Attachment 7

Summary of Professional Accomplishments

Paweł Krupa

Warsaw, 10.09.2024

Table of content

1. Name

Paweł Krupa

2. Diplomas, degrees conferred in specific areas of science or arts

- The degree of Doctor of Chemical Sciences with honors was awarded by the Faculty of Chemistry of the University of Gdańsk on July 8, 2015, based on the doctoral dissertation titled: Extension of the UNRES force field with local potentials and data from comparative methods for better prediction of protein structures. Supervisor: Prof. dr hab. Cezary Czaplewski.
- Master of Chemical Sciences, specialization in cheminformatics, awarded by the Faculty of Chemistry of the University of Gdańsk in 2011. Supervisor: Dr. Magdalena Ślusarz.
- Bachelor of Chemical Sciences, specialization in medicinal chemistry, awarded by the Faculty of Chemistry of the University of Gdańsk in 2009. Supervisor: Dr. Magdalena Ślusarz.

3. Information on employment in research institutes or faculties/departments

- September 1, 2016 - present: Assistant Professor in the Department of Theoretical Physics, Theoretical Biophysics Group, Institute of Physics of the Polish Academy of Sciences, Warsaw.
- August 1, 2015 - July 31, 2016: Postdoctoral associate in the group of Professor Harold Scheraga at the Baker Laboratory of Chemistry and Chemical Biology, Cornell University, Ithaca, NY, USA.
- November 17, 2014 - May 1, 2015: Research Aide in the group of Professor Andrzej Kłoczowski at the Battelle Center for Mathematical Medicine, The Research Institute at Nationwide Children's Hospital and Ohio State University in Columbus, OH, USA.
- October 8, 2014 - July 31, 2015: Specialist analyst at the IT Center of the Tricity Academic Computer Network (CI TASK), Gdańsk University of Technology, Gdańsk.

4. Bibliometric summary of the published works

Current bibliographic data can be verified using the following identifiers of the habilitation candidate:

- Open Researcher and Contributor ID (ORCID): 0000-0002-9710-7837
- ResearcherID: S-2292-2016

Selected bibliometric data:

- **Web of Science Core Collection:**
    - 45 items (counting only full scientific articles)
    - 833 citations (640 without self-citations)
    - 448 citing articles
    - H-index = 18

    - i10-index = 26.
- **Google Scholar:**
    - 95 items (counting all items)
    - 1155 citations
    - H-index = 21 (since 2019: 19)
    - i10-index = 28 (since 2019: 24).

For clarity, I have summarized the current values of impact factors and ministerial points in Table 1.

Table 1. Summary of articles in journals and book chapters with assigned impact factor values (data from Web of Science, May 2024) and/or the number of ministerial points (list from January 5, 2024).

Journal	Impact Factor	Impact Factor 5-year	Ministerial points	Number of publications
Proceedings of the National Academy of Sciences	11.1	12	200	2
Nanoscale	6.7	6.8	140	1
Bioinformatics	5.8	8.3	200	2
Journal of Chemical Information and Modeling	5.6	5.9	100	4
Journal of Molecular Biology	5.6	5.5	140	1
International Journal of Molecular Sciences	5.6	6.2	140	2
Journal of Chemical Theory and Computation	5.5	5.8	140	3
Chemico-Biological Interactions	5.1	4.9	100	1
Scientific Reports	4.6	4.9	140	1
Molecules	4.6	4.9	140	1
Journal of Biomolecular Structure and Dynamics	4.4	3.8	70	1
The Journal of Chemical Physics	4.4	3.5	100	1
Progress In Molecular Biology And Translational Science	4	5.3	100	1
Genes	3.5	3.9	100	1
BBA Biomembranes	3.4	3.5	100	1
The Journal of Physical Chemistry B	3.3	3	140	10
Journal of Computational Chemistry	3	3.3	100	1
Proteins: Structure, Function, and Bioinformatics	2.9	2.9	100	4
Journal of Molecular Graphics and Modelling	2.9	2.4	70	4
Journal of Molecular Modeling	2.2	1.8	40	2
Methods in Molecular Biology	0	0	70	3
Supercomputing Frontiers and Innovations	0	0	70	1
TASK Quarterly	0	0	20	3
Total	**193.8**	**199.1**	**5630**	**51**

5. Description of the achievements

5.1 Title of the scientific achievement

Development and application of force fields at different levels of resolution to study the structure and the dynamics of selected proteins

5.2 List of scientific articles included in the series of publications

H1. **Paweł Krupa**, Magdalena A. Mozolewska, Marta Wiśniewska, Yanping Yin, Yi He, Adam K. Sieradzan, Robert Ganzynkowicz, Agnieszka G. Lipska, Agnieszka Karczyńska, Magdalena Ślusarz, Rafał Ślusarz, Artur Giełdoń, Cezary Czaplewski, D. Jagieła, B. Zaborowski, Harold A. Scheraga*, Adam Liwo, "Performance of protein-structure predictions with the physics-based UNRES force field in CASP11", Bioinformatics 2016, 32(21), 3270-3278.

H2. **Paweł Krupa**, Anna Hałabis, Wioletta Żmudzińska, Stanisław Ołdziej, H.A. Scheraga, Adam Liwo,* "Maximum Likelihood Calibration of the UNRES Force Field for Simulation of Protein Structure and Dynamics", Journal of Chemical Information and Modeling 2017, 57 (9), 2364-2377.

H3. Nguyen Hoang Linh, **Pawel Krupa**, Minh Hai Nguyen, Huynh Quang Linh, Mai Suan Li*, "Structure and physicochemical properties of Aβ42 tetramer: Multi-scale molecular dynamics simulations", Journal of Physical Chemistry B 2019, 123, 7253−7269.

H4. **Pawel Krupa***, Pham Dinh Quoc Huy, and Mai Suan Li*, "Properties of Monomeric Aβ42 Probed by Different Sampling Methods and Force Fields: Role of Energy Components", The Journal of Chemical Physics 2019, 151(5), 055101.

H5. Daniela Marasco[+], Caterina Vicidomini[+], **Pawel Krupa**[+], Federica Cioffi, Pham Dinh Quoc Huy, Mai Suan Li, Daniele Florio, Kerensa Broersen, Maria Francesca De Pandis, Giovanni N Roviello*, "Plant isoquinoline alkaloids as potential neurodrugs: A comparative study of the effects of benzo [c] phenanthridine and berberine-based compounds on β-amyloid aggregation", Chemico-Biological Interactions 2021 334, 109300.

H6. Yuliia Varenyk, Panagiotis E. Theodorakis, Dinh Q. H. Pham, Mai Suan Li, and **Paweł Krupa***, "Exploring Structural Insights of Aβ42 and α-Synuclein Monomers and Heterodimer: A Comparative Study Using Implicit and Explicit Solvent Simulations", J. Phys. Chem. B 2024, 128 (19), 4655–4669.

H7. **Pawel Krupa**, Giovanni La Penna*, Mai Suan Li, "Amyloid-β Tetramers and Divalent Cations at the Membrane/Water Interface: Simple Models Support a Functional Role", International Journal of Molecular Sciences 2023, 24 (16), 12698.

H8. Natalia Karska, Igor Zhukov, Andrea D Lipińska, Sylwia Rodziewicz-Motowidło, **Paweł Krupa*,** "Why does the herpes simplex 1 virus-encoded UL49. 5 protein fail to inhibit the TAP-dependent antigen presentation?", Biochimica et Biophysica Acta (BBA)-Biomembranes 2023, 1865 (8), 184200.

H9. **Paweł Krupa***, "Treatment of disulfide bonds in coarse-grained UNRES force field", TASK Quarterly 2016, 20 (4), 393-398.

H10. **Paweł Krupa**, Adam K. Sieradzan, Magdalena A. Mozolewska, Huiyu Li, Adam Liwo, and Harold A. Scheraga*, "Dynamics of Disulfide-Bond Disruption and Formation in the Thermal Unfolding of Ribonuclease A", Journal of Chemical Theory and Computation 2017, 13 (11), 5721–5730.

H11. Pamela Smardz, Adam K. Sieradzan*, **Paweł Krupa***, "Mechanical Stability of Ribonuclease A Heavily Depends on the Redox Environment", The Journal of Physical Chemistry B 2022, 126, 33, 6240–6249.

H12. **Paweł Krupa**, Agnieszka G. Karczyńska, Magdalena A. Mozolewska, Adam Liwo, Cezary Czaplewski*, "UNRES-Dock - protein-protein and peptide-protein docking by coarse-grained replica-exchange MD simulations", Bioinformatics 2021 37 (11), 1613-1615.

H13. **Paweł Krupa***, Marta Spodzieja, Adam K. Sieradzan, "Prediction of CD28-CD86 protein complex structure using different level of resolution approach", Journal of Molecular Graphics and Modelling 2021 103, 107802

*corresponding author

[+]equal contribution

5.3 Discussion of the scientific objective of the works and the results achieved, along with a discussion of their possible application

5.3.1 Aim, motivation and basic assumptions in conducting research

The main goal of my scientific research was to investigate the structural, physicochemical and dynamic properties of one of the most important biomacromolecules from the point of view of life - proteins - using computational biophysics methods at different levels of resolution. Due to the non-trivial nature of the studied systems, including folded and intrinsically disordered proteins (IDPs), various oligomeric forms, and the presence of other components such as other proteins, metal ions or lipid membranes, my research was multi-stage. In the first step, I determined the capabilities of existing computational methods, then developed or improved them if necessary, and only then used them to investigate biologically relevant properties. The last stage was to compare the obtained results with experimental data and available literature data to determine their accuracy. My research is interdisciplinary in nature and combines computational biophysics, structural bioinformatics, molecular biology, theoretical chemistry and immunology, often based on experimental data obtained through scientific collaborations in order to achieve the highest possible reliability and obtain a complete picture of the studied phenomena.

The two basic tools used and developed during my research were: (i) the coarse-grained UNRES force field (D4), part of the UNICORN package (R3), developed in the group of Professor Adam Liwo from the Faculty of Chemistry, University of Gdańsk, and previously also by the group of Professor Harold Scheraga from Cornell University, and (ii) all-atom Amber force fields, developed as part of the Amber and AmberTools packages mainly under the leadership of Professor David Case from Rutgers University.[1] In some of the presented studies, both methods were used together to increase the reliability of the calculations performed and to obtain a complete picture of the observed physical phenomena at the resolution level of proteins and their amino-acid residues (coarse-grained representation) and individual atoms (all-atom representation).

My research was based on various basic assumptions, starting with the fact that the structure of globular proteins depends on their sequence (Anfinsen's dogma)[2], and the activity of most proteins, peptides, but also nucleic acids is strictly dependent on their structure and even small disturbances can have a significant impact on the activity and physicochemical properties[3]. Probably the most well-known example of how the substitution of one amino-acid residue in a large protein can cause drastic changes for the entire organism is sickle cell anemia, caused by the mutation of glutamic acid to valine (Glu6Val), which causes the formation of hemoglobin aggregates with a different morphology and thus reducing the efficiency of oxygen transport by

erythrocytes compared to the native protein[4]. On the other hand, however, there is a group of peptides and proteins whose fragments or entire molecules do not have a well-defined structure under physiological conditions. These are the so-called intrinsically disordered proteins (IDPs)[5]. This group includes, among others, amyloid β (Aβ) peptides and α-synuclein, the presence of which is closely related to neurodegenerative diseases. In turn, knowledge of the structure and dynamics of proteins is necessary to carry out most research in computational biophysics, for computer-aided design of therapeutic substances interacting with these proteins, or for studying and influencing the aggregation process. Due to the fact that subtle changes in peptides and proteins as well as their environment can significantly affect the studied systems and their properties, it is extremely important that the methods used and the planned experiments enable achieving sufficient accuracy.

