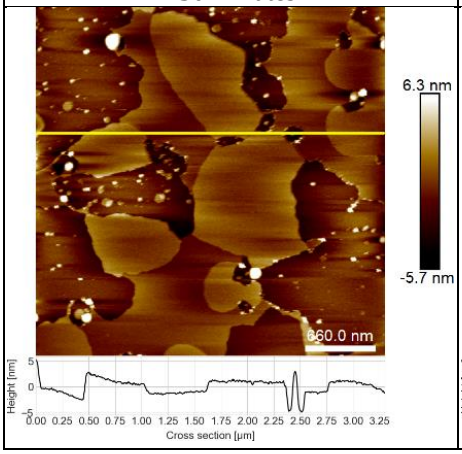
230 minutes

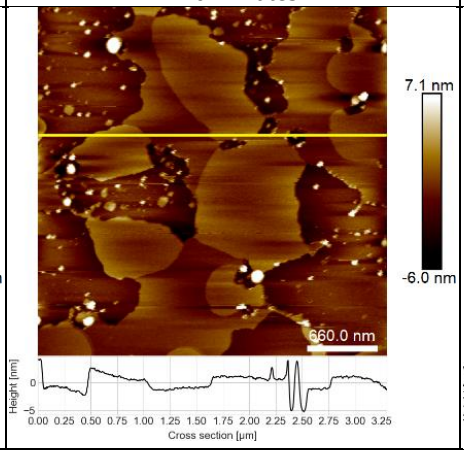


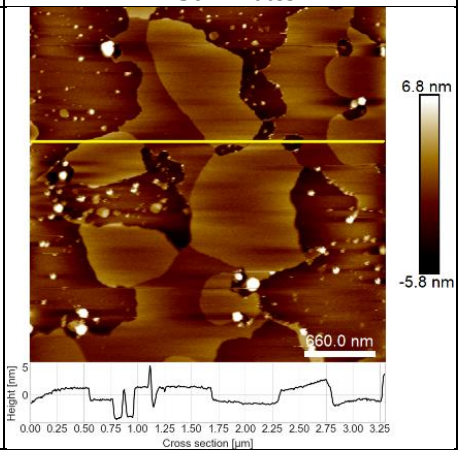
240 minutes

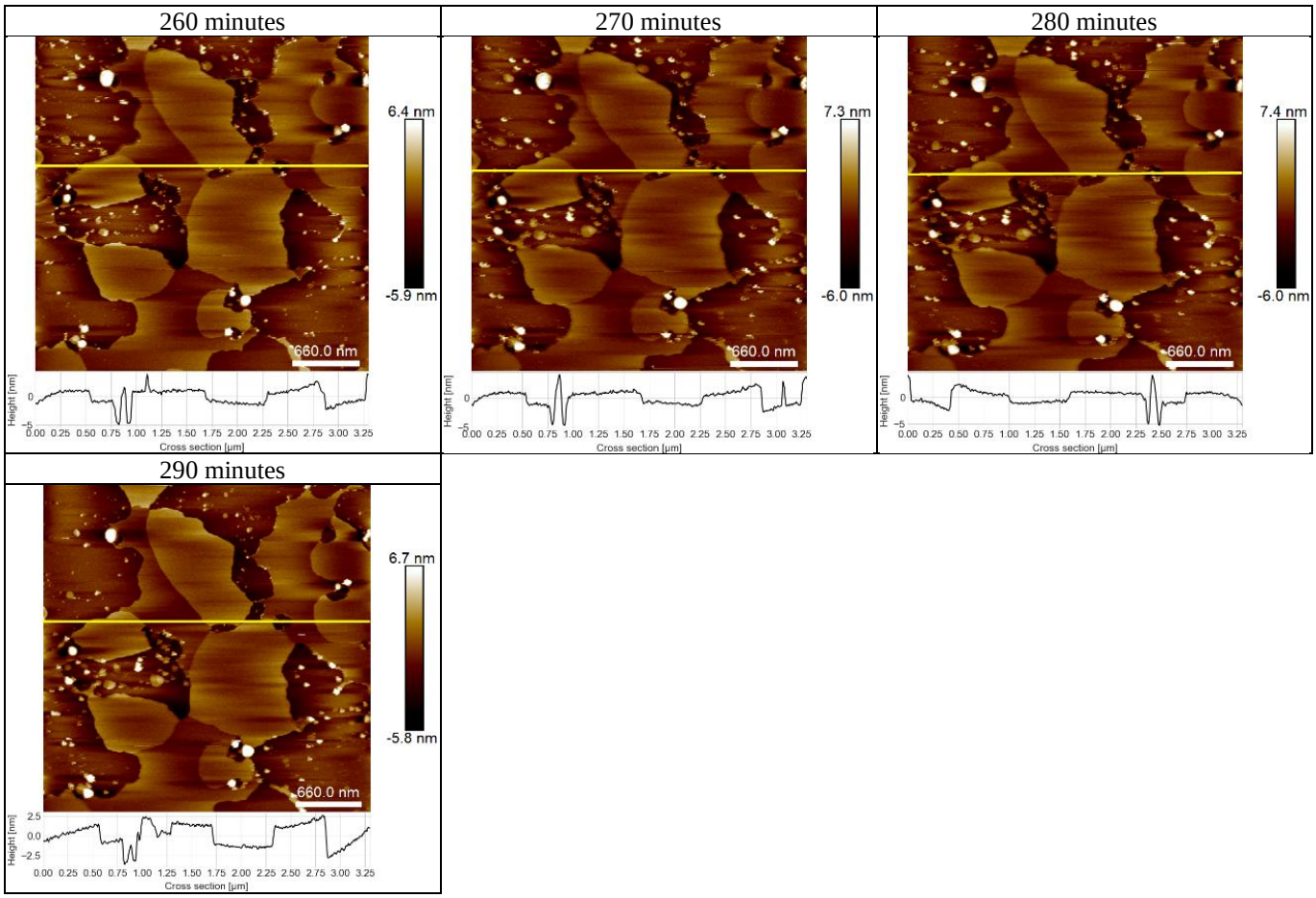


250 minutes







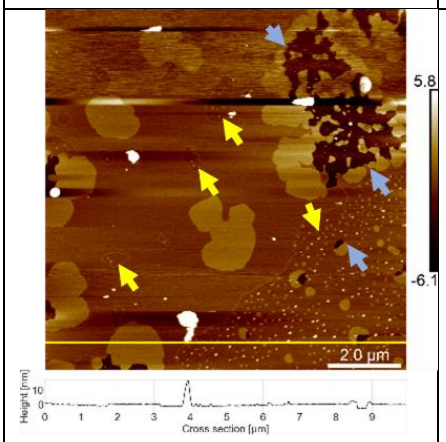
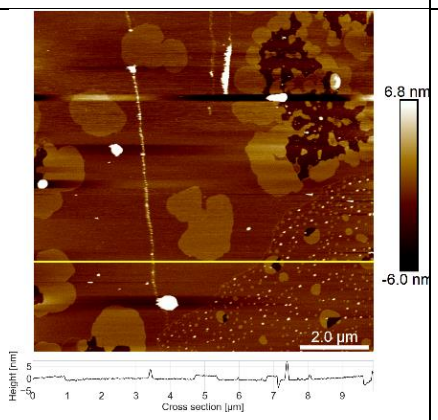
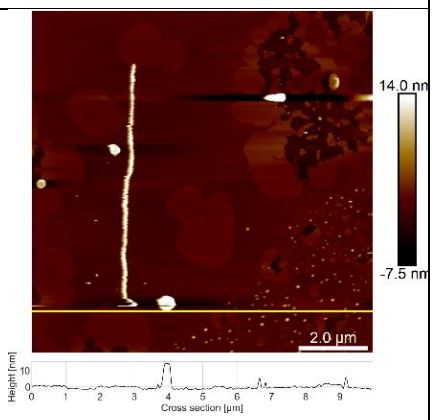
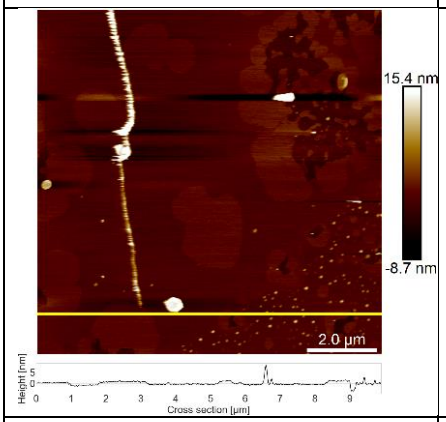
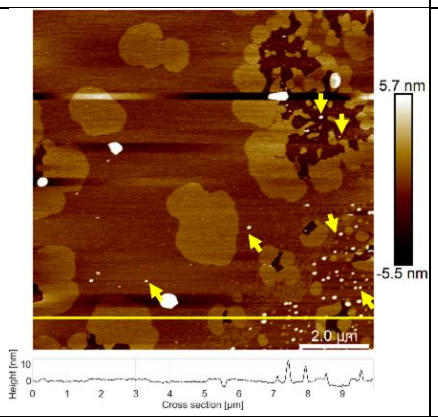
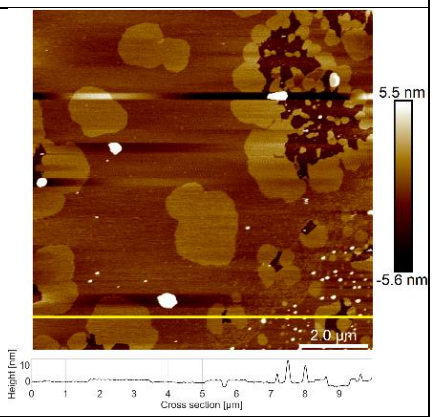
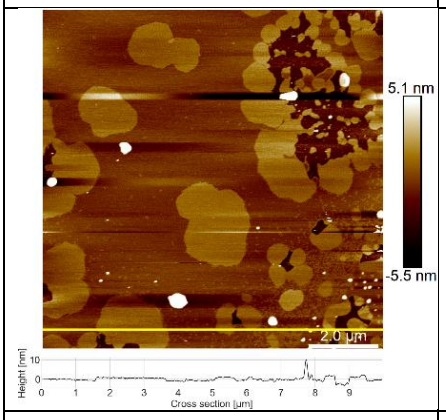
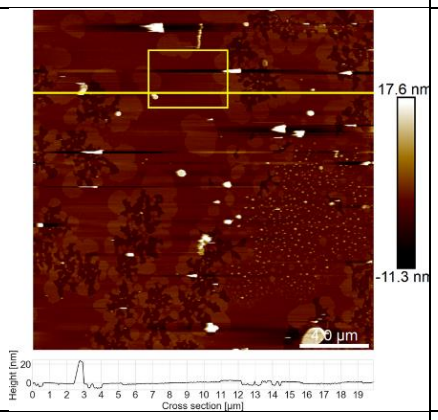
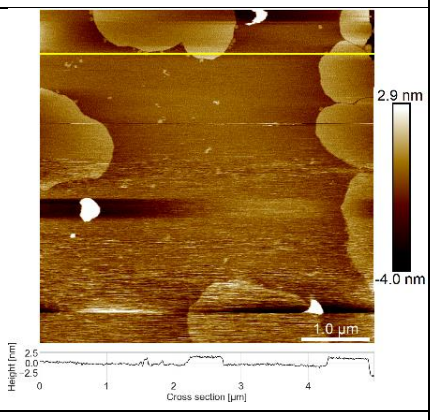


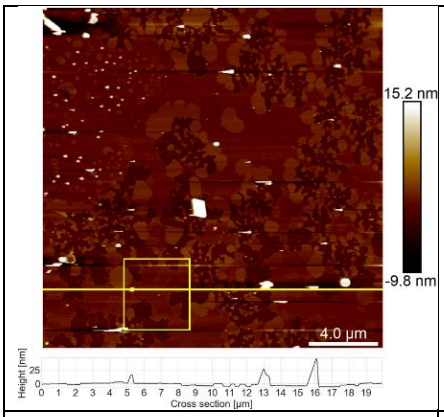
In the second part of this AFM study, monomerized hIAPP at 20 µM concentration was added to preformed SLBs with incorporated Aβ42, after cooling the lipid membrane. These samples should allow investigation of the cross-interaction between Aβ42 and hIAPP, and check whether significant differences can be observed in the membrane disruption (see Table 18). Quite interestingly, the AFM scans show that within 10 minutes of visualization, a long fibril of 1 nm height can be observed on top of the liquid phase (second image in Table 18). However, based on the time evolution, we can observe that after 30 minutes it detached from membrane surface and probably went back to the buffer solution. Additionally, in the liquid-disordered phase, there were much more monomers and they were much brighter than in previous system, with Aβ42 alone. At ~150 minutes, on the AFM image, we can observe that the cluster of peptide's deposits was formed or rearrangements of gel phase occurred (marked with orange arrows). As time progressed, the ongoing disintegration of the gel phase and peptide molecules becomes increasingly evident. The molecules diffusing in the liquid phase of the membrane, tended to accumulate at the edges of this phase. This phenomenon was not observed neither in the system of SLB with embedded Aβ42 (see Table 17). The presence of hIAPP peptide had an effect on the progression of the lipid membrane damage, which may have resulted from the cross-interaction with Aβ42.

Based on this observation, our fourth hypothesis had been partially assess. Unlike the SLB with embedded Aβ42, larger lipid membrane disorganization was observed in the presence of hIAPP peptide on top of the SLB.

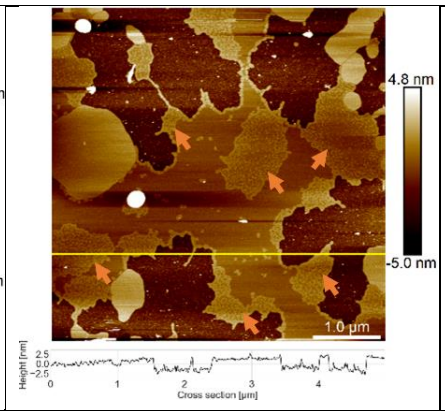
However, one should point out that the exact mechanism is still unclear and further analysis is needed. Previous studies showed that lipid-peptide interaction play crucial role in the self-assembly process of Aβ40 and hIAPP [229]. Seelinger et al. studied cross-interaction of Aβ40 and hIAPP in the presence of complex (raft-like) biomembrane model. They found that within 3 hours hIAPP penetrated the lipid head group region of the disordered domain, then later, hIAPP penetrated deeper, as manifested by the change of hydrocarbon chain region thickness (X-ray Reflectivity and Grazing Incidence Diffraction Measurements). They reported no interactions within the ordered domains. Besides, the interaction of Aβ40 was also studied and it revealed effects similar to those induced by hIAPP, though more impactful. An equimolar mixture of Aβ40 and hIAPP was injected in the solution at the air-water interface in a Langmuir film. The dominant for cross-interaction effect was comparable to only Aβ40 interaction. However, the effect of mixture escalated over the time. The continuous decrease in electron density of the head group over time and the increasing thickness of the head group, are related to oligomer formation. However, the process of hIAPP oligomerization induced by the membrane is delayed by the presence of both peptides in the mixture, as evidenced by the retarded increase of the electron density in the lipid chain region. The decreasing electron density in the lipid head group area can be interpreted as an increased disorder of the packing of lipid molecules in the monolayer. The mechanism of cross-interaction of Aβ40 and hIAPP, proposed by Seelinger et al., showed that peptides interacted mainly with the disordered domain of the raft like lipid membrane, creating heterofibrils [229].

Table 18 Evolution of monomeric form of Aβ42 embedded in POPC/DPPC and hIAPP peptides added to the SLB at room temperature. Topography images from AFM with the cross-sections plots corresponding to the yellow lines of AFM images. Regions marked with yellow rectangle are showed in higher magnification on subsequent pictures. The peptides indicated with yellow arrows and holes in blue color. Cpep = 20 µM.

