Uniwersytet Gdański
[University of Gdańsk]

Związek Uczelni Fahrenheita
[Fahrenheit Universities]

Faculty of Chemistry
ul. Wita Stwosza 63
80-308 Gdańsk

Gdańsk, 10.06.2022

A review of the doctoral dissertation of Mr. Nguyen Truong Co, M.Sc., entitled "Study of factors governing mechanism of protein aggregation by using computer simulation"

The aggregation of peptides and proteins is of great importance in the functioning of living organisms, playing both a functional role (e.g. in blood clotting) and a destructive role, causing conformational diseases. Peptide aggregation is also used in the creation of bionanomaterials that can be applied, for example, as coatings for drug molecules or as part of dressings for wounds or burns. Despite many years of intensive research, the structures of aggregates and the mechanisms of their formation have not been fully understood. Studying the aggregation of peptides and proteins is all the more difficult as only the structures of fully formed aggregates, called amyloids, can be determined experimentally, while oligomers and aggregates of relatively low molecular weight that are the most toxic to cells and peptide aggregates used in dressings have a poorly defined structure. Moreover, the long time of aggregation is a great inconvenience, both in theoretical and experimental studies.

The doctoral dissertation of Mr. Nguyen Truong, M.Sc., concerns theoretical research on the influence of various factors on aggregation and the mechanisms of thermal decomposition of peptide aggregates. The doctoral thesis was carried out under the supervision of prof. dr hab. Maisuan Li *[professor, Ph.D., postdoctoral degree holder]*, who is one of the world's most recognized scientists in theoretical research on aggregation. The dissertation was based on 3 publications, one of which was published in Biomolecules and two in the Journal of Chemical Physics. The Candidate is the first author of two of them and the accompanying statements of the co-authors prove his dominant or very important contribution to their creation. All the journals, where the dissertation material was published, have an excellent reputation in the natural sciences. In addition to the publications constituting the basis for the dissertation, the candidate has 5 other thematically related publications and 1 chapter in the book. He is also the author of one conference presentation.

The text of the dissertation presented for review consists of an introduction (chapter 1), a literature review (chapter 2), a description of the calculation methods (chapter 3), and the attached original three mentioned publications, each of which is preceded by a synthetic

overview, an abstract and statements by the co-authors and the Candidate (chapter 4), and conclusions, with an indication of future research (chapter 5). This main part of the dissertation is preceded by the Candidate's declaration of the work's authorship, a synthetic abstract, a list of publications constituting the basis for the dissertation, a table of contents, and a list of figures and tables. The literature list follows the main text of the dissertation. The entire dissertation was written in English.

The first chapter comprises two parts, the first of which is entitled "Motivation" and contains the objective of the work and a brief description of the research carried out, and the second is entitled "Other important information", containing the table of contents of the main part of the dissertation, a list of publications constituting the basis for the dissertation, a list of other publications and conference presentations as well as information on the number of citations and the Hirsch index of the Candidate. Information about the main objective of the work can be found in the first paragraph of the first part of this chapter. It shows that it was the study, using simulation methods, of factors affecting the aggregation of peptides. In the following paragraphs, specific objectives appear, allowing the reader to see that the purpose was to investigate how interactions with the surface and the content of secondary structures in the monomer affect the aggregation rate and to investigate the mechanisms of aggregate dissociation. The description of the specific objectives of the dissertation, however, is mixed with the reference to the obtained results. I believe that the aims of the dissertation should be more clearly distinguished. On the other hand, the second part of the chapter contains the information already provided in the preamble (part of the table of contents and the list of publications constituting the basis for the dissertation).

The second chapter consists of three parts. In the first part, the Author discusses the kinetics of the aggregation process in the light of the energy landscape theory, then the structure of amyloid fibrils and conformational diseases caused by the aggregation of peptides and proteins. Figure 2 is a very good illustration of the phases of the aggregation process. This part of the chapter is written in very clear language and contains relevant information, but it should not be entitled "Protein self assembly kinetics" as it covers much broader issues. Somewhat confusing is the sentence on page 5 starting in line 3 of part 2.1, *"A protein functions correctly if it can adopt a specific compact and energetically favorable three-dimensional structure called a folded structure or native state"* – this statement does not apply to intrinsically disordered proteins, about which the author writes in further parts of the dissertation, and for which it is precisely plasticity that enables them to fulfill various functions. The second part of this chapter discusses the factors affecting aggregation. This part is brilliantly written and provides an excellent synthesis of the subject. The third part of the chapter discusses the models used in theoretical studies of peptide and protein aggregation, which the author divides into all-atom, coarse-grained and lattice models of aggregation. I believe that this part should be placed in Chapter 3, which is devoted to calculation methods.