Although there are computational methods that enable effective prediction of protein structure, in most cases based on statistical analysis of databases of protein sequences and structures, such as I-TASSER[6], Robetta[7], or AlphaFold3[8], the effectiveness of these methods is mainly limited to single-domain proteins with a structure similar to proteins with experimentally solved conformations. The accuracy of these methods is particularly low for protein complexes and partially unstructured proteins, and additionally they do not provide any information on the dynamics of the studied systems and the influence of external factors on their structure and properties. For this reason, it is necessary to further develop methods that use force fields based on the physics of interactions, rather than statistical data. Moreover, in the case of universal force fields, i.e. those that are not optimized for a specific system, the ability to predict the structure of a protein is an excellent way to approximately assess their efficiency in predicting their dynamics and physicochemical properties as well. It should be noted that these properties are often very difficult or even impossible to obtain experimentally. Therefore, the main goal of my research was to determine the predictive capabilities of available computational methods, select those that are characterized by the best reliability, extend them with necessary functions, and then use them to study selected protein systems that pose a challenge to existing methods, such as various oligomeric forms and aggregation of amyloid β, complexes of proteins involved in the body's immune response, and proteins containing disulfide bonds, which may play different functions in structured proteins and those without a stable structure.[9]

5.3.2 Summary of the conducted research

As part of the presented cycle of thematically related scientific articles, in the first step I focused on developing a unified protocol for predicting protein structure in the UNRES force field and using it to determine, based only on the sequence, the structures of all 55 proteins selected for structure prediction in the 11th edition of the CASP experiment (H1). The obtained results enabled me to make a reliable assessment of the method and plan the optimization of the weights of the potentials included in the energy function in the coarse-grained UNRES force field based on experimental data from nuclear magnetic resonance (NMR) for 7 selected training proteins at different temperatures, containing both folded and partially unfolded conformations (H2) in order to improve the predictive capabilities of the method.

Then, I used the newly optimized coarse-grained UNRES force field, which, thanks to the use of diverse sets of training structures, is characterized by good agreement in predicting both folded and unstructured conformations (H2), to predict the structure and dynamics of the tetrameric form of amyloid β (H3). In order to determine the reliability of the obtained oligomer structures, their stability was tested in two selected all-atom force fields and compared with available literature data. At the same time, I conducted research in five all-atom force fields from the Amber and CHARMM families, determining the efficiency of the selected methods for predicting the structures and dynamics of the inherently unstructured monomeric form of amyloid β along with determining the influence of individual components of the potential energy responsible for various physicochemical properties (H4).

In order to further understand the factors influencing the amyloid β aggregation process, I determined the mechanism of inhibition of this process under the influence of selected isoquinoline alkaloids based on all-atom molecular dynamics simulations (H5). The obtained results provide a theoretical explanation of experimental observations, presenting how the affinity of various isoquinoline alkaloids affects their activity in inhibiting the early stages of the aggregation process (H5). A related study that I conducted was to determine the molecular basis of the occurrence of dementia with Lewy bodies, DLB. For this purpose, I planned and performed studies on the aggregation mechanism of monomeric forms of amyloid β and α-synuclein and a comparison of their monomeric, homo- and heterodimeric forms (H6). The conducted research showed a significant increase in β-type structures in the heterodimer and proposed a stable model of this structure.

Then, I investigated the influence of the lipid membrane and copper ions on the aggregation, structure and dynamics of amyloid β dimers and tetramers, using various variants of all-atom simulation methods (H7). As part of this work, I also determined the affinity of mono- and bivalent metal ions with model homogeneous lipid bilayers using umbrella sampling and estimation of the potential of mean force (PMF) using the weighted histogram analysis method (WHAM). Next, I determined how the composition of a heterogeneous lipid membrane, similar in composition to biological membranes occurring under physiological conditions, affects the structure and dynamics of the herpesvirus protein UL49.5 from two viruses in relation to simple homogeneous models consisting only of 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine (POPC) chains (H8), commonly used in both experimental and theoretical studies. My theoretical studies also made it possible to explain reasons for the lack of inhibitory activity of one of them in the human body (H8).

In parallel with the studies presented above, I became interested in the role of disulfide bonds in proteins, and in particular their influence on the folding pathways and stability of selected proteins. It is worth noting that these bonds occur naturally in over 20% of proteins deposited in the PDB (Protein DataBank) database, but are also often added to both unstructured and folded peptides and proteins during experimental studies to obtain a specific effect. In order to determine the role of disulfide bonds in selected proteins, it turned out to be necessary to introduce and improve the potential enabling dynamic formation and breaking of disulfide bonds during simulations in the coarse-grained UNRES force field (H9 and H10). Then, an analogous potential was developed for the all-atom Amber force field (H11) and studies were performed using classical, steered and replica exchange molecular dynamics simulations in the coarse-grained UNRES force field and all-atom Amber to determine the mechanical and thermodynamic stability of disulfide bonds in ribonuclease A and their impact on the stability of this protein (H10 and H11).

During the above-mentioned research, which often concerned oligomeric forms of proteins, I developed a number of tools to facilitate the preparation and analysis of results, and also designed and implemented a function enabling routine prediction of the structure of protein-peptide and protein-protein complexes in the previously optimized (H2) coarse-grained UNRES force field (H12). I used the initial components of the method to generate the initial structures of the amyloid β and α-synuclein heterodimer (H6) and amyloid β tetramers (H7), and the complete protein-protein docking protocol was used to predict the structure of the CD28-CD86 protein complex, important in the body's immune response. I verified the stability of the obtained CD28-CD86 heterodimers using all-atom methods and available literature data, demonstrating the good efficiency of the method (H13).

In summary, the scientific research carried out as part of the cycle of thematically related publications can be divided according to overlapping research goals, which form a coherent whole leading to the development, understanding of the possibilities and limitations of methods and tools used in computational biophysics, and then using them to study the properties of selected proteins:

- Development and optimization of the coarse-grained UNRES force field (works H1, H2, H10, H12)
- Development and determination of predictive capabilities of all-atom force fields (works H3, H4, H11)
- Use of methods at different levels of resolution to obtain a complete picture of the studied issue (H3, H13)
- Determination of physicochemical properties, structures and dynamics of selected proteins (H1-H13)
    - Unstructured proteins (H3-H7)
    - Proteins with disulfide bonds (H9-H11)
    - Proteins in the lipid membrane environment (H7-H8)
    - Protein complexes (H3, H5-H7, H12-H13)

5.3.3 Conducted research and obtained results

H1. Development of a unified protocol for predicting protein structures in the UNRES force field based on the physics of interactions and evaluation of the predictive capabilities of the method

The first step to perform a reliable assessment of the predictive capabilities of the coarse-grained UNRES force field, previously extended with an additional potential describing the behavior of side chains based on the physics of interactions, which enabled the removal of higher-order correlation terms from the energy equation (D5), was to develop a unified protocol for predicting protein structures (H1) (Figure 1). This protocol is fully based on semi-automatic steps, not requiring manual user intervention at any stage, which is particularly important when selecting representative structures from replica exchange molecular dynamics simulations.[10] In the developed version, representative structures are obtained by performing cluster analysis of conformational ensembles obtained using the weighted histogram analysis method (WHAM)[11] below the obtained phase transition temperature, fully based on the analysis of the estimated free energy of conformations. The development of the protein structure prediction protocol (Figure 1), along with the inclusion of necessary scripts semi-automating the process, not only facilitated and accelerated the process, but also unified the results between different users and based the entire prediction process on the physics of interactions and the selection of representative structures based on the minimum free energy, instead of manual selection.

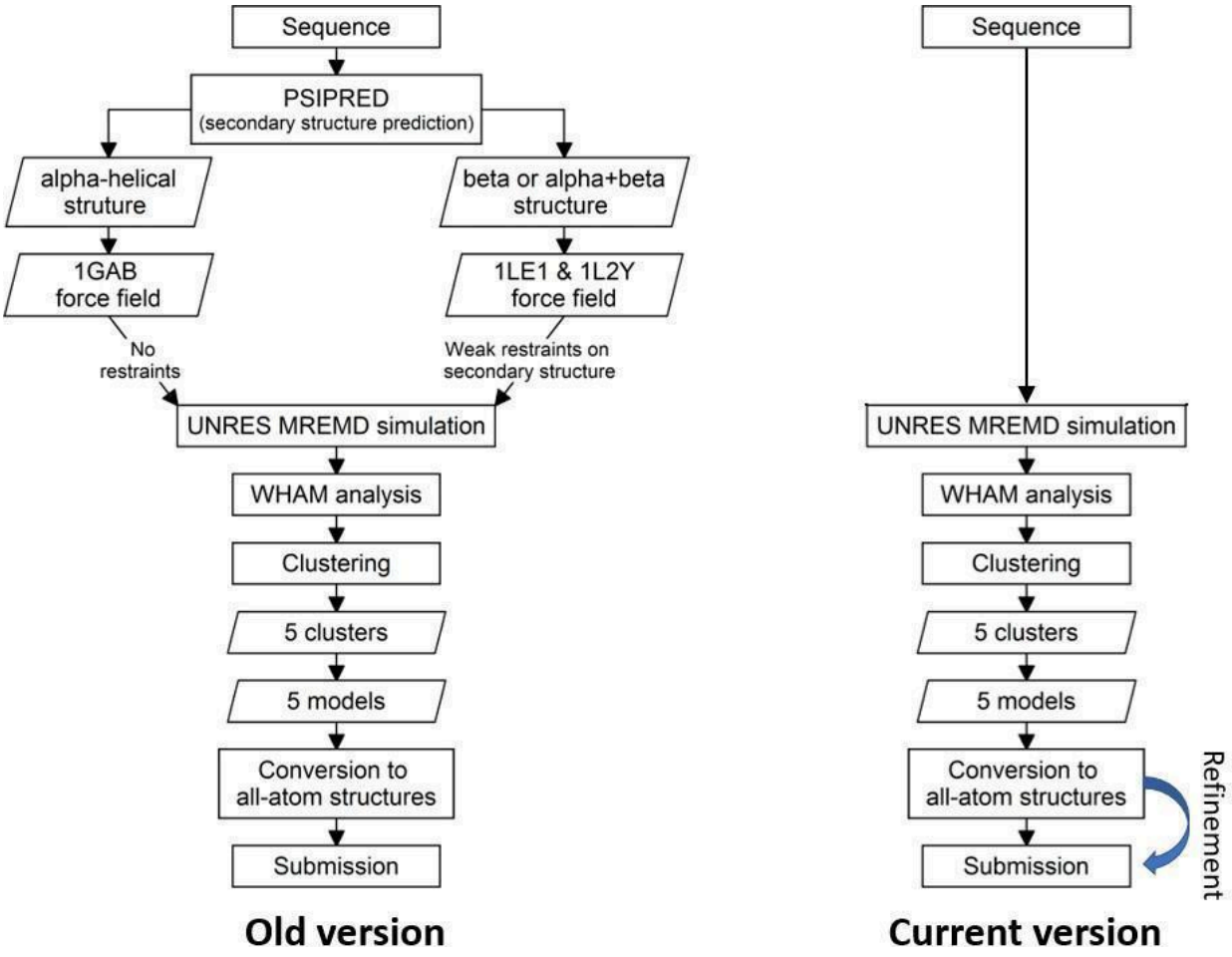


Figure 1. Schematic of the developed unified protein structure prediction protocol in the UNRES force field (work H1, left panel) and after force field optimization (work H2, right panel). The left panel of the figure is from work H1. Reproduction rights obtained (reprinted with permission from Bioinformatics, Volume 32, Issue 21, November 2016, Pages 3270-3278. Copyright 2024 Oxford University Press, license 5740260555795).