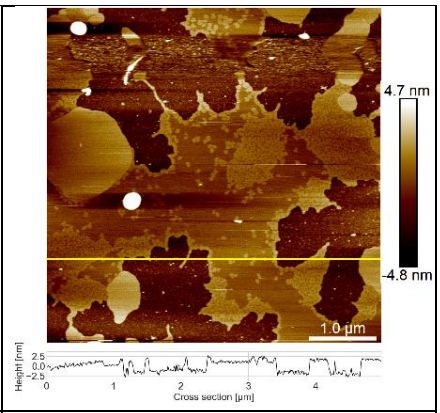
0 minutes	10 minutes	20 minutes
		
30 minutes	90 minutes	100 minutes
		
110 minutes	120 minutes	130 minutes
		
140 minutes	150 minutes	160 minutes



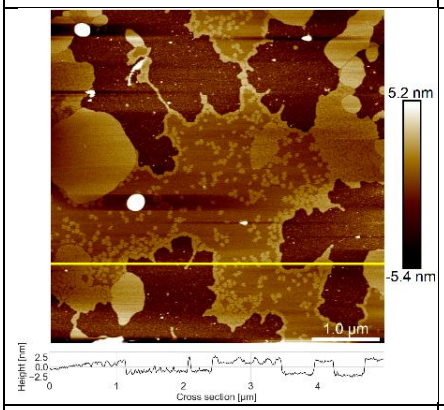
170 minutes



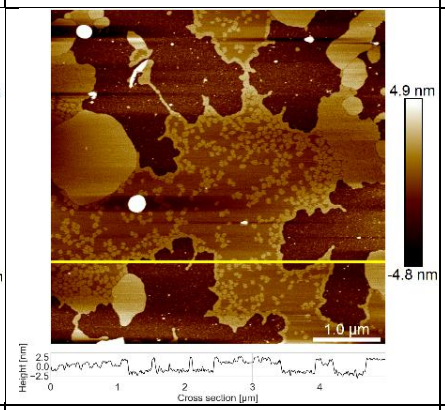
180 minutes



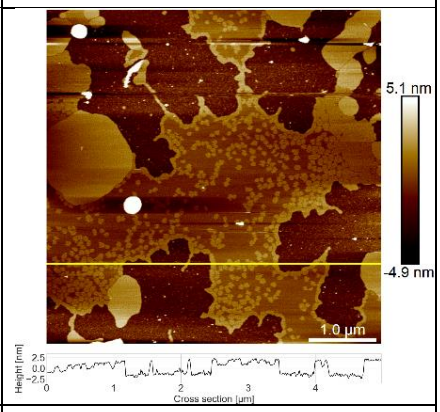
190 minutes



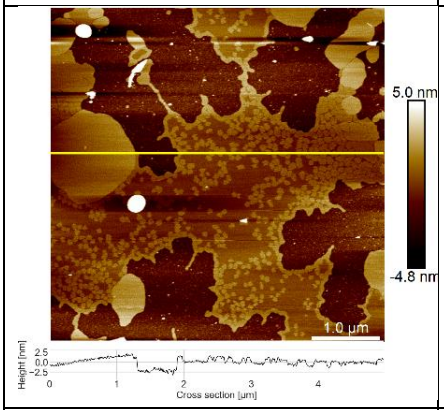
200 minutes



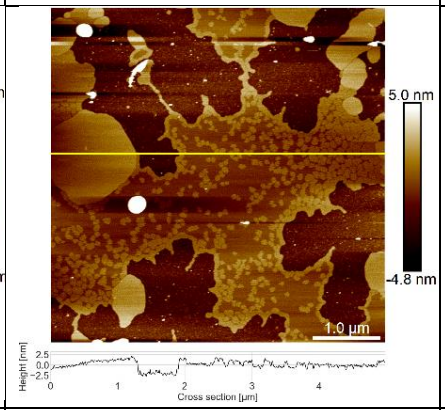
210 minutes



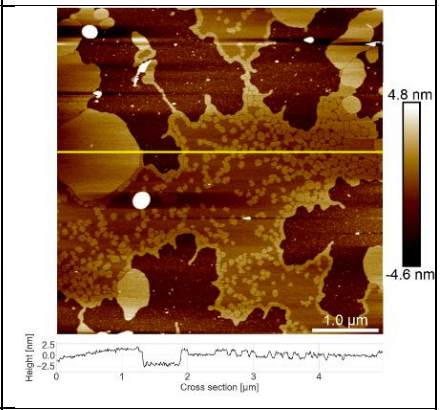
220 minutes



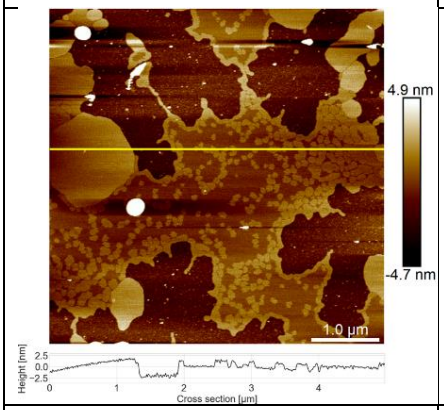
230 minutes



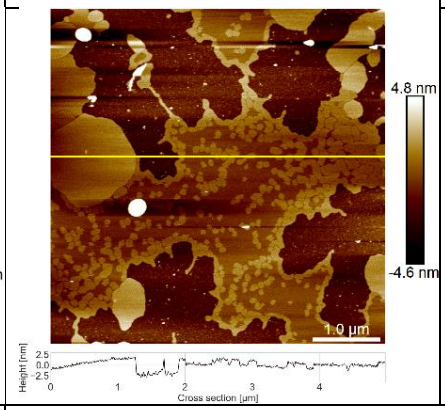
240 minutes

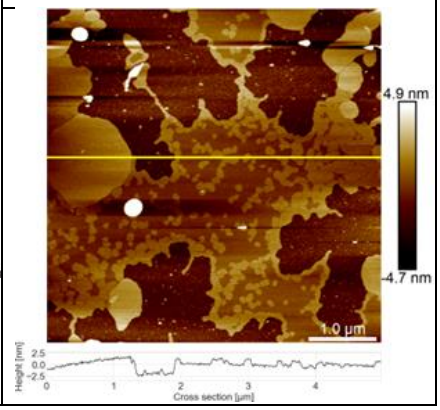


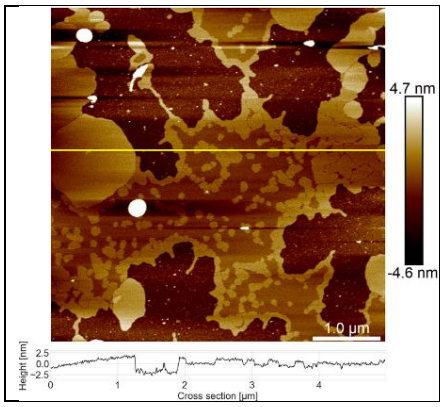
250 minutes



260 minutes







To summarize, this study shows that the incorporation of 20 µM Aβ42 in SLBs formed from POPC/DPPC lipids results in the disruption of the lipid bilayer. The disruptions were observed as holes in SLB (see Table 17). The time evolution also showed disorganization in the ordered and disordered phases, likely caused by the interaction with the Aβ42 peptides. On the topological AFM images, only monomers and oligomers were observed, mainly in the liquid phase (POPC). The peptides present in the membrane have a helical part embedded with the bilayer and probably self-assemble to form the boundaries of liquid domain that detach from the bilayer producing hence holes and defects. These results provide additional support for the 3rd hypothesis, stating that *the presence of the lipid membrane affects the aggregation of the native Aβ42 peptide*.

In the second part of this experimental validation, an equimolar concentration of amylin was added on top of the SLB to investigate how the presence of another peptide influences the fibrillization process of Aβ42. AFM investigation revealed that the presence of amylin caused more disruption of the membrane. Additionally, for a brief moment, fibrils were visible on top. The disruption process of the lipid membrane was much more pronounced in the presence of both Aβ42 and amylin than in the sole presence of embedded Aβ42. The only evidence that the presence of another peptide impacts the general fibrillization process (4th hypothesis) is the higher height value of monomers and oligomers (cross-section plots in Table 18). The only observed side effect of the interaction is the higher membrane disruptions and various disorganizations in the liquid and ordered phases of SLB. In this experimental setup, Aβ42 and amylin cannot be specifically identified. To explore the exact effect further, fluorescence staining/chemical labelling needs to be utilized for each of the peptides

Discussion

The primary objective of this thesis was to examine and evaluate the influence of different internal and external factors on the process of amyloidogenic peptide aggregation. Better understating how such conditions can alter or even inhibit the aggregation pathways would facilitate the advancement of drug development. Moreover, it underscores the potential challenges arising from the ambiguity of experimental outcomes in the context of amyloid investigations.

Firstly, the effect of the solvent on the peptides' tendency to aggregate was studied. Short hexapeptide sequences exhibit a distinct aggregation propensity in response to external factors, like the solvent used, compared to longer sequences (up to 23 amino acid), as shown in preliminary studies. Sequences of 6 amino acids are having a backbone with substantial flexibility and more space to adopt specific conformation, taking into account the interaction with solvent molecules. One of the factors, which influences such phenomena is the concept of symmetry-breaking transitions. Symmetry-breaking transitions happen when there is a comparable energy difference between ordered and disordered structures, which results in the spontaneous formation of patterns with spatial modulation [230]. This concept is especially characteristic for IDPs proteins/peptides such as Aβ or hIAPP. Particularly, in case of D_2O solvent, which changes the classification of R2 fragment of *S. enterica*. This repeat has a modulation role in whole CsgA protein [161]. For R2 fragment in NaOH+PBS, fibrils were not observed. Based on the secondary structure analysis, this fragment had dominant β-turn confirmation (Table 13). In D_2O, the dominant conformation was attributed to intermolecular aggregates, a signature typical for amyloids structures. Additionally, TEM micrographs proved the presence of regular, straight fibrils (Figure 18). Based on that, the R2 fragment could be classified as amyloid or non-amyloid, depending on applied conditions, which proves the first hypothesis. This issue might occur in various experimental methods, that require the use of such solvent, as for example in spectroscopic methods. Note that D_2O has only 1% stronger hydrogen bonds than H_2O [66].

This raises the question whether it is only the solvent effect, or maybe it is the effect of ions present in the solvent. Stephens et al. showed that the presence of NaCl ions decrease the water mobility within the solvation shell and increases the aggregation of α-synuclein. On the other hand, in case of CsI ions, the mobility in the solvation shell is increased and the aggregation process of α-synuclein decreases [231]. The research also revealed that after changing the H_2O to D_2O, the aggregation process was increased due to the solvent, and that the ions can influence the behavior and impact of solvent. In contrast to IDPs, for globular β-lactoglobulin [232] replacing H_2O with D_2O does not have a significant impact on the process of heat-induced denaturation and aggregation. The use of D_2O instead of H_2O results in a higher conformational stability, as indicated by a temperature shift of approximately 3°C

towards higher values in the differential scanning calorimetry thermograms, as well as a reduced rate of overall aggregation [232].

Based therefore on the research presented and the results obtained, we might suggest that the influence can be mainly linked to the overall structure of the peptide/protein. It is well-established that cells possess a full set of proteolytic tools to regulate protein homeostasis, or proteostasis, under diverse environmental conditions. To maintain cellular health, it is vital to regulate the balance between protein production and clearance. If this balance is disrupted, it can result in the accumulation of certain proteins, leading to pathological conditions [233].

What is more, not only the type of the solvent, but also the type of ions present, and their concentration, are important to consider. The study of the mutant sequences of R4 fragments of CsgA protein highlighted the influence of the phosphate buffer ions on the peptides' morphology. The ions affected the local electrostatic interactions involving the polypeptide chains. The MD simulations revealed that the interactions of positively charged amino acids with negatively charged phosphate moieties in the buffer contributed to the morphology (Table 20). Non-specific ion enrichment may occur in the proximity of protein moieties that bear charges opposite to those of the ions. Hence, the dimers structures observed in the simulation may have contributed to the snail-like conformation of the M4 peptide (Figure 27). Possibly, the addition of phosphate anions could also affect the aggregation pathways, however it was not checked in this study. The effect of Hofmeister ions is known for 130 years [234], however the impact of salt selection in solvents on proteins remains unclear. Dogra et al. [235] has shown that ions have a significant impact on both the aggregation rate and the nanoscale structure during assembly of the repeat domain, a proteolytic subproduct of the human Pmel17 protein. Pmel17 is a functional amyloid that contains a pH-responsive intrinsically disordered region [236], which was investigated. Additionally, ions affect the autocatalytic amplification processes via a fascinating dual Hofmeister effect. The propensity for oligomerization is notably higher when exposed to kosmotropes in comparison to chaotropes. In addition, recent investigations into the aggregation mechanism of Aβ and α-synuclein have revealed that the regulation of the secondary nucleation processes is highly dependent on the interplay between electrostatic interactions and salt concentration [237], [238]. In this study, the simulation of hIAPP when placed in solution above the lipid membrane showed that in the system with NaCl ions, the amylin interacts mainly with lipid head groups, while with $CaCl_2$ ions, amylin penetrated deeper inside the bilayer hydrophobic core (Figure 44 and Figure 47). Interestingly this result is consistent with experiments conducted by Sciacca et al. which showed that Ca^{2+} ions promote shallow membrane insertion of hIAPP [239].

The next factor that should be considered in the amyloid classification is the usage of a few experimental method, to support the classification. Here, two IR modes (transmission and ATR), which

can give misleading results even in various measurement modes, were compared. Proteins exhibit an affinity for adsorption onto the surface of the ATR crystal, which subsequently induces alterations in their structural conformation. As a result, this process may accelerate the aggregation of studied peptide/protein (Table 11).

Additionally, in this thesis, the suitability of the Kelvin probe for analyzing amyloids was investigated. The results indicated that the CPD measured with this probe can offer quantitative insights into the molecular features of amyloid aggregates, which are consistent with those obtained using other analytical techniques. As a result, the Kelvin probe has the potential to be employed as an effective tool for characterizing amyloid-based nanomaterials.

The next important studied factor is the amino acid sequence, with the special emphasis on how the sequence properties affect the amyloid aggregation propensity. This is in line with the second hypothesis: *aggregation propensity of closely related homologous functional amyloids may be different*. The R4 fragments of two bacteria strains showed a greater degree of gate-keeper residue modification (Figure 21). It is noteworthy that charged residues are predominantly present in the C terminal regions of peptides that have a tendency to aggregate (Figure 15). Our study provides further evidence for the significance of charge distribution, as demonstrated by the altered aggregation propensity of the R4 fragment. The R4 sequence derived from *E. coli* has two charged amino acids (6K and 17D) located outside of the terminal or linker region. However, in its aggregating counterpart form from *S. enterica*, these charged amino acids are lacking. The substitution of amino acids at positions 17 and 21 results in a shift in the charge distribution. This, along with the loss of charge at position 6, leads to the formation of a structure where all charged amino acids are located on one side of the folded peptide and outside the core of the predicted β-arch. These substitutions are likely to increase the amyloidogenic potential of the R4 fragment from *S. enterica*. The kinetic analysis of ThT revealed an exceedingly rapid aggregation of R4 derived from *S. enterica* without a lag phase, which sharply contrasted with the ThT study results of R4 derived from *E. coli*, that did not display any aggregation (Figure 20). The results of the ATR-FTIR spectroscopy were consistent with these results, exhibiting a higher prevalence of aggregates in R4 from *S. enterica* and random coils in *E. coli*. In turn, previous research had reported that ER4 does not aggregate under similar conditions [161]. R4 from *E. coli* also forms amyloid fibrils, but the process is notably slower than that of R4 from *S. enterica*.

Further investigation of the impact of designed mutations on the aggregation propensity of self-assembling peptides has yielded significant effects in the mutant sequences. These effects were evidenced by various analytical techniques. Notably, the positioning of charged amino acids within the peptide sequence exerts a discernible influence on essential aggregation properties, e.g. aggregation rate, overall structure, and morphology. While the general motif of CsgA protein remains conserved,

the sub-molecular characteristics are evidently sensitive to specific mutations. Moreover, the study has uncovered a novel type of assembly, characterized by spherically packed globular aggregates, in addition to the typical fibrillar structures formed by the self-assembling peptides. The introduced substitutions and additional presence of ions modulate the aggregation propensity of studied peptides. Especially, when in the case of charged amino acids, we take into account the protonation state in the specific environment [240]. As an example, introduction in the STVIIE peptide, a positively charge amino acids (Lys^{+}, Arg^{+}, His^{+}) at the first position, induce fibril formation [158]. Lopez de la Paz et al. pointed out that in case of amyloids, the charge amino acids are located at the edges of the sequence. Their findings underscore the crucial role of electrostatic interactions in the supramolecular self-assembly of peptides/proteins into fibrils, proposing that the preferential formation of amyloid fibrils in peptides/proteins could be attributed to particular charge states of the polypeptide chain. Charged amino acids are crucial in the interaction with lipid membranes. For example, the hIAPP peptide's binding affinity to the lipid membrane is modulated by the protonated state of histidine, at 18^{th} position [241], [242].