In my opinion, the division of models given by the author in section 2.3 is not the best. A lattice or continuous space is a choice of space representation. On the other hand, representing a molecule in the form of atoms (all-atom models) or interaction centers involving many atoms (coarse-grained models) is associated with omitting the appropriate part of the degrees of freedom – electrons in the all-atom representation or other than collective degrees of freedom of entire groups of atoms in the coarse-grained representation. In each of these cases, these neglected degrees of freedom must be accounted for as a function of the effective potential energy. Both all-atom and coarse-grained models can be represented in a continuous or discrete (lattice) space, although discrete space representations are much more commonly used for coarse-grained models. Therefore, it would be more skillful to talk about all-atom models (continuous space) and coarse-grained models in a continuous space or on a lattice; moreover, the Candidate

adopts such a division later in the dissertation. When discussing coarse-grained models in the lattice representation, the Author should mention the CABS model of the group of professor Kolinski, which, contrary to the lattice models discussed in the dissertation, enables simulations of real proteins, as evidenced by the excellent results obtained by this group in the CASP experiments for protein structure prediction.

In the third chapter, the candidate discusses computational methods. The first part of the chapter discusses the Metropolis Monte Carlo method, in the second part the coarse-grained model of polypeptide chains developed by Li-Klimov-Straub-Thirumalai, which was the basis for most of the Candidate's calculations. In this model, the polypeptide chain is reduced to a trace of α-carbon atoms, arranged on a cubic lattice, with simple contact potentials. Lattice models of surface are also discussed in this section. In Part 3, the author briefly characterizes Langevin dynamics. This is where you should find information about the all-atom force fields used in the dissertation. The final two parts concern properly controlled molecular dynamics and methods of analyzing the results of coarse-grained and all-atom simulations.

Chapter four contains the results obtained in the dissertation. Its integral part includes 3 original publications of the Candidate, which are the basis for the dissertation. This material has already been positively assessed in terms of its content by the reviewers of the publication. The obtained results are summarized and discussed in chapter five. In the research, the Candidate conducted a coarse-grained Monte Carlo simulation using the Li-Klimov-Straub-Thirumalai model (all 3 papers) and all-atom molecular dynamics simulations using OPLS (paper 2) and AMBER (paper 3).

Below I present a brief summary of the most important achievements of the Candidate, in my opinion, obtained during the reviewed doctoral dissertation.

1. Introduction of a lattice representation of surfaces to the Li-Klimov-Straub-Thirumalai model, which enabled its extension to the study of aggregation in the presence of surfaces with varying degrees of roughness and strength of interactions with peptides. The Candidate showed that moderately strong interactions of peptides with the surface accelerate aggregation, as a result of a partial limitation of the conformational freedom of the chains, while weak or strong interaction with the surface reduces the aggregation rate. The obtained results are consistent with the results of independent experiments.

2. Demonstrating, by simulating all-atom monomers of the amyloidogenic Aβ42 peptide and its 19 mutants, that the aggregation rate is directly correlated with the content of the secondary structure in the monomer and introducing a correlation relationship between the change of aggregation rate and the content of β structure in the monomer, hydrophobicity and charge change due to mutation.

3. Development of a phenomenological-analytical theory describing the mechanism of dissociation of monomers from fibrils under the influence of temperature, which covers both the case where the dissociated monomers remain in the environment, with the possibility of reattachment to the fibrils and the case of removing monomers from the environment. The derived kinetic equation takes the form of a logistic equation, the solution of which changes into an exponential dependence after reaching the critical concentration of fibrils. This theory was confirmed by the simulations using the Li-Klimov-Straub-Thirumalai model and all-atom simulations. The obtained results are consistent with the available experimental data.

While reading the dissertation, the following three substantive comments came to my mind:

1. In paper 2, which is part of the dissertation (Thu, Co, Tu, Li, *J. Chem. Phys.* **150**, 225101 (2019), I did not find information about the structures from which the all-atom MD calculations were started with the replacement of replicas of the Aβ42 peptide and its mutants. Was it a wild-type experimental structure (PDB: 6SZF)? Although the replica exchange method is relatively insensitive to the initial configuration of the system, information about the start should be provided.

2. Why wasn't the Weighted Histogram Analysis Method (WHAM) used in paper 2 to determine the mean after the replica exchange simulation panels, and only the data from replicas for selected temperatures were taken? Using WHAM in a "binless" version would improve the statistics.

3. What could be the physical point of introducing the input from $\mathbf{\Theta}^2$ in equation 3 in paper 3? This input changes the shape of the time dependence $\mathbf{\Theta}$ to that corresponding to the logistic curve, but there are probably many modifications that give a similar shape to the solution of the kinetic equation.