In order to achieve the most impartial and reliable assessment of the predictive capabilities of the UNRES force field using the developed protein structure prediction protocol, together with a group of scientists, I participated in the eleventh edition of the CASP (Critical Assessment of protein Structure Prediction) initiative[12]. It is an international experiment held every two years, in which at least several dozen research groups from around the world compete to predict the structure of selected proteins as accurately as possible based on the sequence alone (standard mode) and using additional experimental data (e.g. from the SAXS experiment). The most important advantage of this experiment is that under the same conditions, all research groups and computational methods participating in the experiment make blind predictions of protein structures that have not yet been published (most often they are in the process of being obtained by experimental methods), within a strictly defined time frame, usually two weeks for the regular structure prediction mode. The submitted protein structures are then analyzed by an external team of experts, and the results are officially announced after a few months on the CASP experiment website, presenting the accuracy of individual methods using objective measures and criteria. For this reason, in order to reliably and impartially test the developed protein structure prediction protocol, I performed simulations for most of the 55 different proteins, whose size ranged from 44 to 595 amino acid residues, with an average of 251 residues, provided in the form of sequences by the CASP experiment organizers. This was the first such extensive test of the capabilities of the UNRES force field, containing results for all proteins participating in this edition of the CASP experiment. Importantly, the set of

structures included both single-domain and multi-domain proteins with different types of secondary structures, allowing for a representative assessment of the quality of structure prediction for most protein classes.

The improvement of the energy function of the UNRES force field and the application of a unified protein structure prediction protocol eliminating the need for manual selection of representative structures, and using free energy as the selection criterion, resulted in a significant improvement in the obtained results compared to the previous, smaller-scale test in the 10th edition of the CASP experiment (D1). The ability to correctly rank the quality of structures among the five generated representative models was particularly improved thanks to the use of free energy as the selection criterion. Moreover, the performed research showed that the UNRES force field is able to correctly predict the structures of many proteins, competing with the best available methods, especially when the predicted structures were not similar to proteins with known structures. On the other hand, the performed tests showed that due to the simplifications resulting from the coarse-grained representation, the resolution of the UNRES force field is limited, for example, achieving a root mean square deviation (RMSD) value of about 3.8 Å for a 97-residue protein (Figure 2).

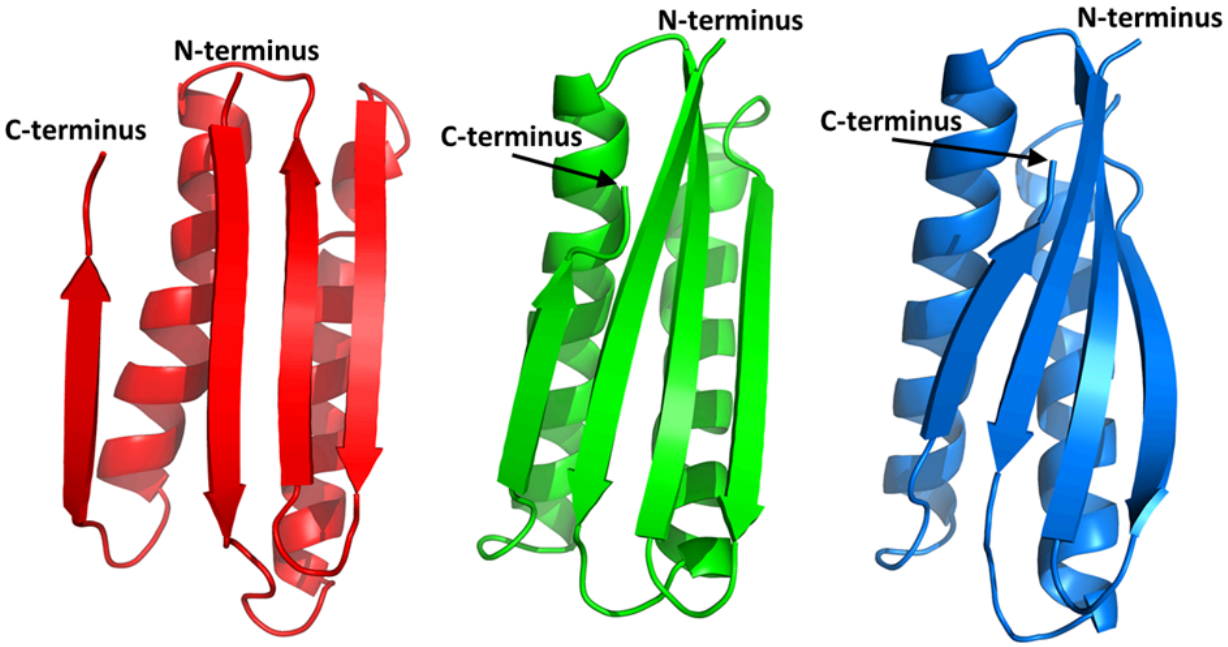


Figure 2. Presentation of the first model (out of five submitted for evaluation) for an exemplary protein predicted by the developed method based on the coarse-grained UNRES force field (red), experimental structure (green) and the highest-rated structure predicted by statistics-based methods (blue) (figure from work H1). Reproduction rights obtained (reprinted with permission from Bioinformatics, Volume 32, Issue 21, November 2016, Pages 3270-3278. Copyright 2024 Oxford University Press, license 5740260555795).

In order to improve the resolution and eliminate the need to use external information during simulations in the form of constraints on the β-type secondary structure (Figure 1), I decided to optimize the energy function in the UNRES force field using the maximum likelihood method (H2). On the other hand, in order to limit the effect of resolution decrease as a result of using a coarse-grained model, noticeable during the reconstruction of all-atom models, I decided to combine coarse-grained and all-atom molecular dynamics simulations to obtain models with atomistic accuracy, while ensuring verification of the obtained structures (works H3, H11 and H13).

H2. Optimization of the UNRES Force Field

To perform the optimization of the weights of the contributions to the energy function in the UNRES force field, I selected a set of structurally diverse small proteins with known experimental structures (Figure 3). The training set consisted of ensembles of protein conformations obtained using nuclear magnetic resonance (NMR) spectroscopy (from 17 to 300 conformations for each system), which allowed for the inclusion of conformational fluctuations observed in experimental conditions, closely resembling physiological conditions. To better account for the fact that a single protein can adopt diverse conformations, experimental data were obtained over a wide range of temperatures for three systems. This approach is crucial for the proper optimization of the UNRES force field (and likely most force fields), as it (i) minimizes the risk of overtraining, (ii) accounts for the temperature dependence of protein conformation, and (iii) does not negatively affect the method's ability to predict the structure and dynamics of inherently disordered proteins. The temperature dependence of conformation is particularly useful when performing molecular dynamics simulations with replica exchange, which are a fundamental tool for predicting protein structures in the UNRES force field (H1).

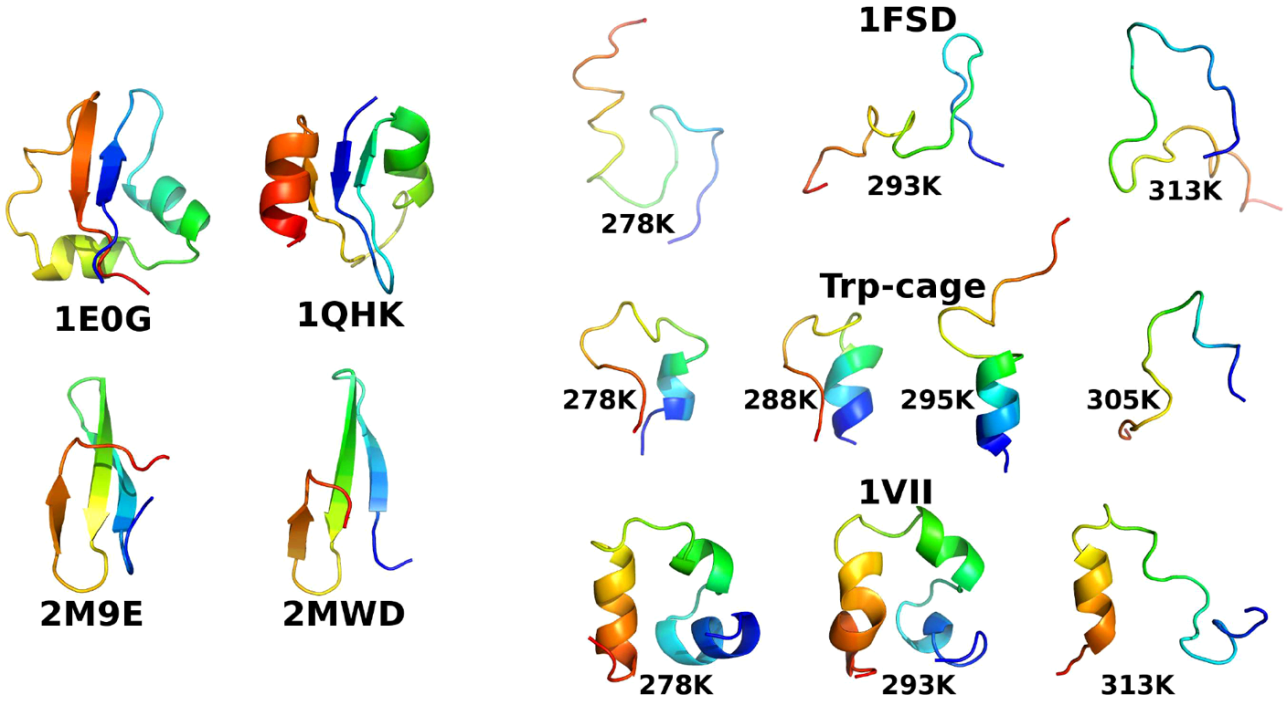


Figure 3. Visualization of the protein structures (first models from the ensembles) used for the optimization of the UNRES force field (P3). Reprinted with permission from J. Chem. Inf. Model. 2017, 57, 9, 2364–2377. Copyright 2024 American Chemical Society.

A test set of proteins was also carefully prepared, consisting of 46 proteins containing 12 to 126 amino acid residues. This set included: 22 proteins with helical structures, 12 proteins with β-type structures, and 12 proteins with α + β structures. These proteins were characterized by low sequence similarity to the training proteins and other test proteins, ensuring the most representative results for the predictive performance of conformation prediction.

The optimization procedure consisted of several steps: (i) molecular dynamics simulations with Hamiltonian replica exchange for the training proteins, (ii) selection of the most probable structures for specific temperatures using the weighted histogram analysis method (WHAM), (iii) optimization of different sets of parameters using the maximum likelihood method, (iv) molecular dynamics simulations with replica exchange for the training and test proteins, (v) selection of the most probable structures for specific temperatures using WHAM, and (vi) evaluation of the results.

The primary simulation tool used for parameter optimization during the training phase was Hamiltonian replica exchange molecular dynamics, based on two parameters: "soft" restraints of

varying strength imposed on conformations based on experimental data and temperature. This approach, known as umbrella sampling, was necessary to thoroughly explore the conformational space of the training systems. Each training simulation consisted of at least 384 trajectories, each with 80,000,000 steps, with a step size of 4.89 fs (not accounting for the acceleration due to the coarse-grained force field), resulting in a total simulation time of approximately 150 ms. In the next step, the weighted histogram analysis method was applied to the final part of the trajectories, where the proteins had reached thermodynamic equilibrium, to select the most probable conformations with the lowest free energy at given temperatures.

To determine the optimal optimization protocol, I tested several variants where only a subset of parameters was optimized, including general weights of individual potentials, internal weights of local and correlation potentials, radii of side chains, and parameters describing their anisotropy, as well as depths of potentials describing interactions between side chains. This approach aimed to find the minimal set of parameters for optimization that would significantly improve predictive capabilities without overfitting the parameters on the training set of proteins.

Next, I used the improved protocol for predicting protein structures without using external information other than the protein sequence, developed in work H1 (Figure 1), to evaluate the performance of the optimized force field. For this purpose, I performed molecular dynamics simulations with replica exchange over a wide temperature range (225-525 K), with each of the 72 trajectories consisting of 50,000,000 steps, resulting in a total simulation time of approximately 17.6 ms. Then, I performed weighted histogram analysis to select the most probable structures at temperatures 10 K below the observed maximum heat capacity, and subjected the obtained ensemble of structures to cluster analysis to select representative structures. In the final step, I analyzed the similarity of the obtained models to experimental structures.