The last phase of the study was related to the presence of the lipid membrane, in which the stability of Aβ42 in DOPC lipid membrane was checked (Figure 32). The MD simulations and AFM investigations partially assessed the 3^{rd} hypothesis of this thesis. Note that APP, the Aβ42 precursor and γ-secretase (lipophilic proteins) are localized in the mitochondria [243], [244]. Furthermore, APP is secreted within the lipid membrane [245]. Additionally, APP is supposed to be processed in the plasma membrane or in the endosomal system [244]. Starting from the premise that nascent Aβ42 is located within a lipid membrane, atomistic simulations carried out have shown that the peptide remains embedded in a DOPC lipid membrane, which suggests that the aggregation of these peptides might occur within the membrane (Figure 31). AFM experimental studies of Aβ42 incorporated within POPC/DPPC SLB, showed that the latter display drastic damage. The peptides are mainly found in the disordered phase (POPC), which was also supported by other studies [141], [229]. On the other hand, some other research suggests that the fibrils found in tissue samples collected from AD patients are attached to lipid particles, primarily cholesterol, phospholipids, and ganglioside GM1 [246].

Our study of the cross-interaction between Aβ42 and hIAPP revealed that the secondary structure exhibited a higher β-sheet content (Figure 53 and Figure 57) in comparison to isolated (Figure 34 and Figure 48). This can be attributed to the fact that these two peptides were interacting since in close proximity, assessing therefore the fourth hypothesis was partially supported by this study. Furthermore, AFM investigations have demonstrated that the SLB was more affected in the presence of Aβ42 and amylin, compared to the damage caused by Aβ42 alone. The exact impact on the Aβ42 structure is difficult to assess. The existence of Aβ42 can be regarded as an environmental and physiological

element that triggers the aggregation process. Altogether, investigating the physiological conditions, that trigger the formation of protein-based amyloid fibrils in the brain, has been a challenging endeavor.

In conclusion, a standardized protocol for studying amyloid peptides/proteins does not exist. The choice of experimental conditions must depend on the physicochemical properties of the sequence, such as the charge, pI, and hydrophobicity index. It is also crucial to acquire knowledge about the physiological conditions in the case of pathological amyloids and the conditions in which the peptide/protein is expressed in the case of functional amyloids, such as specific temperature or pH. These two pieces of information serve as starting points for the experimental investigation. It is crucial to emphasize that certain experimental techniques may impose specific conditions, such as solvent or concentration, which must be considered when designing the experimental setup. For instance, when studying longer sequences (over 20 amino acids) in deuterium oxide, the effect of the solvent must be taken into account, as it might accelerate the self-assembly process. The deuteration process, which involves replacing hydrogen atoms in peptides/proteins with deuterium atoms, is not fully understood. Deuterium oxide have a number of applications in peptide/protein research such as neutron scattering or spectroscopic techniques. What is more, the same effect can be found with the Hofmeister ions.

In more complex systems, such as those involving lipid membranes or the presence of additional peptide, a multitude of additional factors must be considered, which can greatly influence the aggregation tendency of the studied peptide/protein. In order to obtain a better understanding of the role specific factors, it is important to undertake a systematic study, beginning with simpler models and gradually introducing and observing the effects of additional factors. Such an approach will ensure that the validation process is comprehensive and provides a robust understanding of the system under investigation.

Summary

The results of the study indicate the importance of collecting consistent data from experiments conducted under various conditions. Additionally, different experimental methods may yield different results for the same peptides, with this effect being more noticeable for certain sequences. Bioinformatics predictors are quite resilient to data incompatibilities, but their performance may be disrupted if the influence of experimental conditions and insignificant differences in homologous sequences is not taken into account. It is essential to thoroughly acknowledge the non-identical nature of experimental conditions, particularly when constructing machine learning-based bioinformatics predictors.

Therefore, it is crucial to consider these factors when planning experimental research, as appropriate selection of experimental conditions can help modulate the peptide's tendency to aggregate.

Future perspectives

There are numerous opportunities for further research into the impact of intrinsic and extrinsic factors on the tendency of amyloids to aggregate. Especially, in the case of proposed novel mechanism of the oligomerisation of Aβ42, and the interaction with hIAPP, further investigation is necessary. Particularly, to evaluate the influence of several factors, such as lipid composition, cholesterol's function, and the presence of gangliosides, all of which have demonstrated implications for both the progression and pathology of the disease, and on the morphological alterations that have been observed.

Increasingly, scientific research indicates that prevailing force fields are unsuitable for investigating short, non-helical structures, such as the peptides analyzed in this thesis. Owing to their inherent structural flexibility, conventional experimental techniques can not accurately capture the diverse structural features of IDP with confidence, as noted in the study conducted by Mu et al. 2021 [247]. It is evident that further investigations are necessary to utilize force fields that are specifically designed for IDPs, examples: CHARMM36IDP, Amber14IDP [248], [249]. Additionally, it would be interesting to study how various properties of the lipid head group (such as charges), solvent characteristics, protonation state of the peptide influence the cross-interaction of Aβ and hIAPP. It could aid in better comprehending the various factors that influence the outcome of simulations.

Bibliography

[1] M. D. Benson, J. N. Buxbaum, D. S. Eisenberg, G. Merlini, M. J. M. Saraiva, Y. Sekijima, J. D. Sipe, and P. Westermark, "Amyloid nomenclature 2020: update and recommendations by the International Society of Amyloidosis (ISA) nomenclature committee," *Amyloid*, vol. 27, no. 4, pp. 217–222, Oct. 2020,

[2] R. Riek, "The Three-Dimensional Structures of Amyloids," *Cold Spring Harb. Perspect. Biol.*, vol. 9, no. 2, Feb. 2017,

[3] E. Tayeb-Fligelman, O. Tabachnikov, A. Moshe, O. Goldshmidt-Tran, M. R. Sawaya, N. Coquelle, J. P. Colletier, and M. Landau, "The cytotoxic Staphylococcus aureus PSMα3 reveals a cross-α amyloid-like fibril," *Science (80-.).*, vol. 355, no. 6327, pp. 831–833, 2017,

[4] A. Daskalov, N. El Mammeri, A. Lends, J. Shenoy, G. Lamon, Y. Fichou, A. Saad, D. Martinez, E. Morvan, M. Berbon, A. Grélard, B. Kauffmann, M. Ferber, B. Bardiaux, B. Habenstein, S. J. Saupe, and A. Loquet, "Structures of Pathological and Functional Amyloids and Prions, a Solid-State NMR Perspective," *Front. Mol. Neurosci.*, vol. 14, p. 123, Jul. 2021,

[5] R. Tycko, "Amyloid polymorphism: structural basis and neurobiological relevance," *Neuron*, vol. 86, no. 3, May 2015,

[6] M. Kollmer, W. Close, L. Funk, J. Rasmussen, A. Bsoul, A. Schierhorn, M. Schmidt, C. J. Sigurdson, M. Jucker, and M. Fändrich, "Cryo-EM structure and polymorphism of Aβ amyloid fibrils purified from Alzheimer's brain tissue," *Nat. Commun. 2019 101*, vol. 10, no. 1, pp. 1–8, Oct. 2019,

[7] M. R. Sawaya, M. P. Hughes, J. A. Rodriguez, R. Riek, and D. S. Eisenberg, "The expanding amyloid family: Structure, stability, function, and pathogenesis," *Cell*, vol. 184, no. 19, pp. 4857–4873, Sep. 2021,

[8] F. Chiti and C. M. Dobson, "Protein Misfolding, Amyloid Formation, and Human Disease: A Summary of Progress Over the Last Decade," *Annu. Rev. Biochem.*, vol. 86, pp. 27–68, Jun. 2017,

[9] Z. L. Almeida and R. M. M. Brito, "Structure and Aggregation Mechanisms in Amyloids," *Mol. 2020, Vol. 25, Page 1195*, vol. 25, no. 5, p. 1195, Mar. 2020,

[10] M. Fändrich, M. A. Fletcher, and C. M. Dobson, "Amyloid fibrils from muscle myoglobin," *Nat. 2001 4106825*, vol. 410, no. 6825, pp. 165–166, Mar. 2001,

[11] D. Otzen and R. Riek, "Functional Amyloids," *Cold Spring Harb. Perspect. Biol.*, vol. 11, no. 12, Dec. 2019,

[12] P. Westermark, M. D. Benson, J. N. Buxbaum, A. S. Cohen, B. Frangione, S. I. Ikeda, C. L.

Masters, G. Merlini, M. J. Saraiva, and J. D. Sipe, “Amyloid: Toward terminology clarification Report from the Nomenclature Committee of the International Society of Amyloidosis,” *Amyloid*, vol. 12, no. 1, pp. 1–4, Mar. 2009,

[13] C. M. Dobson, “Protein folding and misfolding,” *Nat. 2003 4266968*, vol. 426, no. 6968, pp. 884–890, Dec. 2003,

[14] S. Tomaselli, V. Esposito, P. Vangone, N. A. J. Van Nuland, A. M. J. J. Bonvin, R. Guerrini, T. Tancredi, P. A. Temussi, and D. Picone, “The alpha-to-beta conformational transition of Alzheimer’s Abeta-(1-42) peptide in aqueous media is reversible: a step by step conformational analysis suggests the location of beta conformation seeding,” *Chembiochem*, vol. 7, no. 2, pp. 257–267, Feb. 2006,

[15] K. Kuwajima, Y. Okamoto, T. Knowles, M. Vendruscolo, T. V. Sønderby, Z. Najarzadeh, and D. E. Otzen, “Functional Bacterial Amyloids: Understanding Fibrillation, Regulating Biofilm Fibril Formation and Organizing Surface Assemblies,” *Mol. 2022, Vol. 27, Page 4080*, vol. 27, no. 13, p. 4080, Jun. 2022,

[16] M. Vendruscolo, E. Paci, M. Karplus, and C. M. Dobson, “Structures and relative free energies of partially folded states of proteins,” *Proc. Natl. Acad. Sci. U. S. A.*, vol. 100, no. 25, p. 14817, Dec. 2003,

[17] B. Strodel, C. S. Whittleston, and D. J. Wales, “Thermodynamics and kinetics of aggregation for the GNNQQNY peptide,” *J. Am. Chem. Soc.*, vol. 129, no. 51, pp. 16005–16014, Dec. 2007,

[18] W. Zheng, M. Y. Tsai, and P. G. Wolynes, “Comparing the Aggregation Free Energy Landscapes of Amyloid Beta(1-42) and Amyloid Beta(1-40),” *J. Am. Chem. Soc.*, vol. 139, no. 46, p. 16666, Nov. 2017,

[19] A. Kessel and N. Ben-Tal, “Introduction to Proteins: Structure, Function, and Motion, Second Edition,” *Introd. to Proteins Struct. Funct. Motion, Second Ed.*, pp. 1–932, Jan. 2018,

[20] C. Wells, S. Brennan, M. Keon, and L. Ooi, “The role of amyloid oligomers in neurodegenerative pathologies,” *Int. J. Biol. Macromol.*, vol. 181, pp. 582–604, Jun. 2021,

[21] K. E. Marshall, K. L. Morris, D. Charlton, N. O’Reilly, L. Lewis, H. Walden, and L. C. Serpell, “Hydrophobic, aromatic, and electrostatic interactions play a central role in amyloid fibril formation and stability,” *Biochemistry*, vol. 50, no. 12, pp. 2061–2071, Mar. 2011,

[22] J. H. M. Van Gils, E. Van Dijk, A. Peduzzo, A. Hofmann, N. Vettore, M. P. Schützmann, G. Groth, H. Mouhib, D. E. Otzen, A. K. Buell, and S. Abeln, “The hydrophobic effect characterises the thermodynamic signature of amyloid fibril growth,” *PLoS Comput. Biol.*, vol. 16, no. 5, May 2020,

[23] K. Jong, L. Grisanti, and A. Hassanali, “Hydrogen Bond Networks and Hydrophobic Effects in

the Amyloid β30-35 Chain in Water: A Molecular Dynamics Study," *J. Chem. Inf. Model.*, vol. 57, no. 7, pp. 1548–1562, Jul. 2017,

[24] T. C. T. Michaels, A. Šarić, J. Habchi, S. Chia, G. Meisl, M. Vendruscolo, C. M. Dobson, and T. P. J. Knowles, "Chemical Kinetics for Bridging Molecular Mechanisms and Macroscopic Measurements of Amyloid Fibril Formation," *Annu. Rev. Phys. Chem.*, vol. 69, pp. 273–298, Apr. 2018,

[25] C. Frieden, "Protein aggregation processes: In search of the mechanism," *Protein Sci.*, vol. 16, no. 11, p. 2334, Nov. 2007,

[26] J. D. Harper and P. T. Lansbury, "MODELS OF AMYLOID SEEDING IN ALZHEIMER'S DISEASE AND SCRAPIE: Mechanistic Truths and Physiological Consequences of the Time-Dependent Solubility of Amyloid Proteins," *Annu. Rev. Biochem.*, vol. 66, no. 1, pp. 385–407, Jun. 1997,

[27] P. M. Martins, S. Navarro, A. Silva, M. F. Pinto, Z. Sárkány, F. Figueiredo, P. J. B. Pereira, F. Pinheiro, Z. Bednarikova, M. Burdukiewicz, O. V. Galzitskaya, Z. Gazova, C. M. Gomes, A. Pastore, L. C. Serpell, R. Skrabana, V. Smirnovas, M. Ziaunys, D. E. Otzen, S. Ventura, and S. Macedo-Ribeiro, "MIRRAGGE – Minimum Information Required for Reproducible AGGregation Experiments," *Front. Mol. Neurosci.*, vol. 13, p. 222, Nov. 2020,

[28] T. L. William, B. R. G. Johnson, B. Urbanc, A. T. A. Jenkins, S. D. A. Connell, and L. C. Serpell, "Aβ42 oligomers, but not fibrils, simultaneously bind to and cause damage to ganglioside-containing lipid membranes," *Biochem. J.*, vol. 439, no. 1, pp. 67–77, Oct. 2011,

[29] R. Kayed and C. A. Lasagna-Reeves, "Molecular mechanisms of amyloid oligomers toxicity," *J. Alzheimers. Dis.*, vol. 33 Suppl 1, no. SUPPL. 1, 2013,

[30] B. Mroczko, M. Groblewska, and A. Litman-Zawadzka, "The Role of Protein Misfolding and Tau Oligomers (TauOs) in Alzheimer's Disease (AD)," *Int. J. Mol. Sci.*, vol. 20, no. 19, Oct. 2019,

[31] F. Grigolato and P. Arosio, "The role of surfaces on amyloid formation," *Biophys. Chem.*, vol. 270, p. 106533, Mar. 2021,

[32] Z. Toprakcioglu, A. Kamada, T. C. T. Michaels, M. Xie, J. Krausser, J. Wei, A. Saric, M. Vendruscolo, and T. P. J. Knowles, "Adsorption free energy predicts amyloid protein nucleation rates," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 119, no. 31, p. e2109718119, Aug. 2022,

[33] J. Zhang and X. Y. Liu, "Effect of protein–protein interactions on protein aggregation kinetics," *J. Chem. Phys.*, vol. 119, no. 20, p. 10972, Nov. 2003,

[34] W. Smeralda, M. Since, J. Cardin, S. Corvaisier, S. Lecomte, C. Cullin, and A. Malzert-Fréon, "β-Amyloid peptide interactions with biomimetic membranes: A multiparametric

characterization," *Int. J. Biol. Macromol.*, vol. 181, pp. 769–777, Jun. 2021,

[35] S. Linse, "Monomer-dependent secondary nucleation in amyloid formation," *Biophys. Rev.*, vol. 9, no. 4, pp. 329–338, Aug. 2017,

[36] M. Törnquist, T. C. T. Michaels, K. Sanagavarapu, X. Yang, G. Meisl, S. I. A. Cohen, T. P. J. Knowles, and S. Linse, "Secondary nucleation in amyloid formation," *Chem. Commun.*, vol. 54, no. 63, pp. 8667–8684, Aug. 2018,

[37] P. Arosio, T. P. J. Knowles, and S. Linse, "On the lag phase in amyloid fibril formation," *Phys. Chem. Chem. Phys.*, vol. 17, no. 12, pp. 7606–7618, Mar. 2015,