As the reviewer, I am obliged to pay attention to minor errors and inaccuracies that I noticed in the dissertation:

Page 10, line 12, section "Temperature": *...proteins complicatedly influent....* The word *influent* is used as a verb and it cannot be a verb; it should be *influence* or *affect*.

Page 14, line 1, section "Charge": *...peptides the solubility...*, should be *...the peptide solubility....*

Page 16, section 2.3.1.: After Equation 3, instead of *the detail information*, it should be *the detailed information*.

Page 21: Beginning of the section: *Li at al. were built...* should be *Li et al. have built...* Instead of *The energy of N peptides...* it should be *The energy of a N-peptide-chain system...*, because otherwise, you might as well understand the sentence as referring to the energies of N different (isolated) peptides.

Page 27: There should be a reference to Figures 11A and B, not 10A and B. In Equation 10, the expression under the summation sign should be squared (it is correct in the original publication). The level of *h* is not defined.

Page 49, line 3: Instead of *Contradict to the case...* there should be *Contrary to the case...* or *In contrast to the case...*

Page 62, first paragraph, line 2: Instead of *...much toxic than mature fibrils...* should be *...more toxic than mature fibrils...*

I very highly rate the doctoral dissertation of Mr. Nguyen Truong Co, M.Sc. The Candidate enriched the coarse-grained model developed by Li-Klimov-Straub-Thirumalai with interaction with the surface, which significantly extends the possibilities of aggregation research. He developed an original

analytical phenomenological model of the dissociation of monomers from fibrils, introducing the original logistic equation to the description of this phenomenon. He derived the relationship between the change in the aggregation rate of the amyloidogenic Aβ42 peptide under the influence of mutation and the change in the content of the β structure, hydrophobicity, and charges of individual residues. He conducted the research reliably and diligently. The above-mentioned critical remarks, some of which are questionable, do not significantly diminish its value. The dissertation definitely meets the requirements set for doctoral dissertations by the Act of 14 March 2003 on academic degrees and academic titles as well as academic degrees and titles in the field of art (Journal of Laws No. 65, item 595, as amended), as well as customary standards for doctoral dissertations in the field of natural and exact sciences. All the scientific objectives of the dissertation were achieved. Therefore, with full conviction, I ask the High Scientific Council of the Institute of Physics of the Polish Academy of Sciences to admit Mr. Nguyen Truong Co, M.Sc., to the further stages of the doctoral dissertation. Due to the outstanding nature and high scientific value of the dissertation, I would like to motion for its distinction. Attached is the Justification for this motion.

[illegible signature]
prof. dr hab. Józef Adam Liwo *[professor, Ph.D., postdoctoral degree holder]*

Uniwersytet Gdański
[University of Gdańsk]

Związek Uczelni Fahrenheita
[Fahrenheit Universities]

Faculty of Chemistry
ul. Wita Stwosza 63
80-308 Gdańsk

Gdańsk, 10.06.2022

The justification for the motion for the distinction of the doctoral dissertation of Mr. Nguyen Truong Co, M.Sc., entitled "Study of factors governing mechanism of protein aggregation by using computer simulation"

The doctoral dissertation of Mr. Nguyen Truong Co, M.Sc., presented to me for review, concerns research on the mechanisms of amyloidogenic peptide aggregation. This issue is extremely important in biology and medicine. The Candidate conducted extensive simulation studies, both using the minimal coarse-grained model developed by Li-Klimov-Straub-Thirumalai, which he extended to the interaction of peptides with surfaces, and all-atom models using OPLS and AMBER force fields. The results he obtained significantly contribute to the understanding of the mechanisms and conditions of peptide association. In particular, the Candidate proved that the aggregation rate increases on surfaces with moderately strong interaction with peptides, and he proposed a relationship between the change in the aggregation rate due to mutation of the amyloidogenic Aβ42 peptide and the content of the β structure in the isolated peptide as well as the change in hydrophobicity and charges of amino acid residues, and showed that the kinetics the dissociation of monomers from aggregates is described by a logistic equation. The research was carried out and the results were prepared in a highly professional manner. The material of the dissertation was published in the form of three publications in journals from the ISI list, and in two of them Mr. Nguyen Truong Co, M.Sc., is the first author. I believe that the dissertation stands out above the level of typical doctoral dissertations on the application of computational methods in biophysics. Moreover, the Candidate has 5 other thematically related papers and 1 chapter in the book. Therefore, I would like to motion for the distinction of this doctoral dissertation.

[illegible signature]
prof. dr hab. Józef Adam Liwo *[professor, Ph.D., postdoctoral degree holder]*