The optimization significantly improved the predictive capabilities of the UNRES force field, particularly for α + β and β-type proteins, while the predictive performance for α-type proteins improved only slightly, mainly because previous versions of the force field already yielded good results for these structures. The optimization also enabled the use of a single set of parameters for all protein classes without the need for restraints on secondary structure during conformation prediction (Figure 1), and improved the distribution of folded and unfolded structures as well as the ability to determine the phase transition temperature for proteins. The optimization, thanks to the use of experimental data for partially disordered proteins, also allowed for the use of the force field for proteins with poorly defined structures, such as the monomeric and oligomeric forms of amyloid β.

H3. Studies on the Aggregation of Amyloid β at Different Levels of Resolution

In the next step, I utilized a combination of coarse-grained and all-atom methods to investigate the aggregation pathways of amyloid β (H3). For this purpose, I supervised the performance of coarse-grained molecular dynamics simulations with replica exchange (REMD) in the UNRES force field for four chains of amyloid β 1-42 (Aβ42) placed in different orientations relative to each other. Using the REMD method, the system dissociated in high-temperature trajectories and then re-associated in lower-temperature trajectories, allowing the examination of the protein's aggregation mechanism. The total effective simulation time was approximately 50 ms, considering the acceleration due to the coarse-grained representation, which is 3-4 orders of magnitude faster. This enabled the acquisition of sufficiently large statistical ensembles, reflecting multiple association-dissociation processes, to identify potential different pathways of these phenomena. The analysis of the graphs showed that the primary pathway for forming Aβ42 tetramers is the aggregation of two dimers (Figure 4). This process is significantly more frequent than the formation of tetramers from monomers and trimers.

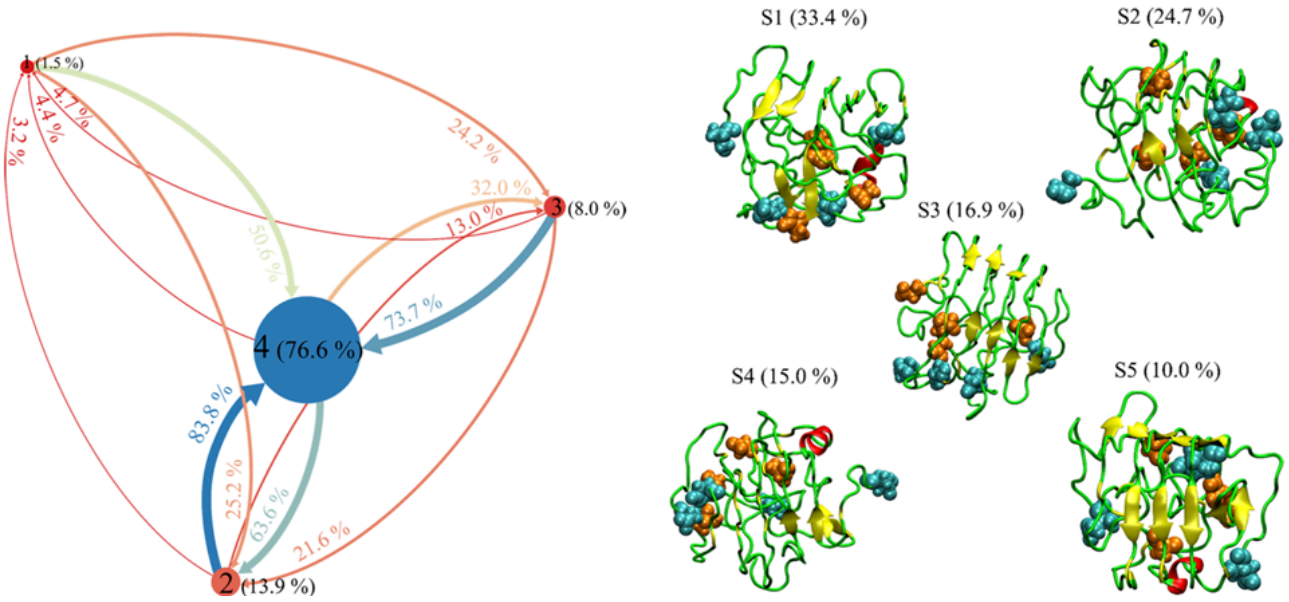


Figure 4. Schematic representation of the amyloid β aggregation mechanism (left part of the figure) and visualization of the five representative structures of the tetramer model (right part of the figure) obtained from the coarse-grained simulations. The figure is from work H3. Reprinted with permission from J. Phys. Chem. B 2019, 123, 34, 7253–7269. Copyright 2024 American Chemical Society.

The conducted studies revealed that amyloid β forms stable tetrameric structures, consisting of a core and an outer part, each built from two chains, with their properties, structure, and dynamics significantly differing from each other. This results from the formation of a hydrophobic core through the internal (central) chains of the tetrameric form and two chains with significantly greater surface contact with the solvent. Interestingly, the most probable tetramer model of Aβ42 (cluster 3) obtained from the coarse-grained simulations exhibits a disk-like structure (oblate spheroid) and does not display characteristics of cross-β structure, typical for fibrils.

Furthermore, we demonstrated that using the experimental protofibril structure of amyloid β (PDB code: 2NAO) does not constitute a good initial model for simulating oligomeric structure. The tetrameric form of amyloid β, obtained by removing all chains from the protofibril model except for the four adjacent ones—a practice used in some scientific studies—proved unstable and underwent significant conformational changes during the simulations, resulting from the exposure of the hydrophobic core of the fibril to water interactions. This result, although expected, indicates the error in assuming that available experimental data for another form of a peptide or protein always serve as a good starting point for simulating other variants.

The decomposition of the energy components responsible for the stability of the Aβ42 tetramer showed that, unlike the monomeric form (H4), van der Waals interactions are the primary type of interactions stabilizing the internal peptides of amyloid β. Unchanged compared to the monomeric form, the interactions of the Aβ42 tetramer with water constitute the dominant contribution to its energy and thus the dominant factor influencing its structure. The interactions of the four chains cut from the protofibrillar form of amyloid β (2NAO) with water are comparable to those of the tetrameric form obtained from the coarse-grained simulations but are characterized by strong electrostatic repulsion of the amyloid β chains, which is only partially compensated by van der Waals interactions. This behavior indicates the potential instability of the structure of such an obtained tetramer model of amyloid β. In the case of tetrameric models obtained from simulations in the coarse-grained UNRES force field, the internal electrostatic interactions are usually close to zero, while the van der Waals interactions are about twice as strong as in the model obtained from the protofibril structure. Additionally, due to the much stronger interaction with water, the Aβ42

tetramer has a lower content of β-type structures, and the charge distribution is more isotropic than in fibrils.

H4. Evaluation of All-Atom Force Fields and the Influence of Different Interactions on the Monomeric Form of Amyloid β

In parallel to the coarse-grained studies, I conducted extensive analyses using all-atom force fields. For this purpose, I selected a small system whose conformational space can be efficiently explored during calculations involving all atoms—the monomeric form of amyloid β, composed of 42 amino-acid residues. This peptide does not have an experimentally resolved structure due to its inherently disordered nature (IDP), yet it plays a crucial role in the development of Alzheimer's disease, which is accompanied by the presence of toxic oligomeric and fibrillar forms of amyloid β.

The goal of the research was not only to evaluate the ability of popular all-atom force fields to accurately predict the conformational ensemble of the amyloid β monomer but also to determine the influence of conformational sampling methods and individual energy components on the stabilization or destabilization of the peptide. To this end, I selected five widely used all-atom force fields: three from the Amber family (universal force fields: ff99SB and ff14SB, and a force field dedicated to disordered proteins: ff14SB_IDPs) and two from the CHARMM family (CHARMM36 and CHARMM36m, with the latter including minor improvements over the former to enhance the accuracy of simulations for both folded and disordered systems).

The conducted simulations unequivocally showed that newer versions of the force fields exhibit higher accuracy in predicting the peptide structure compared to available experimental data, such as the presence of secondary structures, radius of gyration, and chemical shift values. The worst results were obtained for the oldest of the tested but still popular force field, ff99SB, which significantly overestimated the content of β-structure (34.0 ± 7.0%).

A comparison of the results from molecular dynamics simulations with replica exchange (48 trajectories in the temperature range 278.00-373.77 K, each 0.6 μs long, with a total simulation time of 28.8 μs for one force field) with a single 10 μs trajectory of classical molecular dynamics showed that improved sampling enhances agreement with most of the experimental data. The average correlation coefficient of predicted chemical shifts versus experimental values was below 0.90 (0.845-0.898) for classical force fields using single-trajectory sampling (Amber ff99SB, ff14SB, and CHARMM36), significantly increasing after applying replica exchange molecular dynamics (0.903-0.945). The improvement in the agreement of simulation results with experimental data also occurred for newer force fields, although it was not as significant, indicating that a more accurate description of the system facilitates obtaining results close to reality without the need for enhanced sampling techniques. This suggests that newer force fields correctly describe the small energy barriers between different conformations of the disordered monomeric form of amyloid β, allowing them to be overcome even with classical molecular dynamics simulations.

The conducted studies also revealed that Amber and CHARMM force fields are based on different assumptions. While the properties of the system in Amber force fields depend mainly on intra-peptide interactions, in CHARMM force fields, peptide-water interactions play a crucial role (Figure 5), regardless of the sampling method used. The stronger interaction of amyloid β with the solvent in CHARMM force fields results in the peptide being less stable and more hydrophilic compared to Amber force fields. It is also worth noting that for the monomeric form of amyloid β, electrostatic interactions dominate the potential energy, being an order of magnitude larger than van der Waals interactions.

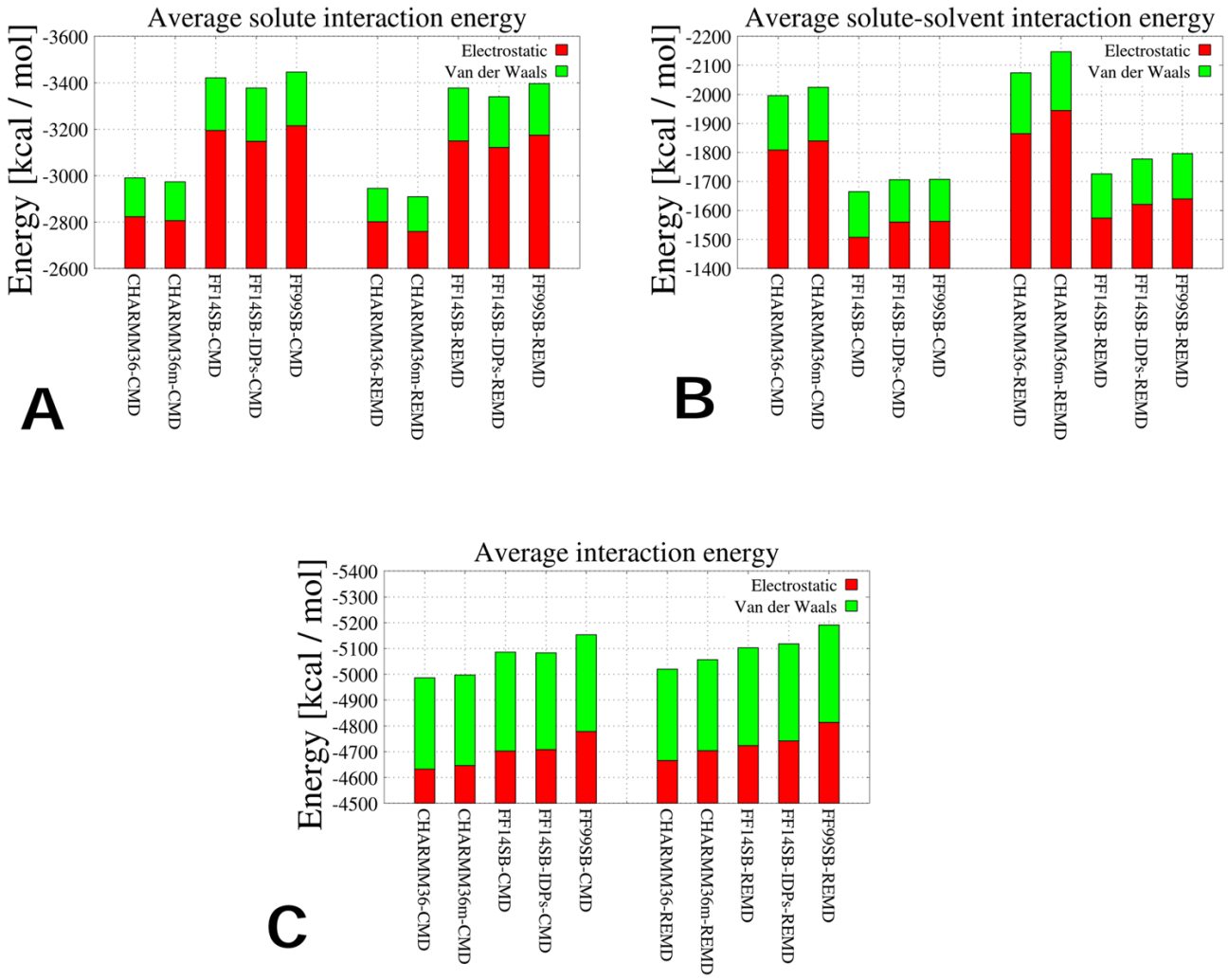


Figure 5. Visualization of the differences in energy contributions of interactions within the peptide and between the peptide and water in classical molecular dynamics (CMD) and replica exchange molecular dynamics (REMD) simulations. Figure from work H4. Reprinted with permission from J. Chem. Phys. 151, 055101 (2019). Copyright 2024 AIP Publishing, license 5744180590277).