[38] K. Gade Malmos, L. M. Blancas-Mejia, B. Weber, J. Buchner, M. Ramirez-Alvarado, H. Naiki, and D. Otzen, "ThT 101: a primer on the use of thioflavin T to investigate amyloid formation," *Amyloid*, vol. 24, no. 1, pp. 1–16, Jan. 2017,

[39] M. Groenning, "Binding mode of Thioflavin T and other molecular probes in the context of amyloid fibrils-current status," *J. Chem. Biol.*, vol. 3, no. 1, pp. 1–18, Mar. 2010,

[40] E. Stroo, M. Koopman, E. A. A. Nollen, and A. Mata-Cabana, "Cellular regulation of amyloid formation in aging and disease," *Front. Neurosci.*, vol. 11, no. FEB, p. 64, Feb. 2017,

[41] F. Chiti, M. Calamai, N. Taddei, M. Stefani, G. Ramponi, and C. M. Dobson, "Studies of the aggregation of mutant proteins in vitro provide insights into the genetics of amyloid diseases," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 99, no. SUPPL. 4, pp. 16419–16426, Dec. 2002,

[42] G. De Baets, J. Schymkowitz, and F. Rousseau, "Predicting aggregation-Prone sequences in proteins," *Essays Biochem.*, vol. 56, no. 1, pp. 41–52, Aug. 2014,

[43] A. M. Fernandez-Escamilla, F. Rousseau, J. Schymkowitz, and L. Serrano, "Prediction of sequence-dependent and mutational effects on the aggregation of peptides and proteins," *Nat. Biotechnol.*, vol. 22, no. 10, pp. 1302–1306, Oct. 2004,

[44] B. Bolognesi and G. G. Tartaglia, "Physicochemical Principles of Protein Aggregation," *Prog. Mol. Biol. Transl. Sci.*, vol. 117, pp. 53–72, Jan. 2013,

[45] F. Chiti, M. Stefani, N. Taddei, G. Ramponi, and C. M. Dobson, "Rationalization of the effects of mutations on peptide andprotein aggregation rates," *Nat. 2003 4246950*, vol. 424, no. 6950, pp. 805–808, 2003,

[46] X. Yang, G. Meisl, B. Frohm, E. Thulin, T. P. J. Knowles, and S. Linse, "On the role of sidechain size and charge in the aggregation of Aβ42 with familial mutations," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 115, no. 26, pp. E5849–E5858, Jun. 2018,

[47] C. W. O'Donnell, J. Waldispühl, M. Lis, R. Halfmann, S. Devadas, S. Lindquist, and B. Berger, "A method for probing the mutational landscape of amyloid structure," *Bioinformatics*, vol. 27, no. 13, p. i34, Jul. 2011,

[48] K. Sankar, S. R. Krystek, S. M. Carl, T. Day, and J. K. X. Maier, “AggScore: Prediction of aggregation-prone regions in proteins based on the distribution of surface patches,” *Proteins*, vol. 86, no. 11, pp. 1147–1156, Nov. 2018,

[49] F. Munir, S. Gul, A. Asif, and F. U. A. A. Minhas, “MILAMP: Multiple Instance Prediction of Amyloid Proteins,” *IEEE/ACM Trans. Comput. Biol. Bioinforma.*, vol. 18, no. 3, pp. 1142–1150, May 2021,

[50] J. Beerten, W. Jonckheere, S. Rudyak, J. Xu, H. Wilkinson, F. De Smet, J. Schymkowitz, and F. Rousseau, “Aggregation gatekeepers modulate protein homeostasis of aggregating sequences and affect bacterial fitness,” *Protein Eng. Des. Sel.*, vol. 25, no. 7, pp. 357–366, Jun. 2012,

[51] X. Wang, Y. Zhou, J. J. Ren, N. D. Hammer, and M. R. Chapman, “Gatekeeper residues in the major curlin subunit modulate bacterial amyloid fiber biogenesis,” *Proc. Natl. Acad. Sci. U. S. A.*, vol. 107, no. 1, pp. 163–168, 2010,

[52] M. Noji, T. Samejima, K. Yamaguchi, M. So, K. Yuzu, E. Chatani, Y. Akazawa-Ogawa, Y. Hagihara, Y. Kawata, K. Ikenaka, H. Mochizuki, J. Kardos, D. E. Otzen, V. Bellotti, J. Buchner, and Y. Goto, “Breakdown of supersaturation barrier links protein folding to amyloid formation,” *Commun. Biol. 2021 41*, vol. 4, no. 1, pp. 1–10, Jan. 2021,

[53] R. H. Choi and H. T. Kim, “Analysis of oligomeric complexes of the amyloid-forming FYLLYY peptide by collision-induced dissociation with electrospray ionization mass spectrometry,” *Eur. J. Mass Spectrom. (Chichester, Eng).*, vol. 26, no. 5, pp. 361–368, Sep. 2020,

[54] P. Hortschansky, V. Schroeckh, T. Christopeit, G. Zandomeneghi, and M. Fändrich, “The aggregation kinetics of Alzheimer’s β-amyloid peptide is controlled by stochastic nucleation,” *Protein Sci.*, vol. 14, no. 7, p. 1753, Jul. 2005,

[55] W. F. Xue, S. W. Homans, and S. E. Radford, “Systematic analysis of nucleation-dependent polymerization reveals new insights into the mechanism of amyloid self-assembly,” *Proc. Natl. Acad. Sci. U. S. A.*, vol. 105, no. 26, pp. 8926–8931, Jul. 2008,

[56] X. Wang, D. R. Smith, J. W. Jones, and M. R. Chapman, “In vitro polymerization of a functional Escherichia coli amyloid protein,” *J. Biol. Chem.*, vol. 282, no. 6, pp. 3713–3719, Jan. 2007,

[57] R. N. Rambaran and L. C. Serpell, “Amyloid fibrils: Abnormal protein assembly,” *Prion*, vol. 2, no. 3, p. 112, 2008,

[58] M. R. Nilsson, “Techniques to study amyloid fibril formation in vitro,” *Methods*, vol. 34, no. 1, pp. 151–160, Sep. 2004,

[59] W. B. Stine, L. Jungbauer, C. Yu, and M. J. Ladu, “Preparing synthetic Aβ in different aggregation states,” *Methods Mol. Biol.*, vol. 670, pp. 13–32, 2011,

[60] T. M. Ryan, J. Caine, H. D. T. Mertens, N. Kirby, J. Nigro, K. Breheney, L. J. Waddington, V.

A. Streltsov, C. Curtain, C. L. Masters, and B. R. Roberts, “Ammonium hydroxide treatment of Aβ produces an aggregate free solution suitable for biophysical and cell culture characterization,” *PeerJ*, vol. 2013, no. 1, p. e73, May 2013,

[61] J. P. Schmittschmitt and J. M. Scholtz, “The role of protein stability, solubility, and net charge in amyloid fibril formation,” *Protein Sci.*, vol. 12, no. 10, p. 2374, Jan. 2003,

[62] G. Wei and J. E. Shea, “Effects of Solvent on the Structure of the Alzheimer Amyloid-β(25–35) Peptide,” *Biophys. J.*, vol. 91, no. 5, p. 1638, Sep. 2006,

[63] A. Berkessel, J. A. Adrio, D. Hüttenhain, and J. M. Neudörfl, “Unveiling the ‘booster effect’ of fluorinated alcohol solvents: Aggregation-induced conformational changes and cooperatively enhanced H-bonding,” *J. Am. Chem. Soc.*, vol. 128, no. 26, pp. 8421–8426, Jul. 2006,

[64] B. Zhang-Haagen, R. Biehl, L. Nagel-Steger, A. Radulescu, D. Richter, and D. Willbold, “Monomeric Amyloid Beta Peptide in Hexafluoroisopropanol Detected by Small Angle Neutron Scattering,” *PLoS One*, vol. 11, no. 2, Feb. 2016,

[65] P. Cioni and G. B. Strambini, “Effect of heavy water on protein flexibility,” *Biophys. J.*, vol. 82, no. 6, pp. 3246–3253, 2002,

[66] T. Clark, J. Heske, and T. D. Kühne, “Opposing Electronic and Nuclear Quantum Effects on Hydrogen Bonds in H2O and D2O,” *ChemPhysChem*, vol. 20, no. 19, pp. 2461–2465, Oct. 2019,

[67] B. Kaplan, S. Yakar, Y. Balta, M. Pras, and B. Martin, “Isolation and purification of two major serum amyloid A isotypes SAA1 and SAA2 from the acute phase plasma of mice,” *J. Chromatogr. B Biomed. Sci. Appl.*, vol. 704, no. 1–2, pp. 69–76, Dec. 1997,

[68] S. Vivekanandan, J. R. Brender, S. Y. Lee, and A. Ramamoorthy, “A partially folded structure of amyloid-beta(1-40) in an aqueous environment,” *Biochem. Biophys. Res. Commun.*, vol. 411, no. 2, pp. 312–316, Jul. 2011,

[69] Y. Li, W. Xu, Y. Mu, and J. Z. H. Zhang, “Acidic pH retards the fibrillization of human Islet Amyloid Polypeptide due to electrostatic repulsion of histidines,” *J. Chem. Phys.*, vol. 139, no. 5, Aug. 2013,

[70] E. Smirnova, I. Safenkova, V. Stein-Margolina, V. Shubin, V. Polshakov, and B. Gurvits, “pH-responsive modulation of insulin aggregation and structural transformation of the aggregates,” *Biochimie*, vol. 109, pp. 49–59, 2015,

[71] T. J. Zbacnik, R. E. Holcomb, D. S. Katayama, B. M. Murphy, R. W. Payne, R. C. Coccaro, G. J. Evans, J. E. Matsuura, C. S. Henry, and M. C. Manning, “Role of Buffers in Protein Formulations,” *J. Pharm. Sci.*, vol. 106, no. 3, pp. 713–733, Mar. 2017,

[72] H. Chaaban, J. J. Vallooran, M. Van De Weert, and V. Foderà, “Ion-Mediated Morphological Diversity in Protein Amyloid Systems,” *J. Phys. Chem. Lett.*, vol. 13, no. 16, pp. 3586–3593,

Apr. 2022,

[73] B. Kang, H. Tang, Z. Zhao, and S. Song, “Hofmeister Series: Insights of Ion Specificity from Amphiphilic Assembly and Interface Property,” *ACS Omega*, vol. 5, no. 12, pp. 6229–6239, Mar. 2020,

[74] K. Garajová, A. Balogová, E. Dušeková, D. Sedláková, E. Sedlák, and R. Varhač, “Correlation of lysozyme activity and stability in the presence of Hofmeister series anions,” *Biochim. Biophys. acta. Proteins proteomics*, vol. 1865, no. 3, pp. 281–288, Mar. 2017,

[75] P. J. Marek, V. Patsalo, D. F. Green, and D. P. Raleigh, “Ionic strength effects on amyloid formation by amylin are a complicated interplay among Debye screening, ion selectivity, and Hofmeister effects,” *Biochemistry*, vol. 51, no. 43, pp. 8478–8490, Oct. 2012,

[76] K. Mikalauskaite, M. Ziaunys, T. Sneideris, and V. Smirnovas, “Effect of Ionic Strength on Thioflavin-T Affinity to Amyloid Fibrils and Its Fluorescence Intensity,” *Int. J. Mol. Sci. 2020, Vol. 21, Page 8916*, vol. 21, no. 23, p. 8916, Nov. 2020,

[77] K. Klement, K. Wieligmann, J. Meinhardt, P. Hortschansky, W. Richter, and M. Fändrich, “Effect of different salt ions on the propensity of aggregation and on the structure of Alzheimer’s abeta(1-40) amyloid fibrils,” *J. Mol. Biol.*, vol. 373, no. 5, pp. 1321–1333, Nov. 2007,

[78] W. Wang, S. Nema, and D. Teagarden, “Protein aggregation—Pathways and influencing factors,” *Int. J. Pharm.*, vol. 390, no. 2, pp. 89–99, May 2010,

[79] E. van Dijk, A. Hoogeveen, and S. Abeln, “The Hydrophobic Temperature Dependence of Amino Acids Directly Calculated from Protein Structures,” *PLoS Comput. Biol.*, vol. 11, no. 5, p. 1004277, May 2015,

[80] M. F. Pignataro, M. G. Herrera, and V. I. Dodero, “Evaluation of Peptide/Protein Self-Assembly and Aggregation by Spectroscopic Methods,” *Mol. 2020, Vol. 25, Page 4854*, vol. 25, no. 20, p. 4854, Oct. 2020,

[81] J. M. Andrews, W. F. Weiss IV, and C. J. Roberts, “Nucleation, growth, and activation energies for seeded and unseeded aggregation of α-chymotrypsinogen A,” *Biochemistry*, vol. 47, no. 8, pp. 2397–2403, Feb. 2008,

[82] W. Wang and C. J. Roberts, “Non-Arrhenius Protein Aggregation,” *AAPS J.*, vol. 15, no. 3, p. 840, Jul. 2013,

[83] S. Y. Ow and D. E. Dunstan, “The effect of concentration, temperature and stirring on hen egg white lysozyme amyloid formation,” *Soft Matter*, vol. 9, no. 40, pp. 9692–9701, Oct. 2013,

[84] E. El-Bastawissy, M. H. Knaggs, and I. H. Gilbert, “Molecular dynamics simulations of wild-type and point mutation human prion protein at normal and elevated temperature,” *J. Mol. Graph. Model.*, vol. 20, no. 2, pp. 145–154, 2001,

[85] A. A. Noorani, H. Yamashita, Y. Gao, S. Islam, Y. Sun, T. Nakamura, H. Enomoto, K. Zou, and M. Michikawa, “High temperature promotes amyloid β-protein production and γ-secretase complex formation via Hsp90,” *J. Biol. Chem.*, vol. 295, no. 52, pp. 18010–18022, Dec. 2020,

[86] D. J. Irwin, V. M. Y. Lee, and J. Q. Trojanowski, “Parkinson’s disease dementia: convergence of α-synuclein, tau and amyloid-β pathologies,” *Nat. Rev. Neurosci. 2013 149*, vol. 14, no. 9, pp. 626–636, Jul. 2013,

[87] R. Morales, I. Moreno-Gonzalez, and C. Soto, “Cross-seeding of misfolded proteins: implications for etiology and pathogenesis of protein misfolding diseases,” *PLoS Pathog.*, vol. 9, no. 9, Sep. 2013,

[88] B. Ren, Y. Zhang, M. Zhang, Y. Liu, D. Zhang, X. Gong, Z. Feng, J. Tang, Y. Chang, and J. Zheng, “Fundamentals of cross-seeding of amyloid proteins: an introduction,” *J. Mater. Chem. B*, vol. 7, no. 46, pp. 7267–7282, Nov. 2019,

[89] J. D. Harper and P. T. Lansbury, “MODELS OF AMYLOID SEEDING IN ALZHEIMER’S DISEASE AND SCRAPIE: Mechanistic Truths and Physiological Consequences of the Time-Dependent Solubility of Amyloid Proteins,” *Annu. Rev. Biochem.*, vol. 66, pp. 385–407, 1997,

[90] T. Härd and C. Lendel, “Inhibition of Amyloid Formation,” *J. Mol. Biol.*, vol. 421, no. 4–5, pp. 441–465, Aug. 2012,

[91] K. Ono, R. Takahashi, T. Ikeda, and M. Yamada, “Cross-seeding effects of amyloid β-protein and α-synuclein,” *J. Neurochem.*, vol. 122, no. 5, pp. 883–890, Sep. 2012,

[92] T. Tripathi and H. Khan, “Direct Interaction between the β-Amyloid Core and Tau Facilitates Cross-Seeding: A Novel Target for Therapeutic Intervention,” *Biochemistry*, vol. 59, no. 4, pp. 341–342, Feb. 2020,

[93] M. Zhang, R. Hu, B. Ren, H. Chen, B. Jiang, J. Ma, and J. Zheng, “Molecular Understanding of Aβ-hIAPP Cross-Seeding Assemblies on Lipid Membranes,” *ACS Chem. Neurosci.*, vol. 8, no. 3, pp. 524–537, Mar. 2017,

[94] M. Burdukiewicz, D. Rafacz, A. Barbach, K. Hubicka, L. Bąkała, A. Lassota, J. Stecko, N. Szymańska, J. W. Wojciechowski, D. Kozakiewicz, N. Szulc, J. Chilimoniuk, I. Jęśkowiak, M. Gąsior-Głogowska, and M. Kotulska, “AmyloGraph: a comprehensive database of amyloid–amyloid interactions,” *Nucleic Acids Res.*, vol. 51, no. D1, pp. D352–D357, Jan. 2023,