Interestingly, despite the completely different distributions of individual potential energy components, which translate into different tendencies of amyloid β to adopt extended conformations, the analysis of structures corresponding to the minima of free energy showed a very high similarity in the representative conformations (most frequently occurring metastable forms) obtained for Amber ff14SB and CHARMM36m force fields (Figure 6).

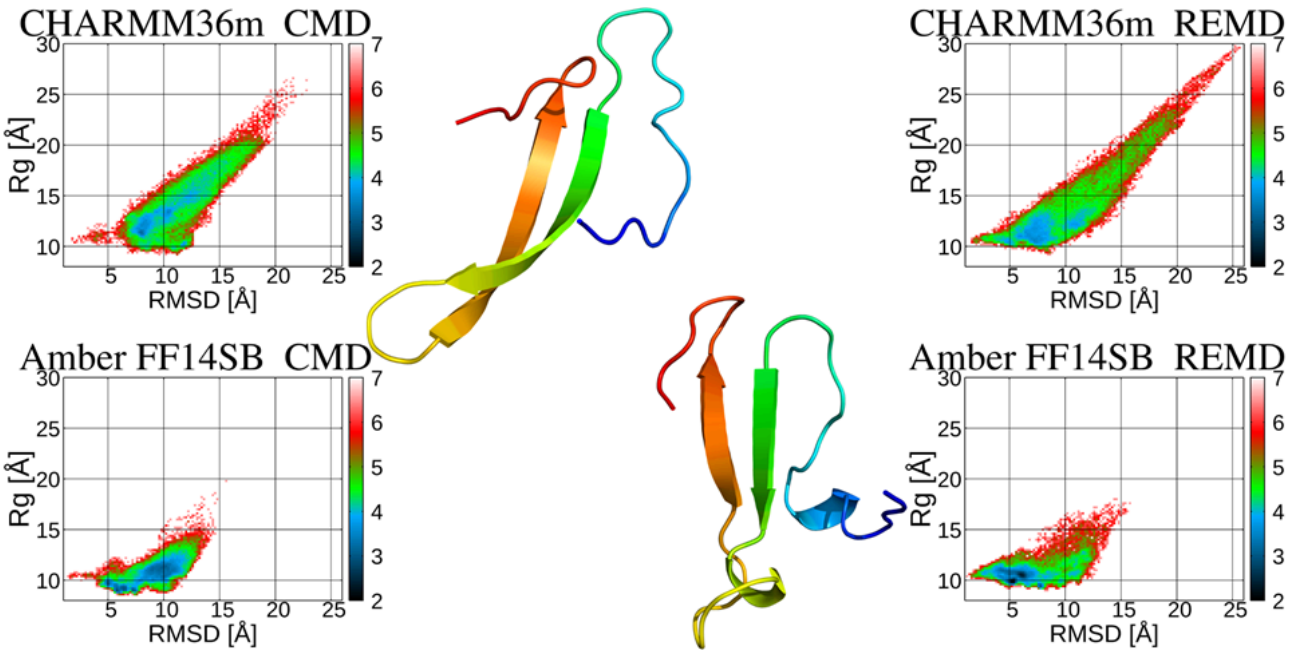


Figure 6. Visualization of the differences in structural properties of amyloid β in classical molecular dynamics (CMD) and replica exchange molecular dynamics (REMD) simulations, along with the representation of the representative CHARMM and Amber structures in a cartoon format. Figure from work H4. Reprinted with permission from J. Chem. Phys. 151, 055101 (2019). Copyright 2024 AIP Publishing, license 5744180590277).

The conducted studies confirmed previous observations that the use of replica exchange molecular dynamics improves the ability of force fields to identify the most probable conformations of amyloid β. However, it was found that using newer force fields or those dedicated to disordered proteins increases the efficiency of sampling, reducing the difference between classical molecular dynamics and the replica exchange variant. This is an important observation because classical molecular dynamics simulations can be successfully performed using graphics processing units (GPUs), which accelerate calculations by up to 100 times compared to typical central processing units (CPUs), while their application in replica exchange simulations is significantly more complicated and yields a much smaller increase in performance. Additionally, especially in the case of the CHARMM36m force field, replica exchange molecular dynamics likely leads to obtaining too many highly extended structures with high radius of gyration values. I thus demonstrated that an intermediate approach involving a series of several independent but relatively long classical molecular dynamics simulations can be an optimal combination of conformational space exploration efficiency with hardware and time requirements.

Moreover, achieving results with high agreement with experimental data and the similarity of the most probable metastable structures between Amber and CHARMM force fields unequivocally indicates the utility of these methods for simulating amyloid β. These results suggest that with the appropriate choice of force field and sampling method, all-atom simulations can provide reliable information about the structure and dynamics of the monomeric form of amyloid β, which is crucial for understanding the aggregation mechanisms of this peptide and its role in Alzheimer's disease.

H5. Studies on the Mechanism of Inhibition of Amyloid β Aggregation by Isoquinoline Alkaloids

The next study provided a theoretical explanation for the experimental observations regarding the different effects of selected isoquinoline alkaloids on amyloid β aggregation. Collaborating

experimental research groups, led by Giovanni N. Roviello, determined that two similar isoquinoline alkaloids, sanguinarine and coralyne (Figure 7), have entirely different effects on amyloid β aggregation. Specifically, sanguinarine possesses strong inhibitory properties, while coralyne accelerates the aggregation process.

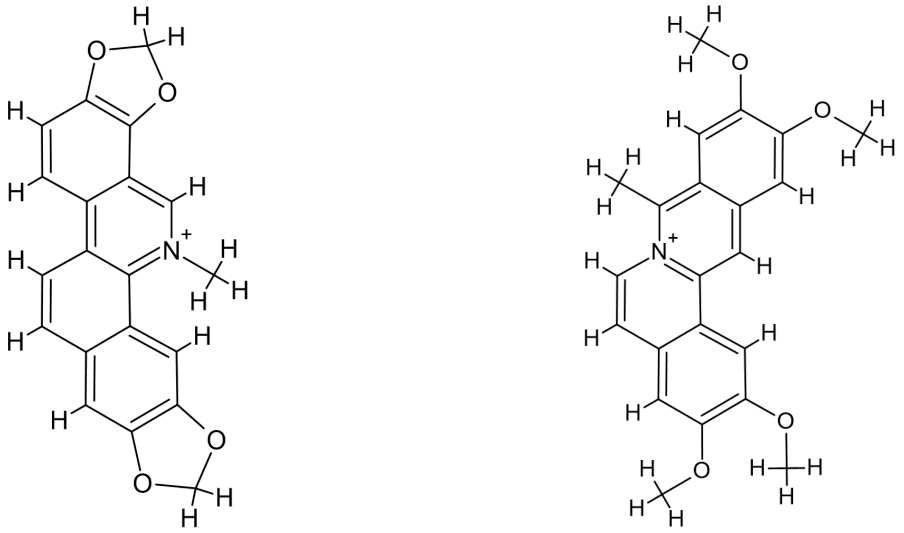


Figure 7. Structural formulas of the studied molecules: sanguinarine (left panel) and coralyne (right panel); own drawings.

To explain the observed differences in the influence of sanguinarine and coralyne on amyloid β aggregation, I performed a series of molecular docking simulations of these molecules to the previously obtained forms of amyloid β: three representative models of the monomer (H4), tetrameric (H3), and protofibrillar (U-type, PDB code: 2LMN, and LS-type, PDB code: 2MXU). The conducted studies showed that both molecules bind to all forms of amyloid β with comparable strength and form similar types of interactions, confirming the experimental observations but not explaining the different effects on amyloid β aggregation. The main observed difference in the binding of isoquinoline alkaloids to different forms of amyloid β is the different number of binding sites—sanguinarine docks more selectively than coralyne, which can interact with more diverse fragments of amyloid β. Subsequent molecular dynamics simulations for sanguinarine and coralyne interacting with two protofibril models confirmed nearly identical free energy values of interaction, indicating that both molecules strongly interact with amyloid β.

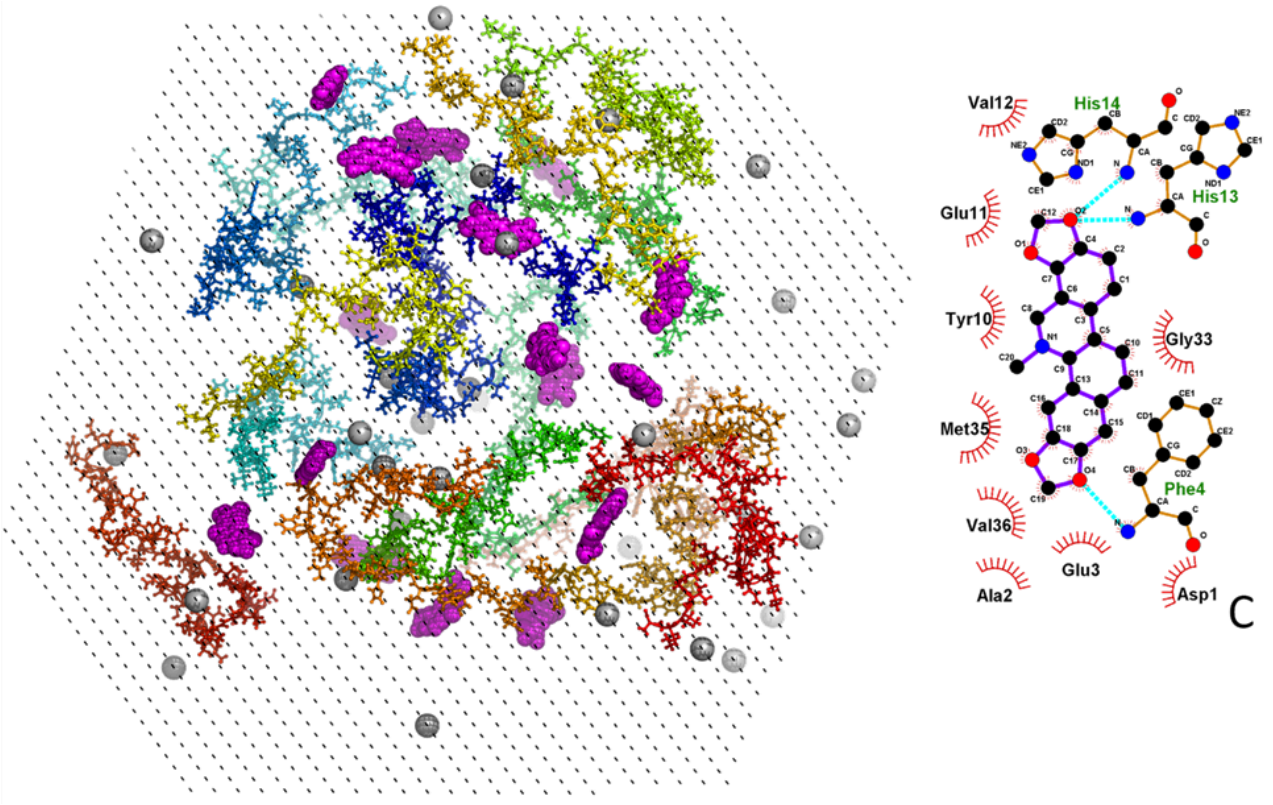


Figure 8. Presentation of the initial setup consisting of 12 amyloid β chains (skeletal structures in 12 colors) along with water molecules (small dark spheres), counterions (gray spheres), and the studied isoquinoline alkaloids (purple spheres; left part of the figure) and a schematic visualization of the interactions between amyloid β and the selected isoquinoline alkaloid (right part of the figure). Figures from work H5. Reprinted with permission from Chemico-Biological Interactions Volume 334, 25 January 2021, 109300. Copyright 2024 Elsevier.

To understand the mechanism of the influence of these molecules on amyloid β aggregation, I performed molecular dynamics simulations for a system consisting of 12 amyloid β chains placed in a 1:1 ratio with sanguinarine and coralyne (Figure 8). The initial arrangement of the molecules was specially prepared so that all components of the system were at least 8 Å apart to allow for the rotation of the molecules before potential association. It is necessary to note that the obtained concentration of amyloid β in the molecular dynamics simulations is higher than those used in experimental studies but does not reach the glass phase. Using such high concentration of amyloid β had a direct relation to the simulation time, which, to capture physiological (pathological) conditions, would have to be on the order of hours and days, which are computationally unattainable even with coarse-grained methods. However, increasing the concentration allowed for observing differences in the early stages of the aggregation process on the microsecond timescale. The analysis of the simulations showed that sanguinarine increases the tendency of amyloid β to form helical structures while reducing its tendency to form β-type structures, which is opposite to the effect of coralyne. Previous studies have shown that the content of β-type structures in monomeric and small oligomeric forms is a crucial factor influencing the aggregation rate of amyloid β. Interestingly, the analysis of the obtained oligomeric forms showed that tetramers and heptamers are the dominant types of oligomers in the absence of isoquinoline alkaloids, confirming our previous assumption that tetrameric forms are an important stage in amyloid β aggregation (H3), while the presence of coralyne in the aggregation simulations reduces the tendency of amyloid β to form higher-order oligomeric structures.

H6. Obtaining of the Structure and Structural Studies of the Heterodimer of α-Synuclein and Amyloid β42

The next step involved investigating the interactions between the inherently disordered monomeric forms of α-synuclein (α-Syn) and amyloid β42 (Aβ42) and the impact of these interactions on their structure and dynamics, as well as comparing them to their monomeric and homodimeric forms (H6). An additional goal was to evaluate the feasibility of molecular dynamics simulations using implicit and explicit solvent models in popular force fields like Amber and CHARMM. The first type of solvent allowed the use of replica exchange molecular dynamics, while simulations with explicit water were performed in the classical molecular dynamics variant, but with the use of graphics processing units (GPUs), which significantly accelerated the simulations.

To achieve effective conformational sampling, the simulations of the α-synuclein-Aβ42 heterodimer began with the generation of 20 highly diverse initial positions of the molecules relative to each other, using the initial algorithms employed in the UNRES-Dock method developed by me (H12) (Figure 9). The combination of replica exchange molecular dynamics with diverse initial structures enabled the dissociation of unstable orientations in high-temperature replicas, thus allowing for the convergence of simulations within a total time of 40 μs (20 replicas, each 2 μs long).

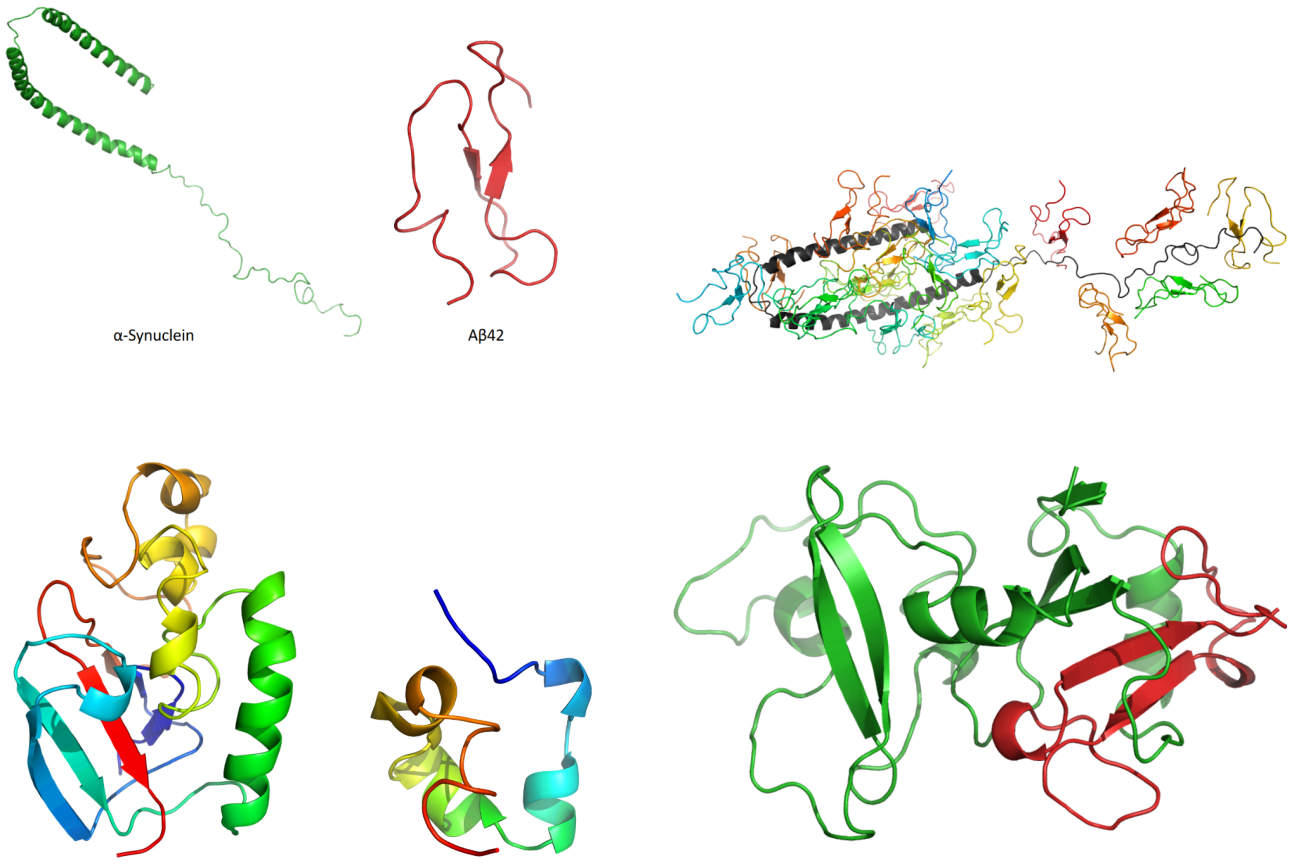


Figure 9. Visualization of the structures of α-synuclein (left panels), amyloid β (middle panels), and their heterodimer (right panels): initial (top row) and most probable (bottom row) based on simulation analysis. Synthesis of two figures from work H6. Used and modified under CC BY 4.0 license.

The analysis of the obtained results showed that the selected all-atom force field with an implicit solvent model (AMBER ff14SBonlysc with the GB-Neck2 model) tended to generate overly compact and structured conformations, particularly evident for the monomeric form of

α-synuclein. However, for the α-synuclein-Aβ42 heterodimer, the most probable structure remained stable in both implicit and explicit solvent simulations using the AMBER ff14SBonlysc with the GB-Neck2 implicit solvent model, and explicit solvent simulations with AMBER-FB15 and CHARMM36m force fields. The obtained sets of structures for the monomeric form of Aβ42 fell within the ranges observed in previous experimental and theoretical studies.

Additionally, I performed a comparative analysis of the impact of Aβ42 on the aggregation propensity of α-synuclein in simulations with implicit and explicit solvent models. I found that the presence of Aβ42 significantly increased the content of β-type structures in α-synuclein in explicit solvent simulations, consistent with experimental data indicating that Aβ42 accelerates α-synuclein aggregation. This effect was not observed in implicit solvent simulations, further confirming the significant overestimation of the stability of the monomeric form of α-synuclein in this approach.

To identify the most probable conformation of the α-synuclein-Aβ42 heterodimer, I conducted a detailed analysis of the structural and physicochemical properties of both the monomeric and heterodimeric structures. The results showed that the binding of the molecules significantly affected their mobility and flexibility. Interestingly, I also discovered that the presence of α-synuclein had a much greater impact on the conformation of Aβ42 than the reverse situation, which I attribute to the hydrophobic nature of the inter-chain interactions and its larger size.

Emphasizing the crucial role of solvent models in protein aggregation modeling, I demonstrated that the impact of Aβ42 on α-synuclein's aggregation propensity is accurately captured by explicit solvent models but not by implicit ones. This observation confirms previous findings (H3 and H4) about the significant influence of protein-solvent interactions, especially on proteins lacking stable secondary structures under physiological conditions, such as the monomeric forms of Aβ42 and α-synuclein.

Using the MM-PBSA method to estimate the free energy of binding between molecules based on molecular dynamics trajectories, I compared the α-synuclein-Aβ42 heterodimer with models of Aβ42 and α-synuclein homodimers, showing that it binds about twice as strongly as Aβ42 dimers but slightly weaker than α-synuclein homodimers. This suggests that α-synuclein and Aβ42 can indeed form stable complexes, potentially serving as "nuclei" for fibrillar structures and competing with the aggregation process of Aβ42.

The molecular dynamics simulations of the α-synuclein-Aβ42 heterodimer also revealed that the fragments 16-19 and 29-34 of Aβ42 are most involved in interactions with α-synuclein, consistent with the β-strand regions in Aβ42 peptides and oligomers identified in previous studies.[14] My findings are also consistent with previous predictions made through discrete molecular dynamics (DMD) simulations indicating the interaction sites between α-synuclein and Aβ42.[15]

In summary, in work H6, I demonstrated the possibility of forming stable α-synuclein-Aβ42 heterodimers, which may play a crucial role in the early stages of aggregation of these proteins associated with neurodegenerative diseases. The provided structure (in PDB format) of the most probable α-synuclein-Aβ42 heterodimer can be used for further studies on the structure, dynamics, and design of potential inhibitors for the aggregation of these proteins.

H7. Study on the Influence of Metal Ions on Lipid Membranes with the Presence of Amyloid β

In the next study, I demonstrated that simple models can be effectively used to illustrate complex physical phenomena. Within the project, I determined the affinity and energy barriers associated with the penetration of a homogeneous lipid bilayer model by various mono- and divalent metal ions, as well as the conformational differences of tetrameric forms of amyloid β in the presence and absence of lipid membranes and copper ions.

In this work, I applied a recently proposed more realistic model for divalent cations using 12-6-4 potentials instead of the typical 12-6 potentials, known as the Lennard-Jones potential, to

determine the affinity of ions to the dimyristoylphosphatidylcholine (DMPC) bilayer. This model partially mitigates the issues arising from the point charge treatment of ions and should allow for a good description of changes in hydration of charged and polar groups upon ion binding to lipid atoms.