[95] D. A. Stakos, K. Stamatelopoulos, D. Bampatsias, M. Sachse, E. Zormpas, N. I. Vlachogiannis, S. Tual-Chalot, and K. Stellos, “The Alzheimer’s Disease Amyloid-Beta Hypothesis in Cardiovascular Aging and Disease: JACC Focus Seminar,” *J. Am. Coll. Cardiol.*, vol. 75, no. 8, pp. 952–967, Mar. 2020,

[96] P. H. St George-Hyslop and A. Petit, “Molecular biology and genetics of Alzheimer’s disease,”

C. R. Biol., vol. 328, no. 2, pp. 119–130, 2005,

[97] C. Aisenbrey, T. Borowik, R. Byström, M. Bokvist, F. Lindström, H. Misiak, M. A. Sani, and G. Gröbner, "How is protein aggregation in amyloidogenic diseases modulated by biological membranes?," *Eur. Biophys. J.*, vol. 37, no. 3, pp. 247–255, Mar. 2008,

[98] R. P. Friedland and M. R. Chapman, "The role of microbial amyloid in neurodegeneration," *PLoS Pathog.*, vol. 13, no. 12, Dec. 2017,

[99] K. Kowalski and A. Mulak, "Brain-Gut-Microbiota Axis in Alzheimer's Disease," *J. Neurogastroenterol. Motil.*, vol. 25, no. 1, p. 48, Jan. 2019,

[100] Y. Liu, B. Ren, Y. Zhang, Y. Sun, Y. Chang, G. Liang, L. Xu, and J. Zheng, "Molecular simulation aspects of amyloid peptides at membrane interface," *Biochim. Biophys. Acta - Biomembr.*, vol. 1860, no. 9, pp. 1906–1916, Sep. 2018,

[101] G. Musteikytė, A. K. Jayaram, C. K. Xu, M. Vendruscolo, G. Krainer, and T. P. J. Knowles, "Interactions of α-synuclein oligomers with lipid membranes," *Biochim. Biophys. acta. Biomembr.*, vol. 1863, no. 4, Apr. 2021,

[102] M. Pannuzzo, A. Raudino, D. Milardi, C. La Rosa, and M. Karttunen, "α-Helical Structures Drive Early Stages of Self-Assembly of Amyloidogenic Amyloid Polypeptide Aggregate Formation in Membranes," *Sci. Reports 2013 31*, vol. 3, no. 1, pp. 1–10, Sep. 2013,

[103] J. A. Hardy and G. A. Higgins, "Alzheimer's disease: the amyloid cascade hypothesis," *Science*, vol. 256, no. 5054, pp. 184–185, 1992,

[104] J. K. Kasim, I. Kavianinia, P. W. R. Harris, and M. A. Brimble, "Three decades of amyloid beta synthesis: Challenges and advances," *Front. Chem.*, vol. 7, no. JUL, p. 472, Jul. 2019,

[105] S. K. Ng, K. L. Nyam, I. A. Nehdi, G. H. Chong, O. M. Lai, and C. P. Tan, "Impact of stirring speed on β-lactoglobulin fibril formation," *Food Sci. Biotechnol.*, vol. 25, no. Suppl 1, p. 15, Mar. 2016,

[106] T. R. Heyn, J. Mayer, H. R. Neumann, C. Selhuber-Unkel, A. Kwade, K. Schwarz, and J. K. Keppler, "The threshold of amyloid aggregation of beta-lactoglobulin: Relevant factor combinations," *J. Food Eng.*, vol. 283, p. 110005, Oct. 2020,

[107] K. L. Zapadka, F. J. Becher, A. L. Gomes dos Santos, and S. E. Jackson, "Factors affecting the physical stability (aggregation) of peptide therapeutics," *Interface Focus*, vol. 7, no. 6. Interface Focus, Dec. 06, 2017.

[108] X. Ye, M. S. Hedenqvist, M. Langton, and C. Lendel, "On the role of peptide hydrolysis for fibrillation kinetics and amyloid fibril morphology," *RSC Adv.*, vol. 8, no. 13, pp. 6915–6924, Feb. 2018,

[109] S. Raccosta, M. Manno, D. Bulone, D. Giacomazza, V. Militello, V. Martorana, and P. L. San

Biagio, “Irreversible gelation of thermally unfolded proteins: Structural and mechanical properties of lysozyme aggregates,” *Eur. Biophys. J.*, vol. 39, no. 6, pp. 1007–1017, May 2010,

[110] T. Sneideris, D. Darguzis, A. Botyriute, M. Grigaliunas, R. Winter, and V. Smirnovas, “pH-Driven Polymorphism of Insulin Amyloid-Like Fibrils,” *PLoS One*, vol. 10, no. 8, p. e0136602, Aug. 2015,

[111] M. E. Gąsior-Głogowska, N. Szulc, and M. Szefczyk, “Challenges in Experimental Methods,” *Methods Mol. Biol.*, vol. 2340, pp. 281–307, 2022,

[112] A. Micsonai, F. Wien, L. Kernya, Y. H. Lee, Y. Goto, M. Réfrégiers, and J. Kardos, “Accurate secondary structure prediction and fold recognition for circular dichroism spectroscopy,” *Proc. Natl. Acad. Sci. U. S. A.*, vol. 112, no. 24, pp. E3095–E3103, Jun. 2015,

[113] B. Ren, R. Hu, M. Zhang, Y. Liu, L. Xu, B. Jiang, J. Ma, B. Ma, R. Nussinov, and J. Zheng, “Experimental and Computational Protocols for Studies of Cross-Seeding Amyloid Assemblies,” *Methods Mol. Biol.*, vol. 1777, p. 429, 2018,

[114] A. Barth, “Infrared spectroscopy of proteins,” *Biochim. Biophys. Acta - Bioenerg.*, vol. 1767, no. 9, pp. 1073–1101, Sep. 2007,

[115] J. M. Ruysschaert and V. Raussens, “ATR-FTIR Analysis of Amyloid Proteins,” *Methods Mol. Biol.*, vol. 1777, pp. 69–81, 2018,

[116] S. D. Moran and M. T. Zanni, “How to get insight into amyloid structure and formation from infrared spectroscopy,” *J. Phys. Chem. Lett.*, vol. 5, no. 11, pp. 1984–1993, Jun. 2014,

[117] M. E. Goldberg and A. F. Chaffotte, “Undistorted structural analysis of soluble proteins by attenuated total reflectance infrared spectroscopy,” *Protein Sci.*, vol. 14, no. 11, p. 2781, Nov. 2005,

[118] R. Sarroukh, E. Goormaghtigh, J. M. Ruysschaert, and V. Raussens, “ATR-FTIR: a ‘rejuvenated’ tool to investigate amyloid proteins,” *Biochim. Biophys. Acta*, vol. 1828, no. 10, pp. 2328–2338, 2013,

[119] K. Bagińska, J. Makowska, W. Wiczk, F. Kasprzykowski, and L. Chmurzyński, “Conformational studies of alanine-rich peptide using CD and FTIR spectroscopy,” *J. Pept. Sci.*, vol. 14, no. 3, pp. 283–289, Mar. 2008,

[120] J. Kong and S. Yu, “Fourier Transform Infrared Spectroscopic Analysis of Protein Secondary Structures,” *Acta Biochim. Biophys. Sin. (Shanghai).*, vol. 39, no. 8, pp. 549–559, Aug. 2007,

[121] S. A. Tatulian, “FTIR Analysis of Proteins and Protein-Membrane Interactions,” *Methods Mol. Biol.*, vol. 2003, pp. 281–325, 2019,

[122] J. Seo, W. Hoffmann, S. Warnke, X. Huang, S. Gewinner, W. Schöllkopf, M. T. Bowers, G. Von Helden, and K. Pagel, “An infrared spectroscopy approach to follow β-sheet formation in

peptide amyloid assemblies," *Nat. Chem. 2016 91*, vol. 9, no. 1, pp. 39–44, Sep. 2016,

[123] T. Šneideris, L. Baranauskiene, J. G. Cannon, R. Rutkiene, R. Meškys, and V. Smirnovas, "Looking for a generic inhibitor of amyloid-like fibril formation among flavone derivatives," *PeerJ*, vol. 2015, no. 9, 2015,

[124] N. Szulc, M. Burdukiewicz, M. Gąsior-Głogowska, J. W. Wojciechowski, J. Chilimoniuk, P. Mackiewicz, T. Šneideris, V. Smirnovas, and M. Kotulska, "Bioinformatics methods for identification of amyloidogenic peptides show robustness to misannotated training data," *Sci. Reports 2021 111*, vol. 11, no. 1, pp. 1–11, Apr. 2021,

[125] I. Ettah and L. Ashton, "Engaging with Raman Spectroscopy to Investigate Antibody Aggregation," *Antibodies 2018, Vol. 7, Page 24*, vol. 7, no. 3, p. 24, Jul. 2018,

[126] B. Li, R. Zhang, and X. Shi, "Aggregation of amyloid peptides into fibrils driven by nanoparticles and their curvature effect," *Phys. Chem. Chem. Phys.*, vol. 21, no. 4, pp. 1784–1790, Jan. 2019,

[127] S. Wang, J. Zhao, H. Lui, Q. He, and H. Zeng, "A modular Raman microspectroscopy system for biological tissue analysis," *Spectroscopy*, vol. 24, pp. 577–583, 2010,

[128] D. Kurouski, T. Postiglione, T. Deckert-Gaudig, V. Deckert, and I. K. Lednev, "Amide I vibrational mode suppression in surface (SERS) and tip (TERS) enhanced Raman spectra of protein specimens," *Analyst*, vol. 138, no. 6, p. 1665, Mar. 2013,

[129] H. Takeuchi, "Raman structural markers of tryptophan and histidine side chains in proteins," *Biopolymers*, vol. 72, no. 5, pp. 305–317, Jan. 2003,

[130] J. D. Flynn and J. C. Lee, "Raman fingerprints of amyloid structures," *Chem. Commun.*, vol. 54, no. 51, pp. 6983–6986, Jun. 2018,

[131] D. Kurouski, R. P. Van Duyne, and I. K. Lednev, "Exploring the structure and formation mechanism of amyloid fibrils by Raman spectroscopy: a review," *Analyst*, vol. 140, no. 15, pp. 4967–4980, Jul. 2015,

[132] M. Baranska, "Optical spectroscopy and computational methods in biology and medicine," *Opt. Spectrosc. Comput. Methods Biol. Med.*, pp. 1–540, Jan. 2014,

[133] M. Ishigaki, K. Morimoto, E. Chatani, and Y. Ozaki, "Exploration of Insulin Amyloid Polymorphism Using Raman Spectroscopy and Imaging," *Biophys. J.*, vol. 118, no. 12, pp. 2997–3007, Jun. 2020,

[134] D. Partouche, V. Militello, A. Gomez-Zavaglia, F. Wien, C. Sandt, and V. Arluison, "In Situ Characterization of Hfq Bacterial Amyloid: A Fourier-Transform Infrared Spectroscopy Study," *Pathog. 2019, Vol. 8, Page 36*, vol. 8, no. 1, p. 36, Mar. 2019,

[135] M. Biancalana and S. Koide, "Molecular mechanism of Thioflavin-T binding to amyloid fibrils,"

Biochim. Biophys. Acta, vol. 1804, no. 7, pp. 1405–1412, Jul. 2010,

[136] R. Sabaté and S. J. Saupe, "Thioflavin T fluorescence anisotropy: An alternative technique for the study of amyloid aggregation," *Biochem. Biophys. Res. Commun.*, vol. 360, no. 1, pp. 135–138, Aug. 2007,

[137] O. Galzitskaya, "New Mechanism of Amyloid Fibril Formation," *Curr. Protein Pept. Sci.*, vol. 20, no. 6, pp. 630–640, Jan. 2019,

[138] J. J. Bozzola and L. D. Russell, "Electron microscopy : principles and techniques for biologists," p. 542, 1992.

[139] F. S. Ruggeri, T. Šneideris, M. Vendruscolo, and T. P. J. Knowles, "Atomic force microscopy for single molecule characterisation of protein aggregation," *Arch. Biochem. Biophys.*, vol. 664, pp. 134–148, Mar. 2019,

[140] M. Charnley, J. Gilbert, O. G. Jones, and N. P. Reynolds, "Characterization of Amyloid Fibril Networks by Atomic Force Microscopy," *Bio-protocol*, vol. 8, no. 4, Feb. 2018,

[141] M. Azouz, C. Cullin, S. Lecomte, and M. Lafleur, "Membrane domain modulation of Aβ1–42 oligomer interactions with supported lipid bilayers: an atomic force microscopy investigation," *Nanoscale*, vol. 11, no. 43, pp. 20857–20867, Nov. 2019,

[142] C. Goldsbury and J. Green, "Time-lapse atomic force microscopy in the characterization of amyloid-like fibril assembly and oligomeric intermediates," *Methods Mol. Biol.*, vol. 299, pp. 103–128, 2005,

[143] J. Colchero, A. Gil, and A. M. Baró, "Resolution enhancement and improved data interpretation in electrostatic force microscopy," *Phys. Rev. B*, vol. 64, no. 24, p. 245403, Nov. 2001,

[144] W. A. Zisman, "A new method of measuring contact potential differences in metals," *Rev. Sci. Instrum.*, vol. 3, no. 7, p. 367, Dec. 2004,

[145] I. D. Baikie, A. C. Grain, J. Sutherland, and J. Law, "Dual Mode Kelvin Probe: Featuring Ambient Pressure Photoemission Spectroscopy and Contact Potential Difference," *Energy Procedia*, vol. 60, no. C, pp. 48–56, Jan. 2014,

[146] G. Lee, W. Lee, H. Lee, S. Woo Lee, D. Sung Yoon, K. Eom, and T. Kwon, "Mapping the surface charge distribution of amyloid fibril," *Appl. Phys. Lett.*, vol. 101, no. 4, p. 043703, Jul. 2012,

[147] W. Lee, H. Lee, Y. Choi, K. S. Hwang, S. W. Lee, G. Lee, and D. S. Yoon, "Nanoelectrical characterization of amyloid-β 42 aggregates via Kelvin probe force microscopy," *Macromol. Res.*, vol. 25, no. 12, pp. 1187–1191, Dec. 2017,

[148] M. Kotulska and J. W. Wojciechowski, "Bioinformatics Methods in Predicting Amyloid Propensity of Peptides and Proteins," *Methods Mol. Biol.*, vol. 2340, pp. 1–15, 2022,

[149] I. Walsh, F. Seno, S. C. E. Tosatto, and A. Trovato, “PASTA 2.0: An improved server for protein aggregation prediction,” *Nucleic Acids Res.*, vol. 42, no. W1, p. W301, Jul. 2014,

[150] S. O. Garbuzynskiy, M. Y. Lobanov, and O. V. Galzitskaya, “FoldAmyloid: a method of prediction of amyloidogenic regions from protein sequence,” *Bioinformatics*, vol. 26, no. 3, pp. 326–332, Feb. 2010,

[151] R. Zambrano, M. Jamroz, A. Szczasiuk, J. Pujols, S. Kmiecik, and S. Ventura, “AGGRESCAN3D (A3D): server for prediction of aggregation properties of protein structures,” *Nucleic Acids Res.*, vol. 43, no. W1, pp. W306–W313, Jul. 2015,

[152] A. R. Leach, “Molecular modelling : principles and applications,” p. 595, 1996,

[153] Y. Sugita and Y. Okamoto, “Replica-exchange molecular dynamics method for protein folding,” *Chem. Phys. Lett.*, vol. 314, no. 1–2, pp. 141–151, Nov. 1999,

[154] Y. Mori and H. Okumura, “Simulated tempering based on global balance or detailed balance conditions: Suwa-Todo, heat bath, and Metropolis algorithms,” *J. Comput. Chem.*, vol. 36, no. 31, pp. 2344–2349, Dec. 2015,

[155] J. Huang, S. Rauscher, G. Nawrocki, T. Ran, M. Feig, B. L. De Groot, H. Grubmüller, and A. D. MacKerell, “CHARMM36m: an improved force field for folded and intrinsically disordered proteins,” *Nat. Methods 2016 141*, vol. 14, no. 1, pp. 71–73, Nov. 2016,

[156] A. Paul, S. Samantray, M. Anteghini, M. Khaled, and B. Strodel, “Thermodynamics and kinetics of the amyloid-β peptide revealed by Markov state models based on MD data in agreement with experiment,” *Chem. Sci.*, vol. 12, no. 19, pp. 6652–6669, May 2021,

[157] B. Strodel, “Amyloid aggregation simulations: challenges, advances and perspectives,” *Curr. Opin. Struct. Biol.*, vol. 67, pp. 145–152, Apr. 2021,

[158] M. López De La Paz and L. Serrano, “Sequence determinants of amyloid fibril formation,” *Proc. Natl. Acad. Sci. U. S. A.*, vol. 101, no. 1, pp. 87–92, Jan. 2004,

[159] M. J. Thompson, S. A. Sievers, J. Karanicolas, M. I. Ivanova, D. Baker, and D. Eisenberg, “The 3D profile method for identifying fibril-forming segments of proteins,” *Proc. Natl. Acad. Sci. U. S. A.*, vol. 103, no. 11, pp. 4074–4078, Mar. 2006,

[160] N. Louros, K. Konstantoulea, M. De Vleeschouwer, M. Ramakers, J. Schymkowitz, and F. Rousseau, “WALTZ-DB 2.0: an updated database containing structural information of experimentally determined amyloid-forming peptides,” *Nucleic Acids Res.*, vol. 48, no. D1, pp. D389–D393, Jan. 2020,

[161] X. Wang, N. D. Hammer, and M. R. Chapman, “The Molecular Basis of Functional Bacterial Amyloid Polymerization and Nucleation,” *J. Biol. Chem.*, vol. 283, no. 31, p. 21530, Aug. 2008,

[162] E. P. Debenedictis, D. Ma, and S. Keten, “Structural predictions for curli amyloid fibril subunits

CsgA and CsgB," *RSC Adv.*, vol. 7, no. 76, pp. 48102–48112, Oct. 2017,

[163] N. Szulc, M. Gasior-Głogowska, J. W. Wojciechowski, M. Szefczyk, A. M. Żak, M. Burdukiewicz, and M. Kotulska, "Variability of amyloid propensity in imperfect repeats of csga protein of salmonella enterica and escherichia coli," *Int. J. Mol. Sci.*, vol. 22, no. 10, p. 5127, May 2021,

[164] G. N. Derry, M. E. Kern, and E. H. Worth, "Recommended values of clean metal surface work functions," *J. Vac. Sci. Technol. A Vacuum, Surfaces, Film.*, vol. 33, no. 6, p. 060801, Oct. 2015,

[165] N. Barrett, O. Renault, H. Lemaître, P. Bonnaillie, F. Barcelo, F. Miserque, M. Wang, and C. Corbel, "Microscopic work function anisotropy and surface chemistry of 316L stainless steel using photoelectron emission microscopy," *J. Electron Spectros. Relat. Phenomena*, vol. 195, pp. 117–124, Aug. 2014,

[166] J. W. Wojciechowski and M. Kotulska, "PATH - Prediction of Amyloidogenicity by Threading and Machine Learning," *Sci. Reports 2020 101*, vol. 10, no. 1, pp. 1–9, May 2020,

[167] S. Maurer-Stroh, M. Debulpaep, N. Kuemmerer, M. L. De La Paz, I. C. Martins, J. Reumers, K. L. Morris, A. Copland, L. Serpell, L. Serrano, J. W. H. Schymkowitz, and F. Rousseau, "Exploring the sequence determinants of amyloid structure using position-specific scoring matrices," *Nat. Methods 2010 73*, vol. 7, no. 3, pp. 237–242, Feb. 2010,

[168] A. C. Tsolis, N. C. Papandreou, V. A. Iconomidou, and S. J. Hamodrakas, "A Consensus Method for the Prediction of 'Aggregation-Prone' Peptides in Globular Proteins," *PLoS One*, vol. 8, no. 1, p. e54175, Jan. 2013,

[169] M. Emily, A. Talvas, and C. Delamarche, "MetAmyl: A METa-Predictor for AMYLoid Proteins," *PLoS One*, vol. 8, no. 11, p. e79722, Nov. 2013,

[170] A. B. Ahmed, N. Znassi, M. T. Château, and A. V. Kajava, "A structure-based approach to predict predisposition to amyloidosis," *Alzheimers. Dement.*, vol. 11, no. 6, pp. 681–690, Jun. 2015,

[171] A. Bateman, M. J. Martin, S. Orchard, M. Magrane, R. Agivetova, S. Ahmad, E. Alpi, E. H. Bowler-Barnett, R. Britto, B. Bursteinas, H. Bye-A-Jee, R. Coetzee, A. Cukura, A. da Silva, P. Denny, T. Dogan, T. G. Ebenezer, J. Fan, L. G. Castro, P. Garmiri, G. Georghiou, L. Gonzales, E. Hatton-Ellis, A. Hussein, A. Ignatchenko, G. Insana, R. Ishtiaq, P. Jokinen, V. Joshi, D. Jyothi, A. Lock, R. Lopez, A. Luciani, J. Luo, Y. Lussi, A. MacDougall, F. Madeira, M. Mahmoudy, M. Menchi, A. Mishra, K. Moulang, A. Nightingale, C. S. Oliveira, S. Pundir, G. Qi, S. Raj, D. Rice, M. R. Lopez, R. Saidi, J. Sampson, T. Sawford, E. Speretta, E. Turner, N. Tyagi, P. Vasudev, V. Volynkin, K. Warner, X. Watkins, R. Zaru, H. Zellner, A. Bridge, S. Poux, N. Redaschi, L. Aimo, G. Argoud-Puy, A. Auchincloss, K. Axelsen, P. Bansal, D. Baratin,

M. C. Blatter, J. Bolleman, E. Boutet, L. Breuza, C. Casals-Casas, E. de Castro, K. C. Echioukh, E. Coudert, B. Cuche, M. Doche, D. Dornevil, A. Estreicher, M. L. Famiglietti, M. Feuermann, E. Gasteiger, S. Gehant, V. Gerritsen, A. Gos, N. Gruaz-Gumowski, U. Hinz, C. Hulo, N. Hyka-Nouspikel, F. Jungo, G. Keller, A. Kerhornou, V. Lara, P. Le Mercier, D. Lieberherr, T. Lombardot, X. Martin, P. Masson, A. Morgat, T. B. Neto, S. Paesano, I. Pedruzzi, S. Pilbout, L. Pourcel, M. Pozzato, M. Pruess, C. Rivoire, C. Sigrist, K. Sonesson, A. Stutz, S. Sundaram, M. Tognolli, L. Verbregue, C. H. Wu, C. N. Arighi, L. Arminski, C. Chen, Y. Chen, J. S. Garavelli, H. Huang, K. Laiho, P. McGarvey, D. A. Natale, K. Ross, C. R. Vinayaka, Q. Wang, Y. Wang, L. S. Yeh, J. Zhang, P. Ruch, and D. Teodoro, "UniProt: the universal protein knowledgebase in 2021," *Nucleic Acids Res.*, vol. 49, no. D1, pp. D480–D489, Jan. 2021,

[172] K. Katoh, K. Misawa, K. I. Kuma, and T. Miyata, "MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform," *Nucleic Acids Res.*, vol. 30, no. 14, pp. 3059–3066, Jul. 2002,

[173] K. Katoh and D. M. Standley, "MAFFT multiple sequence alignment software version 7: improvements in performance and usability," *Mol. Biol. Evol.*, vol. 30, no. 4, pp. 772–780, Apr. 2013,

[174] J. Jumper, R. Evans, A. Pritzel, T. Green, M. Figurnov, O. Ronneberger, K. Tunyasuvunakool, R. Bates, A. Žídek, A. Potapenko, A. Bridgland, C. Meyer, S. A. A. Kohl, A. J. Ballard, A. Cowie, B. Romera-Paredes, S. Nikolov, R. Jain, J. Adler, T. Back, S. Petersen, D. Reiman, E. Clancy, M. Zielinski, M. Steinegger, M. Pacholska, T. Berghammer, S. Bodenstein, D. Silver, O. Vinyals, A. W. Senior, K. Kavukcuoglu, P. Kohli, and D. Hassabis, "Highly accurate protein structure prediction with AlphaFold," *Nat. 2021 5967873*, vol. 596, no. 7873, pp. 583–589, Jul. 2021,

[175] M. Mirdita, K. Schütze, Y. Moriwaki, L. Heo, S. Ovchinnikov, and M. Steinegger, "ColabFold: making protein folding accessible to all," *Nat. Methods 2022 196*, vol. 19, no. 6, pp. 679–682, May 2022,

[176] V. Mariani, M. Biasini, A. Barbato, and T. Schwede, "lDDT: a local superposition-free score for comparing protein structures and models using distance difference tests," *Bioinformatics*, vol. 29, no. 21, p. 2722, Nov. 2013,

[177] E. Jurrus, D. Engel, K. Star, K. Monson, J. Brandi, L. E. Felberg, D. H. Brookes, L. Wilson, J. Chen, K. Liles, M. Chun, P. Li, D. W. Gohara, T. Dolinsky, R. Konecny, D. R. Koes, J. E. Nielsen, T. Head-Gordon, W. Geng, R. Krasny, G. W. Wei, M. J. Holst, J. A. McCammon, and N. A. Baker, "Improvements to the APBS biomolecular solvation software suite," *Protein Sci.*, vol. 27, no. 1, pp. 112–128, Jan. 2018,

[178] J. Lee, D. S. Patel, J. Ståhle, S. J. Park, N. R. Kern, S. Kim, J. Lee, X. Cheng, M. A. Valvano, O. Holst, Y. A. Knirel, Y. Qi, S. Jo, J. B. Klauda, G. Widmalm, and W. Im, "CHARMM-GUI Membrane Builder for Complex Biological Membrane Simulations with Glycolipids and Lipoglycans," *J. Chem. Theory Comput.*, vol. 15, no. 1, pp. 775–786, Jan. 2019,

[179] P. Eastman, J. Swails, J. D. Chodera, R. T. McGibbon, Y. Zhao, K. A. Beauchamp, L. P. Wang, A. C. Simmonett, M. P. Harrigan, C. D. Stern, R. P. Wiewiora, B. R. Brooks, and V. S. Pande, "OpenMM 7: Rapid development of high performance algorithms for molecular dynamics," *PLOS Comput. Biol.*, vol. 13, no. 7, p. e1005659, Jul. 2017,

[180] J. B. Klauda, R. M. Venable, J. A. Freites, J. W. O'Connor, D. J. Tobias, C. Mondragon-Ramirez, I. Vorobyov, A. D. MacKerell, and R. W. Pastor, "Update of the CHARMM All-Atom Additive Force Field for Lipids: Validation on Six Lipid Types," *J. Phys. Chem. B*, vol. 114, no. 23, pp. 7830–7843, Jun. 2010,

[181] Z. Zhang, X. Liu, K. Yan, M. E. Tuckerman, and J. Liu, "Unified Efficient Thermostat Scheme for the Canonical Ensemble with Holonomic or Isokinetic Constraints via Molecular Dynamics," *J. Phys. Chem. A*, vol. 123, no. 28, pp. 6056–6079, May 2019,

[182] Y. Gao, J. Lee, I. P. S. Smith, H. Lee, S. Kim, Y. Qi, J. B. Klauda, G. Widmalm, S. Khalid, and W. Im, "CHARMM-GUI Supports Hydrogen Mass Repartitioning and Different Protonation States of Phosphates in Lipopolysaccharides," *J. Chem. Inf. Model.*, vol. 61, no. 2, pp. 831–839, Feb. 2021,

[183] K. H. Chow and D. M. Ferguson, "Isothermal-isobaric molecular dynamics simulations with Monte Carlo volume sampling," *Comput. Phys. Commun.*, vol. 91, no. 1–3, pp. 283–289, Sep. 1995,

[184] R. J. Gowers, M. Linke, J. Barnoud, T. J. E. Reddy, M. N. Melo, S. L. Seyler, J. Domanski, D. L. Dotson, S. Buchoux, I. M. Kenney, and O. Beckstein, "MDAnalysis: A Python Package for the Rapid Analysis of Molecular Dynamics Simulations," *Proc. 15th Python Sci. Conf.*, pp. 98–105, Sep. 2019,

[185] A. S. Urban, K. V. Pavlov, A. V. Kamynina, I. S. Okhrimenko, A. S. Arseniev, and E. V. Bocharov, "Structural studies providing insights into production and conformational behavior of amyloid-β peptide associated with Alzheimer's disease development," *Molecules*, vol. 26, no. 10, May 2021,

[186] A. Waterhouse, M. Bertoni, S. Bienert, G. Studer, G. Tauriello, R. Gumienny, F. T. Heer, T. A. P. De Beer, C. Rempfer, L. Bordoli, R. Lepore, and T. Schwede, "SWISS-MODEL: homology modelling of protein structures and complexes," *Nucleic Acids Res.*, vol. 46, no. W1, pp. W296–W303, Jul. 2018,

[187] M. J. Abraham, T. Murtola, R. Schulz, S. Páll, J. C. Smith, B. Hess, and E. Lindah, “Gromacs: High performance molecular simulations through multi-level parallelism from laptops to supercomputers,” *SoftwareX*, vol. 1–2, pp. 19–25, Sep. 2015,

[188] D. J. Evans and B. L. Holian, “The Nose–Hoover thermostat,” *J. Chem. Phys.*, vol. 83, no. 8, p. 4069, Aug. 1998,

[189] M. Parrinello and A. Rahman, “Polymorphic transitions in single crystals: A new molecular dynamics method,” *J. Appl. Phys.*, vol. 52, no. 12, p. 7182, Aug. 1998,

[190] Berk Hess, Henk Bekker, Herman J. C. Berendsen, and Johannes G. E. M. Fraaije, “LINCS: A linear constraint solver for molecular simulations - Hess - 1997 - Journal of Computational Chemistry - Wiley Online Library,” *J. Comput. Chem.*, 1998,

[191] T. Darden, D. York, and L. Pedersen, “Particle mesh Ewald: An N·log(N) method for Ewald sums in large systems,” *J. Chem. Phys.*, vol. 98, no. 12, p. 10089, Aug. 1998,

[192] W. Humphrey, A. Dalke, and K. Schulten, “VMD: Visual molecular dynamics,” *J. Mol. Graph.*, vol. 14, no. 1, pp. 33–38, 1996,

[193] R. P. R. Nanga, J. R. Brender, S. Vivekanandan, and A. Ramamoorthy, “Structure and membrane orientation of IAPP in its natively amidated form at physiological pH in a membrane environment,” *Biochim. Biophys. Acta*, vol. 1808, no. 10, pp. 2337–2342, Oct. 2011,

[194] C. Röder, T. Kupreichyk, L. Gremer, L. U. Schäfer, K. R. Pothula, R. B. G. Ravelli, D. Willbold, W. Hoyer, and G. F. Schröder, “Cryo-EM structure of islet amyloid polypeptide fibrils reveals similarities with amyloid-β fibrils,” vol. 27, no. 7, pp. 660–667, 2020,

[195] N. Benseny-Cases, M. Cócera, and J. Cladera, “Conversion of non-fibrillar β-sheet oligomers into amyloid fibrils in Alzheimer’s disease amyloid peptide aggregation,” *Biochem. Biophys. Res. Commun.*, vol. 361, no. 4, pp. 916–921, Oct. 2007,

[196] A. Krüger, A. Bürkle, A. Mangerich, and K. Hauser, “A combined approach of surface passivation and specific immobilization to study biomolecules by ATR-FTIR spectroscopy,” *Biomed. Spectrosc. Imaging*, vol. 7, no. 1–2, pp. 25–33, Jan. 2018,

[197] P. L. Tsiolaki, N. N. Louros, and V. A. Iconomidou, “Hexapeptide Tandem Repeats Dictate the Formation of Silkmoth Chorion, a Natural Protective Amyloid,” *J. Mol. Biol.*, vol. 430, no. 20, pp. 3774–3783, Oct. 2018,

[198] I. Cherny, L. Rockah, O. Levy-Nissenbaum, U. Gophna, E. Z. Ron, and E. Gazit, “The formation of Escherichia coli curli amyloid fibrils is mediated by prion-like peptide repeats,” *J. Mol. Biol.*, vol. 352, no. 2, pp. 245–252, Sep. 2005,

[199] U. Göbel, C. Sander, R. Schneider, and A. Valencia, “Correlated mutations and residue contacts in proteins,” *Proteins Struct. Funct. Bioinforma.*, vol. 18, no. 4, pp. 309–317, Apr. 1994,