The studies consisted of a series of molecular dynamics simulations in the classical variant and with umbrella sampling. In the first step, I arranged models of amyloid β dimers without copper ions and with copper ions in as many diverse ways as possible, using the simplified UNRES-Dock protocol (H12). The 128 different initial positions obtained were subjected to all-atom molecular dynamics simulations, and then, using the binding energy between dimers as a preliminary selection criterion, I performed cluster analysis of a subset of structures to identify the five most probable models of amyloid β tetramers. These tetramer models of amyloid β were placed near a homogeneous lipid bilayer model, and I observed how the presence of the bilayer affects the structure and stability of the amyloid β tetramers and how they interact with and influence the lipid membrane. Since the binding of copper ions, present in high concentrations in the vicinity of synaptic membranes, is crucial for both the stability of amyloid β tetramers and lipid membranes, in the second step, I performed umbrella sampling and classical molecular dynamics simulations to determine the positions and energy barriers associated with the penetration of the lipid bilayer by mono- and divalent metal ions. Based on the simulations, I analyzed the weighted histograms and determined the potential of mean force (PMF) for the interaction with the lipid bilayer (Figures 10 and 11).

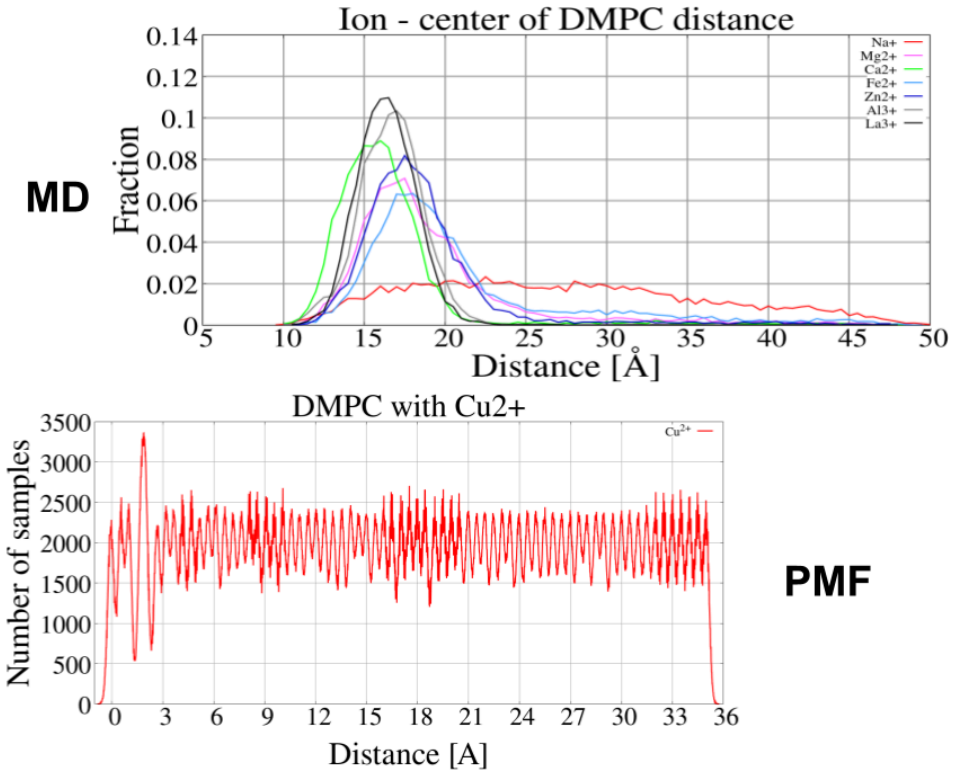


Figure 10. Plots showing differences in sampling distances between metal ions and the center of the lipid bilayer for classical molecular dynamics simulations (MD) - Gaussian-like distribution and umbrella sampling method used in PMF calculations. Unpublished figures.

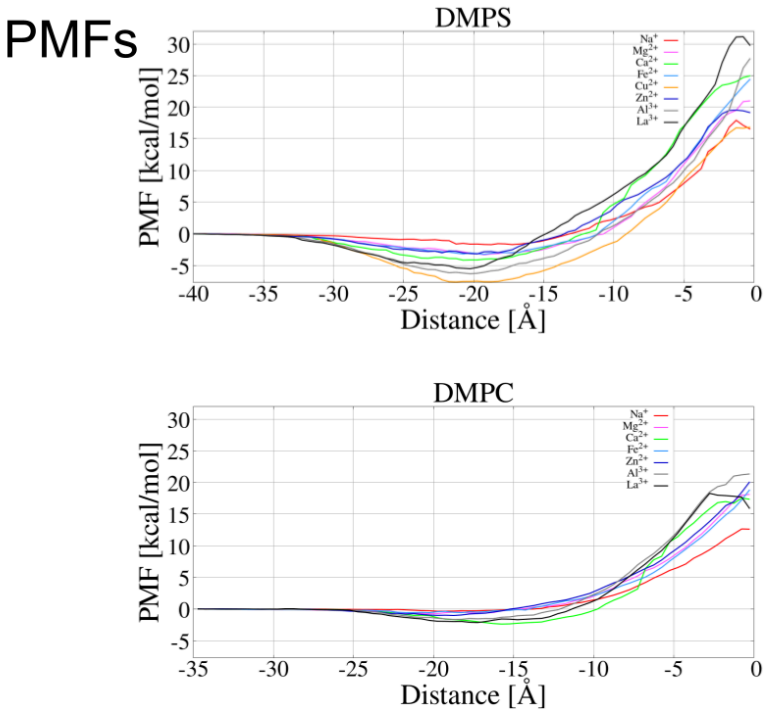


Figure 11. Plots of the mean force potentials as a function of distance for selected metal ions from the center of lipid bilayers: DMPS (top part of the figure) and DMPC (bottom part of the figure). Unpublished figures.

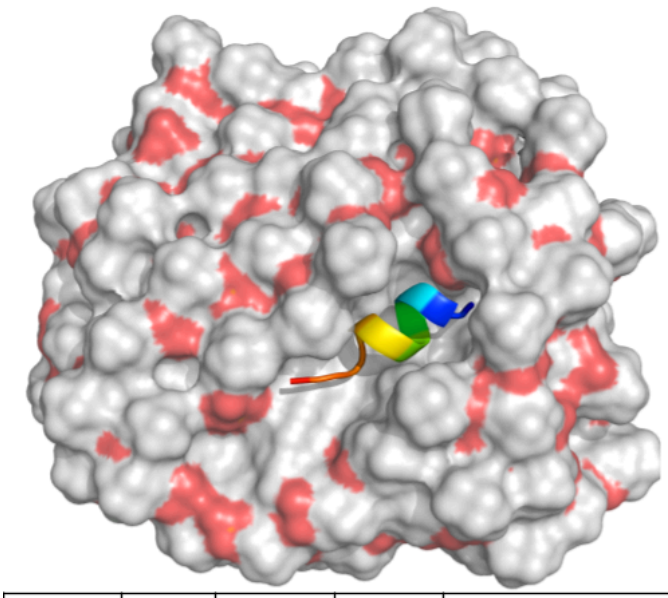
The analysis of the simulations showed that Cu^{2+} ions behave similarly to Ca^{2+} ions in terms of the type of interactions with DMPC and binding strength. However, Mg^{2+}, Fe^{2+}, and Zn^{2+} ions more strongly maintain their hydration spheres, exhibiting weaker interactions with the lipid bilayer. I also observed that the presence of copper ions interacting with amyloid β reduces the stability of amyloid β tetramers. Interestingly, amyloid β tetramers without bound copper ions interact more strongly with DMPC lipid heads, showing the possibility of reaching ion-binding regions and potential extraction of divalent cations from the lipid-water interface. Although the depth of penetration is similar to that achieved by monomeric forms of Aβ42 (P13), the effect of the tetramer on the membrane properties is larger than in the case of the Aβ42 monomer. However, when amyloid β tetramers have bound copper ions, their tendency to interact with the lipid-water interface decreases. This effect is particularly evident when copper ions form bridges between peptides, linking them into associated dimers, in contrast to situations where copper ions are present only within individual dimers.

In summary, the conducted studies indicate that the formation of copper complexes with different oligomeric forms of amyloid β may protect lipid membranes from destabilization and oxidation by free divalent cations, which are present in high concentrations in the synaptic environment, and the release of unstructured amyloid β peptides may be a mechanism for restoring ion and lipid membrane homeostasis. On the other hand, we observed that copper ions effectively restrict the ability of tetrameric amyloid β to penetrate the lipid bilayer, so the simultaneous presence of copper ions and amyloid β peptides mitigates the negative impact on lipid membranes observed with only one type of molecule.

H8. Influence of Lipid Membrane Composition on Herpesvirus Protein UL49.5

Herpesvirus proteins UL49.5 are an interesting subject of study due to their ability to evade the host's immune response by blocking the transport of peptides through the TAP (transporter associated with antigen processing) protein to the MHC-1 complex. Our previous experimental and theoretical studies showed that the UL49.5 protein from BoHV-1 (bovine herpesvirus 1) contains special anchoring regions in the membrane, which, upon mutation, cause significant structural changes and loss of inhibitory properties against TAP (P16). Despite achieving very good agreement between experimental and theoretical studies in that work (P16), it should be noted that these studies were conducted using a simple homogeneous lipid bilayer model of POPC, which does not reflect the physicochemical properties of the endoplasmic reticulum (ER) membranes where human TAP proteins are found. For this reason, in the next project (H8), we conducted studies using a lipid bilayer model containing two types of tails (DM and PO) and three types of lipid heads (uncharged PC and PE, and charged PS), which reflect the physiological composition of the ER membrane: POPC (1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine), POPE (1-palmitoyl-2-oleoyl-sn-glycero-3-phosphoethanolamine), POPS (1-palmitoyl-2-oleoyl-sn-glycero-3-phosphoserine), DMPC (1,2-dimyristoyl-sn-glycero-3-phosphocholine), DMPE (1,2-dimyristoyl-sn-glycero-3-phosphoethanolamine), and DMPS (1,2-dimyristoyl-sn-glycero-3-phosphoserine).

Simultaneously, for the first time in the literature, we obtained experimental structures of fragments of the UL49.5 protein from HSV-1 (human herpesvirus 1), which were subjected to molecular dynamics simulations in micelles (Figure 12). NMR studies showed that the N-terminal fragment of the UL49.5 protein from HSV-1 adopts a highly elastic, disordered structure in the extracellular part due to the presence of a large number of proline and glycine residues. In contrast, the transmembrane region consists of a single long α-helix, unlike its homolog from BoHV-1, which contains two shorter helices oriented at an angle of 90°.