[200] L. Sewell, F. Stylianou, Y. Xu, J. Taylor, L. Sefer, and S. Matthews, “NMR insights into the pre-amyloid ensemble and secretion targeting of the curli subunit CsgA,” *Sci. Reports 2020 101*, vol. 10, no. 1, pp. 1–10, May 2020,

[201] G. Devitt, A. Crisford, W. Rice, H. A. Weismiller, Z. Fan, C. Commins, B. T. Hyman, M. Margittai, S. Mahajan, and A. Mudher, “Conformational fingerprinting of tau variants and strains by Raman spectroscopy,” *RSC Adv.*, vol. 11, no. 15, pp. 8899–8915, Feb. 2021,

[202] M. Hoshi, M. Sato, S. Matsumoto, A. Noguchi, K. Yasutake, N. Yoshida, and K. Sato, “Spherical aggregates of beta-amyloid (amylospheroid) show high neurotoxicity and activate tau protein kinase I/glycogen synthase kinase-3beta,” *Proc. Natl. Acad. Sci. U. S. A.*, vol. 100, no. 11, pp. 6370–6375, May 2003,

[203] S. Gul, B. Bahadir, A. Dusak, M. Kalayci, N. Edebali, and B. Acikgoz, “Spherical amyloid deposition in a prolactin-producing pituitary adenoma,” *Neuropathology*, vol. 29, no. 1, pp. 81–84, Feb. 2009,

[204] D. R. Hinton, R. K. Polk, K. D. Linse, M. H. Weiss, K. Kovacs, and J. A. Garner, “Characterization of spherical amyloid protein from a prolactin-producing pituitary adenoma,” *Acta Neuropathol.*, vol. 93, no. 1, pp. 43–49, Dec. 1997,

[205] K. I. Yuyama, M. Ueda, S. Nagao, S. Hirota, T. Sugiyama, and H. Masuhara, “A Single Spherical Assembly of Protein Amyloid Fibrils Formed by Laser Trapping,” *Angew. Chem. Int. Ed. Engl.*, vol. 56, no. 24, pp. 6739–6743, Jun. 2017,

[206] K. A. Conway, S. J. Lee, J. C. Rochet, T. T. Ding, R. E. Williamson, and P. T. Lansbury, “Acceleration of oligomerization, not fibrillization, is a shared property of both alpha-synuclein mutations linked to early-onset Parkinson’s disease: implications for pathogenesis and therapy,” *Proc. Natl. Acad. Sci. U. S. A.*, vol. 97, no. 2, pp. 571–576, Jan. 2000,

[207] E. Tayeb-Fligelman, O. Tabachnikov, A. Moshe, O. Goldshmidt-Tran, M. R. Sawaya, N. Coquelle, J. P. Colletier, and M. Landau, “The cytotoxic Staphylococcus aureus PSMα3 reveals a cross-α amyloid-like fibril,” *Science (80-.).*, vol. 355, no. 6327, pp. 831–833, Feb. 2017,

[208] S. S. Strickler, A. V. Gribenko, A. V. Gribenko, T. R. Keiffer, J. Tomlinson, T. Reihle, V. V. Loladze, and G. I. Makhatadze, “Protein stability and surface electrostatics: A charged relationship,” *Biochemistry*, vol. 45, no. 9, pp. 2761–2766, Mar. 2006,

[209] S. Sokalingam, G. Raghunathan, N. Soundrarajan, and S. G. Lee, “A study on the effect of surface lysine to arginine mutagenesis on protein stability and structure using green fluorescent protein,” *PLoS One*, vol. 7, no. 7, Jul. 2012,

[210] H. X. Zhou and X. Pang, “Electrostatic Interactions in Protein Structure, Folding, Binding, and Condensation,” *Chem. Rev.*, vol. 118, no. 4, pp. 1691–1741, Feb. 2018,

[211] S. Brudar and B. Hribar-Lee, “Effect of buffer on protein stability in aqueous solutions: A simple protein aggregation model,” *J. Phys. Chem. B*, vol. 125, no. 10, pp. 2504–2512, Mar. 2021,

[212] L. Zongo, H. Lange, and C. Crestini, “A Study of the Effect of Kosmotropic and Chaotropic Ions on the Release Characteristics of Lignin Microcapsules under Stimuli-Responsive Conditions,” *ACS Omega*, vol. 4, no. 4, pp. 6979–6993, Apr. 2019,

[213] K. Sikora, M. Jaśkiewicz, D. Neubauer, D. Migoń, and W. Kamysz, “The Role of Counter-Ions in Peptides-An Overview,” *Pharmaceuticals (Basel).*, vol. 13, no. 12, pp. 1–26, Dec. 2020,

[214] P. Westermark, A. Andersson, and G. T. Westermark, “Islet amyloid polypeptide, islet amyloid, and diabetes mellitus,” *Physiol. Rev.*, vol. 91, no. 3, pp. 795–826, Jul. 2011,

[215] P. Bharadwaj, T. Solomon, B. R. Sahoo, K. Ignasiak, S. Gaskin, J. Rowles, G. Verdile, M. J. Howard, C. S. Bond, A. Ramamoorthy, R. N. Martins, and P. Newsholme, “Amylin and beta amyloid proteins interact to form amorphous heterocomplexes with enhanced toxicity in neuronal cells,” *Sci. Reports 2020 101*, vol. 10, no. 1, pp. 1–14, Jun. 2020,

[216] N. Schultz, S. Janelidze, E. Byman, L. Minthon, K. Nägga, O. Hansson, and M. Wennström, “Levels of islet amyloid polypeptide in cerebrospinal fluid and plasma from patients with Alzheimer’s disease,” *PLoS One*, vol. 14, no. 6, p. e0218561, Jun. 2019,

[217] E. Andreetto, L. M. Yan, M. Tatarek-Nossol, A. Velkova, R. Frank, and A. Kapurniotu, “Identification of hot regions of the Abeta-IAPP interaction interface as high-affinity binding sites in both cross- and self-association,” *Angew. Chem. Int. Ed. Engl.*, vol. 49, no. 17, pp. 3081–3085, Apr. 2010,

[218] J. Miklossy, H. Qing, A. Radenovic, A. Kis, B. Vileno, F. Làszló, L. Miller, R. N. Martins, G. Waeber, V. Mooser, F. Bosman, K. Khalili, N. Darbinian, and P. L. McGeer, “Beta amyloid and hyperphosphorylated tau deposits in the pancreas in type 2 diabetes,” *Neurobiol. Aging*, vol. 31, no. 9, pp. 1503–1515, Sep. 2010,

[219] D. Siniscalco, G. Francius, M. Tarek, S. K. Bali, O. Laprévote, C. Malaplate, T. Oster, L. Pauron, and F. Quilès, “Molecular Insights for Alzheimer’s Disease: An Unexplored Storyline on the Nanoscale Impact of Nascent Aβ1–42 toward the Lipid Membrane,” *ACS Appl. Mater. Interfaces*, Mar. 2023,

[220] J. J. Balbach, Y. Ishii, O. N. Antzutkin, R. D. Leapman, N. W. Rizzo, F. Dyda, J. Reed, and R. Tycko, “Amyloid Fibril Formation by Aβ16-22, a Seven-Residue Fragment of the Alzheimer’s β-Amyloid Peptide, and Structural Characterization by Solid State NMR†,” *Biochemistry*, vol. 39, no. 45, pp. 13748–13759, Nov. 2000,

[221] M. A. Wälti, F. Ravotti, H. Arai, C. G. Glabe, J. S. Wall, A. Böckmann, P. Güntert, B. H. Meier, and R. Riek, “Atomic-resolution structure of a disease-relevant Aβ(1-42) amyloid fibril,” *Proc.*

Natl. Acad. Sci. U. S. A., vol. 113, no. 34, pp. E4976–E4984, Aug. 2016,

[222] S. J. Moore, K. Sonar, P. Bharadwaj, E. Deplazes, and R. L. Mancera, "Characterisation of the Structure and Oligomerisation of Islet Amyloid Polypeptides (IAPP): A Review of Molecular Dynamics Simulation Studies," *Mol. A J. Synth. Chem. Nat. Prod. Chem.*, vol. 23, no. 9, Aug. 2018,

[223] Q. Cao, D. R. Boyer, M. R. Sawaya, P. Ge, and D. S. Eisenberg, "Cryo-EM structure and inhibitor design of human IAPP (amylin) fibrils," *Nat. Struct. Mol. Biol. 2020 277*, vol. 27, no. 7, pp. 653–659, Jun. 2020,

[224] S. Gilead and E. Gazit, "The Role of the 14–20 Domain of the Islet Amyloid Polypeptide in Amyloid Formation," *Exp. Diabetes Res.*, vol. 2008, p. 256954, 2008,

[225] E. T. A. S. Jaikaran, C. E. Higham, L. C. Serpell, J. Zurdo, M. Gross, A. Clark, and P. E. Fraser, "Identification of a novel human islet amyloid polypeptide beta-sheet domain and factors influencing fibrillogenesis," *J. Mol. Biol.*, vol. 308, no. 3, pp. 515–525, May 2001,

[226] M. C. Owen, D. Gnutt, M. Gao, S. K. T. S. Wärmländer, J. Jarvet, A. Gräslund, R. Winter, S. Ebbinghaus, and B. Strodel, "Effects of in vivo conditions on amyloid aggregation," *Chem. Soc. Rev.*, vol. 48, no. 14, pp. 3946–3996, Jul. 2019,

[227] R. Gallardo, M. G. Iadanza, Y. Xu, G. R. Heath, R. Foster, S. E. Radford, and N. A. Ranson, "Fibril structures of diabetes-related amylin variants reveal a basis for surface-templated assembly," *Nat. Struct. Mol. Biol. 2020 2711*, vol. 27, no. 11, pp. 1048–1056, Sep. 2020,

[228] S. Samantray, F. Yin, B. Kav, and B. Strodel, "Different Force Fields Give Rise to Different Amyloid Aggregation Pathways in Molecular Dynamics Simulations," *J. Chem. Inf. Model.*, vol. 60, no. 12, pp. 6462–6475, Dec. 2020,

[229] J. Seeliger, F. Evers, C. Jeworrek, S. Kapoor, K. Weise, E. Andreetto, M. Tolan, A. Kapurniotu, and R. Winter, "Cross-amyloid interaction of Aβ and IAPP at lipid membranes," *Angew. Chem. Int. Ed. Engl.*, vol. 51, no. 3, pp. 679–683, Jan. 2012,

[230] C. La Rosa, M. Condorelli, G. Compagnini, F. Lolicato, D. Milardi, T. N. Do, M. Karttunen, M. Pannuzzo, A. Ramamoorthy, F. Fraternali, F. Collu, H. Rezaei, B. Strodel, and A. Raudino, "Symmetry-breaking transitions in the early steps of protein self-assembly," *Eur. Biophys. J.*, vol. 49, no. 2, pp. 175–191, Mar. 2020,

[231] A. D. Stephens, J. Kölbel, R. Moons, C. W. Chung, M. T. Ruggiero, N. Mahmoudi, T. A. Shmool, T. M. McCoy, D. Nietlispach, A. F. Routh, F. Sobott, J. A. Zeitler, and G. S. Kaminski Schierle, "Decreased Water Mobility Contributes To Increased α-Synuclein Aggregation**," *Angew. Chemie Int. Ed.*, vol. 62, no. 7, p. e202212063, Feb. 2023,

[232] M. Verheul, S. P. F. M. Roefs, and K. G. De Kruif, "Aggregation of β-lactoglobulin and

influence of D2O," *FEBS Lett.*, vol. 421, no. 3, pp. 273–276, Jan. 1998,

[233] C. L. Klaips, G. G. Jayaraj, and F. U. Hartl, "Pathways of cellular proteostasis in aging and disease," *J. Cell Biol.*, vol. 217, no. 1, pp. 51–63, Jan. 2018,

[234] K. P. Gregory, G. R. Elliott, H. Robertson, A. Kumar, E. J. Wanless, G. B. Webber, V. S. J. Craig, G. G. Andersson, and A. J. Page, "Understanding specific ion effects and the Hofmeister series," *Phys. Chem. Chem. Phys.*, vol. 24, no. 21, pp. 12682–12718, Jun. 2022,

[235] P. Dogra, S. S. Roy, A. Joshi, and S. Mukhopadhyay, "Hofmeister Ions Modulate the Autocatalytic Amyloidogenesis of an Intrinsically Disordered Functional Amyloid Domain via Unusual Biphasic Kinetics," *J. Mol. Biol.*, vol. 432, no. 23, pp. 6173–6186, Nov. 2020,

[236] D. N. Dean and J. C. Lee, "Purification and characterization of an amyloidogenic repeat domain from the functional amyloid Pmel17," *Protein Expr. Purif.*, vol. 187, p. 105944, Nov. 2021,

[237] A. K. Buell, C. Galvagnion, R. Gaspar, E. Sparr, M. Vendruscolo, T. P. J. Knowles, S. Linse, and C. M. Dobson, "Solution conditions determine the relative importance of nucleation and growth processes in α-synuclein aggregation," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 111, no. 21, pp. 7671–7676, May 2014,

[238] G. Meisl, X. Yang, C. M. Dobson, S. Linse, and T. P. J. Knowles, "Modulation of electrostatic interactions to reveal a reaction network unifying the aggregation behaviour of the Aβ42 peptide and its variants," *Chem. Sci.*, vol. 8, no. 6, pp. 4352–4362, May 2017,

[239] M. F. M. Sciacca, D. Milardi, G. M. L. Messina, G. Marletta, J. R. Brender, A. Ramamoorthy, and C. La Rosa, "Cations as Switches of Amyloid-Mediated Membrane Disruption Mechanisms: Calcium and IAPP," *Biophys. J.*, vol. 104, no. 1, p. 173, Jan. 2013,

[240] H. F. Zein, I. Alam, P. Asanithi, and T. Sutthibutpong, "Molecular dynamics study on the effects of charged amino acid distribution under low pH condition to the unfolding of hen egg white lysozyme and formation of beta strands," *PLoS One*, vol. 17, no. 3, p. e0249742, Mar. 2022,

[241] K. K. Skeby, O. J. Andersen, T. V. Pogorelov, E. Tajkhorshid, and B. Schiøtt, "Conformational Dynamics of the Human Islet Amyloid Polypeptide in a Membrane Environment: Toward the Aggregation Prone Form," *Biochemistry*, vol. 55, no. 13, pp. 2031–2042, Apr. 2016,

[242] Z. Qian, Y. Zou, Q. Zhang, P. Chen, B. Ma, G. Wei, and R. Nussinov, "Atomistic-level study of the interactions between hIAPP protofibrils and membranes: Influence of pH and lipid composition," *Biochim. Biophys. acta. Biomembr.*, vol. 1860, no. 9, p. 1818, Sep. 2018,

[243] H. K. Anandatheerthavarada, G. Biswas, M. A. Robin, and N. G. Avadhani, "Mitochondrial targeting and a novel transmembrane arrest of Alzheimer's amyloid precursor protein impairs mitochondrial function in neuronal cells," *J. Cell Biol.*, vol. 161, no. 1, pp. 41–54, Apr. 2003,

[244] T. A. Strope and H. M. Wilkins, "Amyloid precursor protein and mitochondria," *Curr. Opin.*

Neurobiol., vol. 78, p. 102651, Feb. 2023,

[245] M. S. Uddin, M. T. Kabir, M. S. Rahman, T. Behl, P. Jeandet, G. M. Ashraf, A. Najda, M. N. Bin-Jumah, H. R. El-Seedi, and M. M. Abdel-Daim, "Revisiting the Amyloid Cascade Hypothesis: From Anti-Aβ Therapeutics to Auspicious New Ways for Alzheimer's Disease," *Int. J. Mol. Sci.*, vol. 21, no. 16, pp. 1–33, Aug. 2020,

[246] J. M. Sanderson, "The association of lipids with amyloid fibrils," *J. Biol. Chem.*, vol. 298, no. 8, Aug. 2022,

[247] J. Mu, H. Liu, J. Zhang, R. Luo, and H. F. Chen, "Recent Force Field Strategies for Intrinsically Disordered Proteins," *J. Chem. Inf. Model.*, vol. 61, no. 3, p. 1037, Mar. 2021,

[248] J. Chen, X. Liu, and J. Chen, "Targeting Intrinsically Disordered Proteins through Dynamic Interactions," *Biomolecules*, vol. 10, no. 5, May 2020,

[249] J. Preto and I. Krimm, "The intrinsically disordered N-terminus of the voltage-dependent anion channel," *PLoS Comput. Biol.*, vol. 17, no. 2, Feb. 2021,

[250] K. Gade Malmos, L. M. Blancas-Mejia, B. Weber, J. Buchner, M. Ramirez-Alvarado, H. Naiki, and D. Otzen, "ThT 101: a primer on the use of thioflavin T to investigate amyloid formation," *Amyloid*, vol. 24, no. 1, pp. 1–16, Jan. 2017,

Appendix

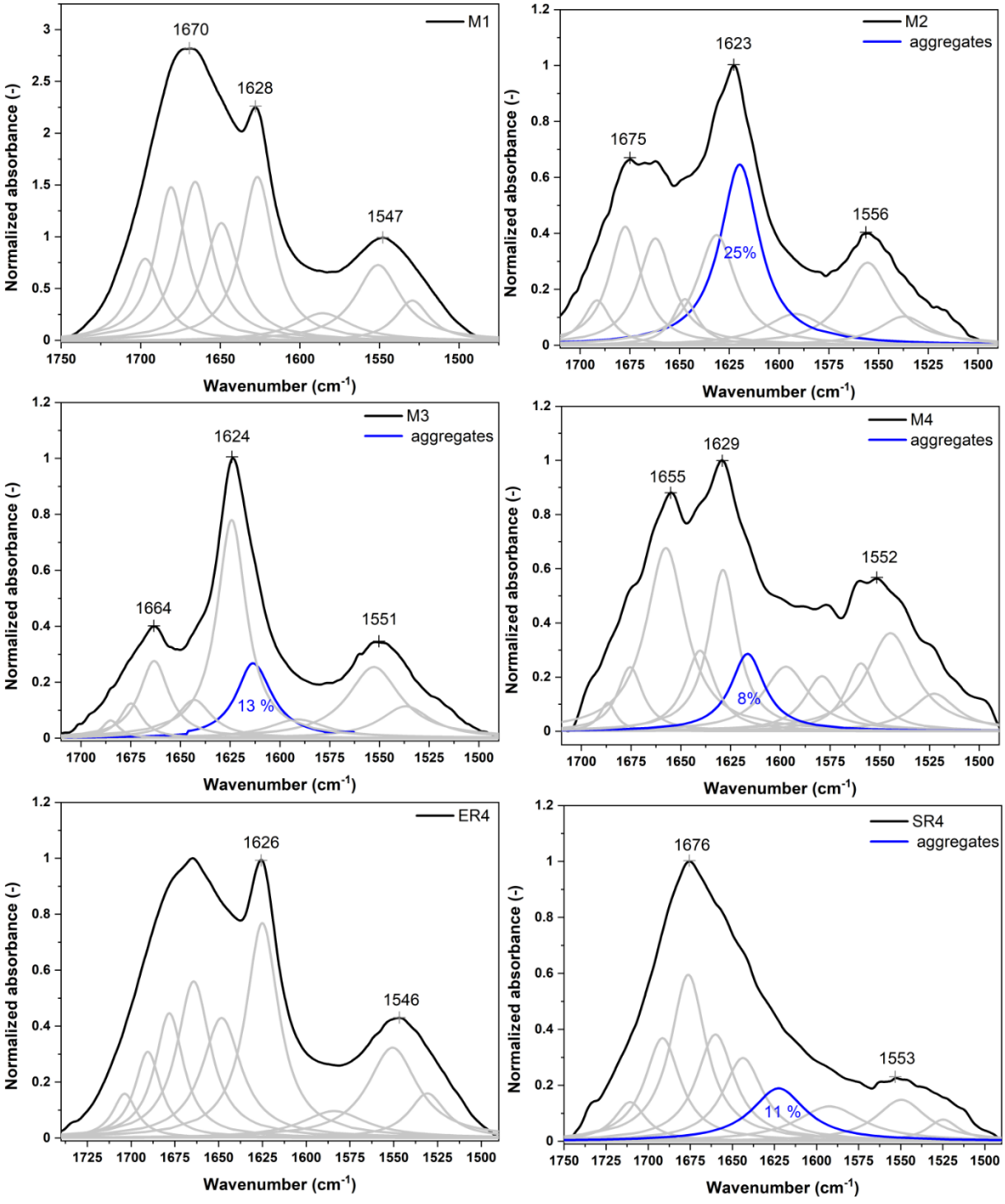
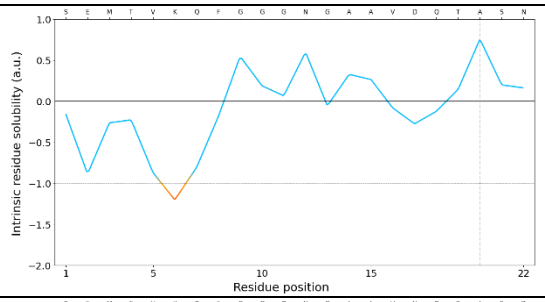
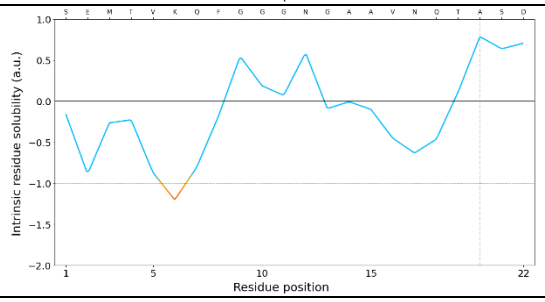
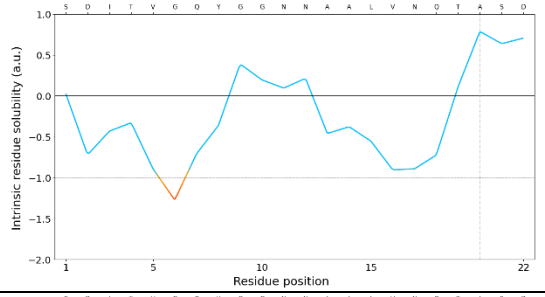
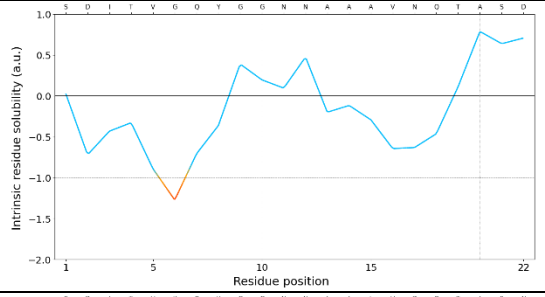
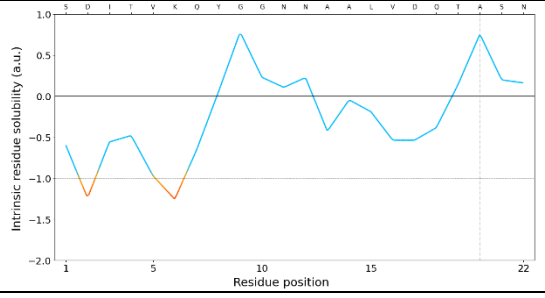
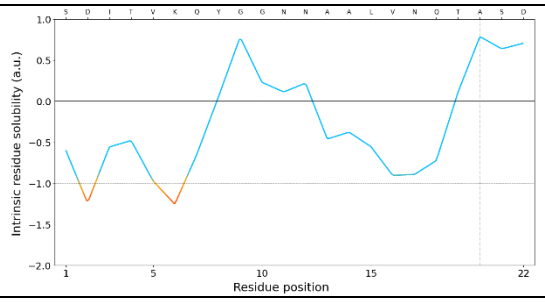


Figure 60 Normalized ATR-FTIR spectra of air-dried films of mutants and R4 fragments with sub-bands obtained from the curve fitting procedure in the amide bands region (1750–1490 cm^{-1}).

Table 19 The results obtained from CamSol for mutants and R4 fragments of CsgA. In red marked introduced mutations.

Name	Sequence	Variant intrinsic solubility score	Solubility profile
ER4	SEMTVKQFGGG-NGAAVDQTASN	1.64	
M1	SEMTVKQFGGG-NGAAVNQTASD	1.64	
SR4	SDITVGQYGG-NNAALVNQTASD	1.57	
M2	SDITVGQYGG-NNAAAVNQTASD	1.65	
M3	SDITVKQYGG-NNAALVDQTASN	1.60	

M4	SDITVKQYGG- NNAALVNQTASD	1.50	

The solubility plot shows profile, in which regions with scores larger than 1 denote highly soluble regions, while scores smaller than -1 poorly soluble ones (one score per residue in the protein sequence).

1. ThT fluorescence assay.

The aggregation kinetic curves obtained for studied peptides represent classic example of a ThT-monitored kinetic curves and so were fitted by the sigmoidal curves using the following equation (1)[250]:

$$Y = y_i + m_i t + \frac{y_f + m_f t}{1 + e^{-[(t - t_{1/2})/\tau]}} \quad (1)$$

where, Y is the fluorescence intensity as a function of time t, y_i and y_f are the intercepts of the initial and final baselines with the y-axis, mi and mf are slopes of the initial and final baselines, t½ is the time needed to reach halfway through the elongation phase and τ is the elongation time constant. The apparent rate constant, k_{app}, for the growth of fibrils is given by 1/τ, and the lag time is defined as t_{lag} = t½ –2τ.

ThT fluorescence assay was used to assess the kinetics of the aggregation process. We observed a typical sigmoidal nucleation-polymerization curve for M2, M4, M1 peptides. The fluorescence curve for peptides SR4, ER4 and M3 cannot be fitted by equation 1. However, after the lag phase, the curve has a more complex shape (SR4, ER4). In the case of peptide M3 the increased fluorescence before the elongation process can be observed. This can mean that some phase transition can occur, that transition may correspond to the change of self-assembly pathway of the studied peptide.

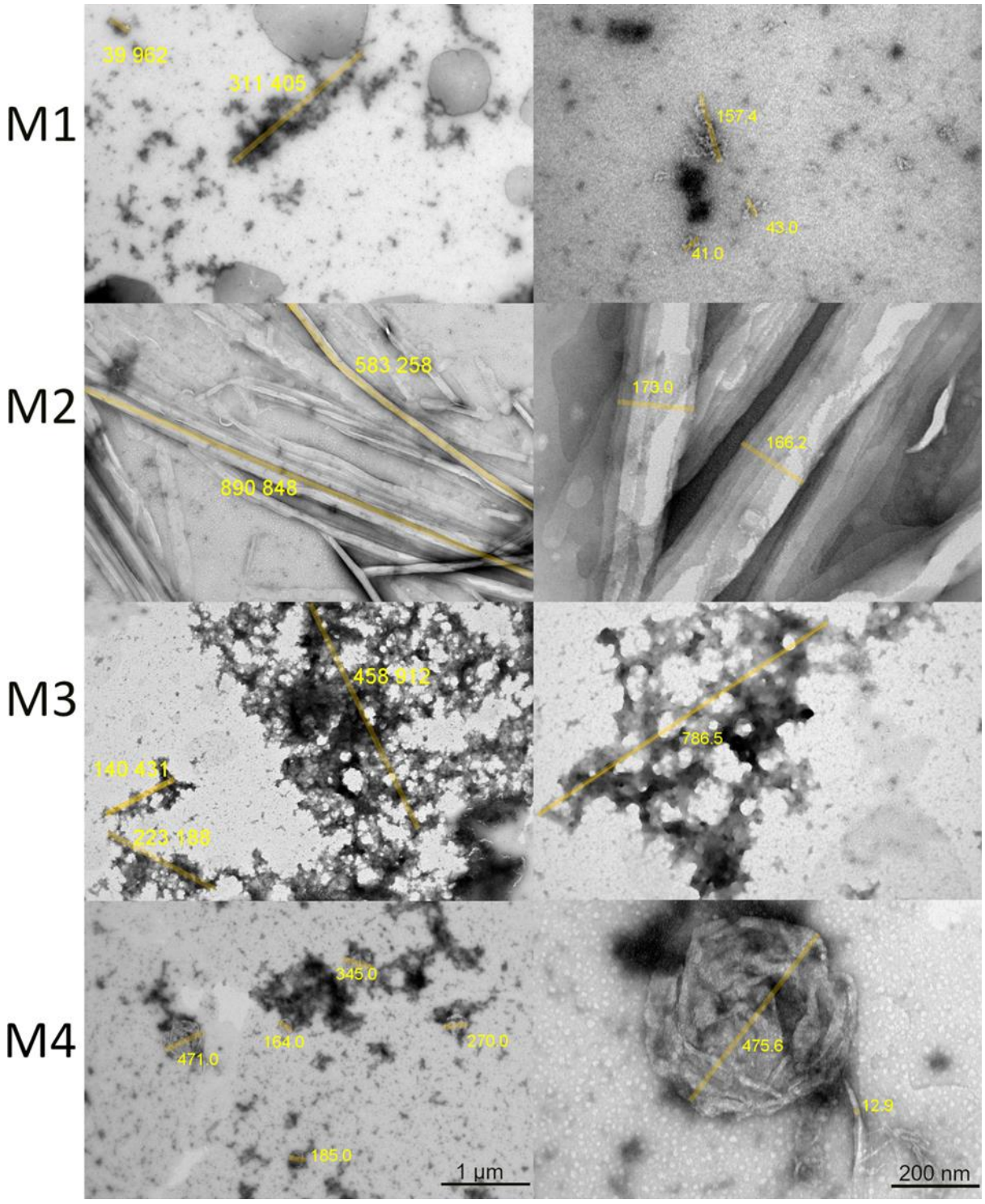
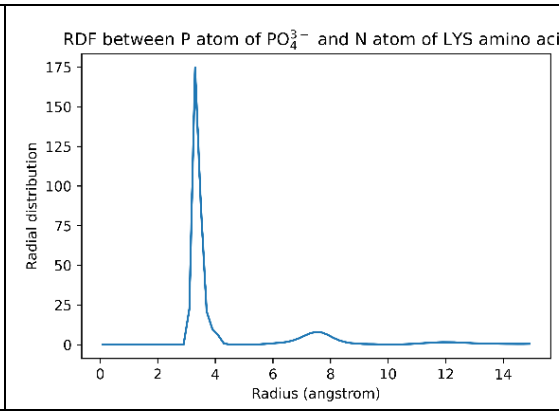
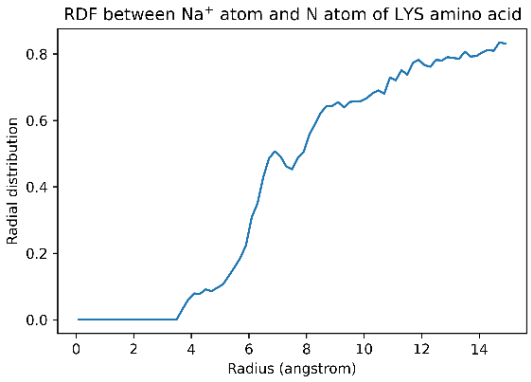
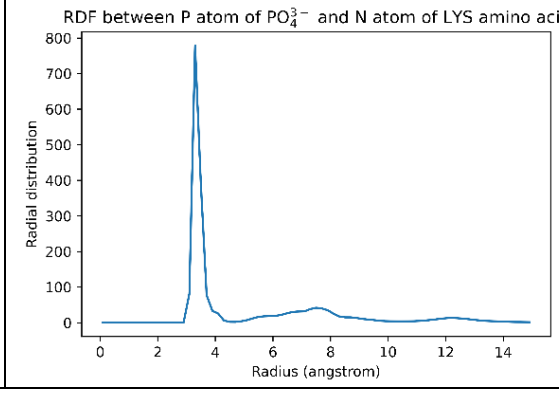
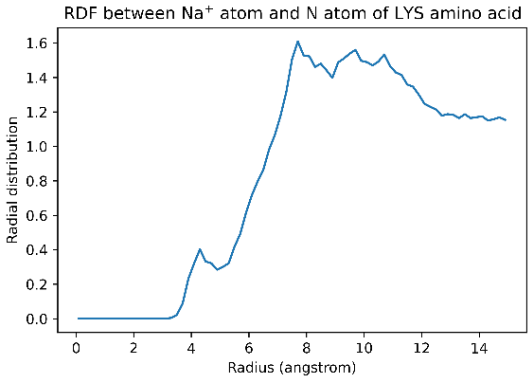


Figure 61 Electron micrographs with the marked size (in nm) of mutants on the day of the dissolving (left column: magnification of 10 000, right column: magnification of 40 000). Cpep = 500 µM.

Table 20 The RDF plots of M3 and M4 mutants.

M3		
M4		

List of Figures

List of Tables