Res number	Res name	Max SASA*	Ave SASA	Fraction exposed to solvent
1	ABU	167	17.93	0.107
2	GLN	223	95.99	0.430
3	ARG	274	79.27	0.289
4	LEU	201	20.93	0.104
5	LEU	201	27.97	0.139
6	ALA	129	47.05	0.365
7	ALA	129	23.04	0.179
8	GLY	104	7.74	0.074
9	HIE	224	59.97	0.268
10	ALA	129	74.95	0.581

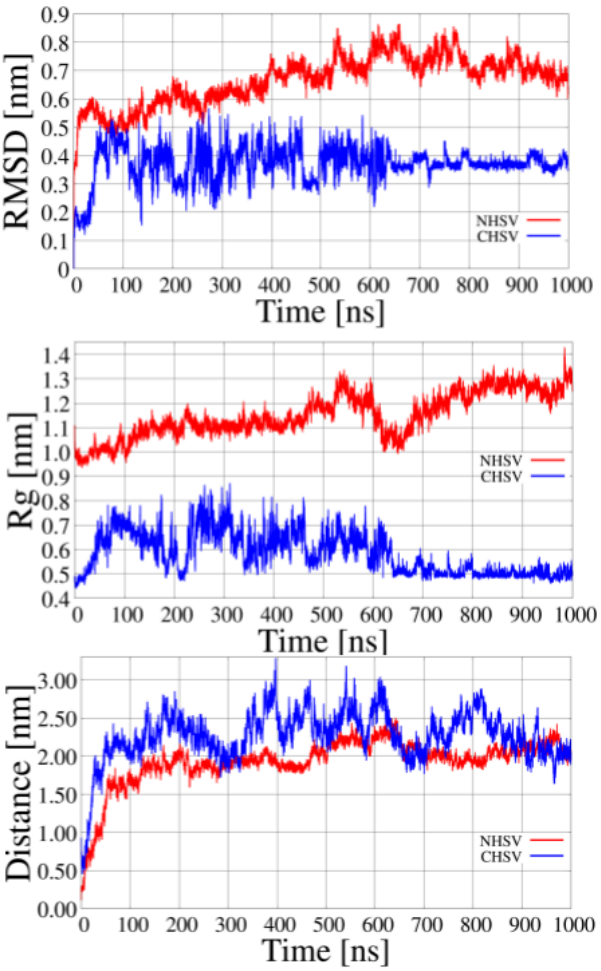


Figure 12. Visualization of the experimentally determined part of the UL49.5 protein during molecular dynamics simulations in DPC micelles (the micelle is presented as a solvent-accessible surface, where gray and red colors represent hydrogen and oxygen atoms exposed to the solvent, and the peptide is shown in a ribbon representation in rainbow colors) (upper left part of the figure); plots of RMSD, Rg, and distances from the center of the micelle for the two parts of the UL49.5 protein during molecular dynamics simulations in DPC micelles (dodecylphosphocholine) (right part of the figure) and presentation of the exposure of individual parts of the peptide to the solvent. Part of the figure from work H8. Used under CC BY 4.0 Deed license.

The equilibrium forms of the structures of the N- and C-terminal fragments of the protein obtained from molecular dynamics simulations in micelles were combined to reconstruct the complete structure of the UL49.5 protein from HSV-1. This model was placed in the heterogeneous lipid bilayer described above and subjected to a series of molecular dynamics simulations. The results based on the converged parts of the trajectories were compared with the UL49.5 protein from BoHV-1 (Figure 13). The conducted simulations showed that, particularly, the protein from BoHV-1 exhibits high specificity for binding to one type of lipid—charged PS heads, which causes additional stabilization of the N-terminal fragment of this protein compared to simulations in the homogeneous POPC model. Additionally, the UL49.5 protein from BoHV-1 shows greater affinity for the ER membrane and forms more interactions with lipid heads compared to the protein from HSV-1.

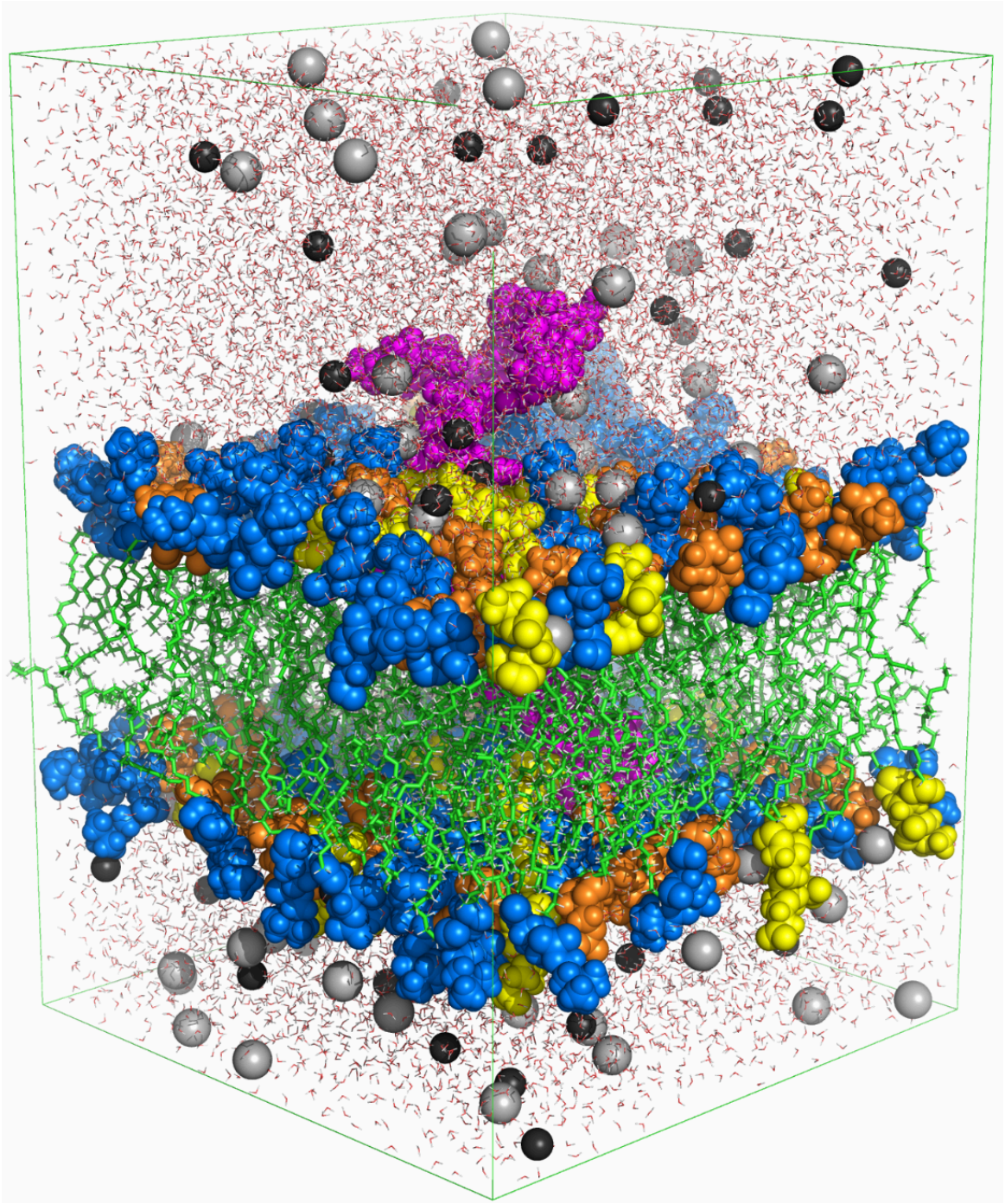


Figure 13. Visualization of the developed model of the UL49.5 protein from HSV-1 (pink spheres) in a heterogeneous lipid bilayer model (different lipid heads marked with different colored spheres), surrounded by Na+ and Cl- ions (light and dark gray spheres, respectively) and water (thin lines) in a periodic box (green outline). Figure from work H8. Used under CC BY 4.0 Deed license.

The analysis of the dynamics of the UL49.5 protein from HSV-1 showed high mobility of the N-terminal fragment due to the presence of numerous proline and glycine residues, which likely causes the loss of inhibitory activity against TAP (Figure 14). The protein from HSV-1 also has a long and rigid transmembrane helix, which, unlike the two shorter helices connected by an elastic linker in the UL49.5 protein from BoHV-1, fits poorly to the transmembrane part of the TAP protein. Furthermore, the C-terminal fragment of the UL49.5 protein from BoHV-1 protrudes from the luminal side of the membrane, while in HSV-1, it is completely embedded in the lipid bilayer.

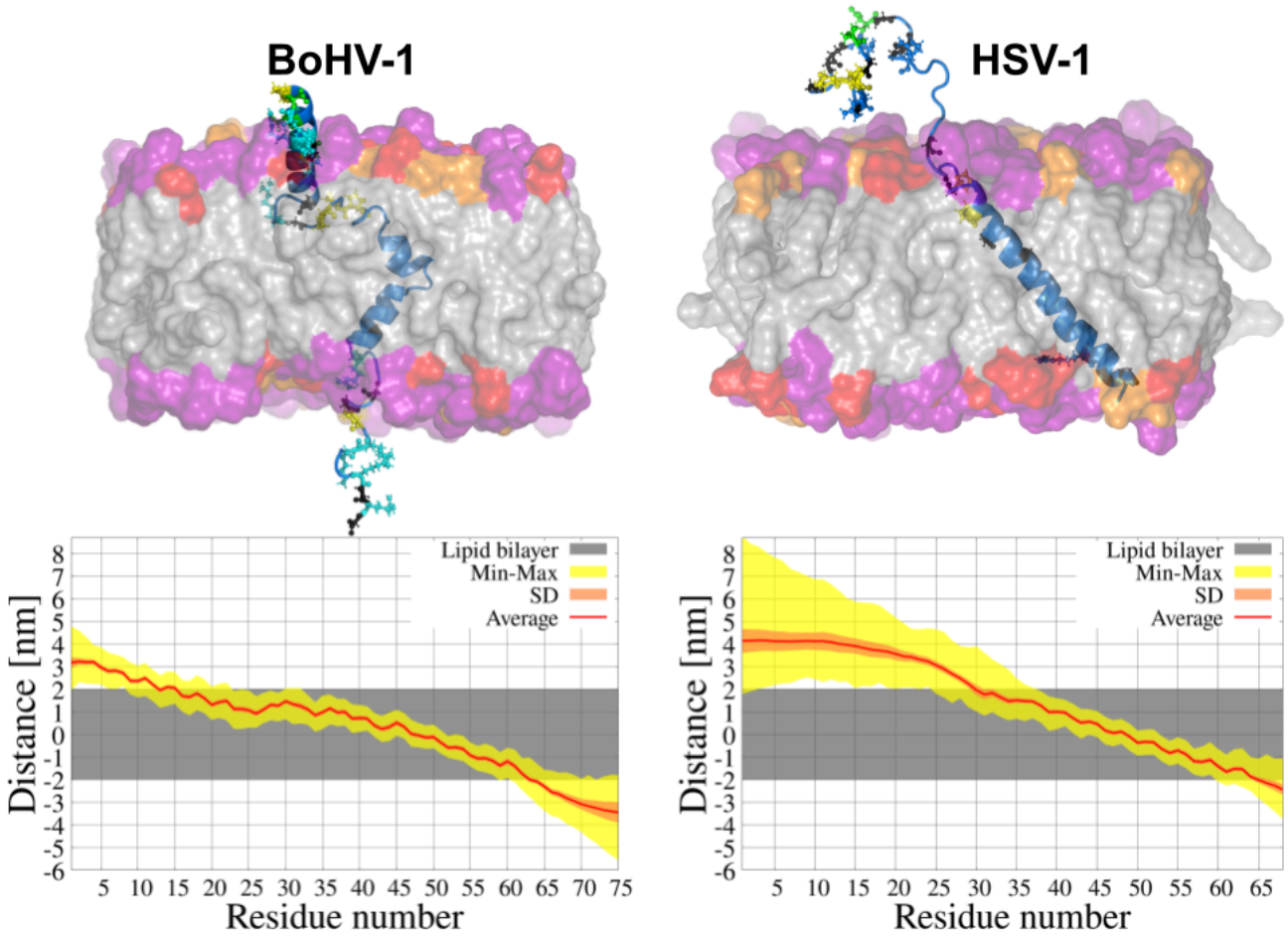


Figure 14. Visualization of the representative structures of the UL49.5 proteins from herpesviruses BoHV-1 and HSV-1 (upper part of the figure) and plots of mean distance (red line), standard deviation (orange contour), and observed minimum-maximum values (yellow contour) from the center of the lipid bilayer (lower part of the figure). Figure from work H8. Used under CC BY 4.0 Deed license.

Comparing the structure of the N-terminal fragment of both proteins, it is clear that UL49.5 from BoHV-1 forms a helix, while UL49.5 from HSV-1 adopts a mainly disordered structure with a single, not very stable α-helix turn. Additionally, analyzing the mobility of this part of both proteins, it can be seen that UL49.5 from BoHV-1 forms a stable N-terminal helix with low mobility, which undergoes partial deformation in the hydrophilic part of the membrane and is maintained in the membrane by anchoring with amino acid residues 30RRE32. In the case of the protein from HSV-1, the N-terminal fragment is extremely flexible due to the numerous Gly and Pro residues, forming variable structures during molecular dynamics simulations and orienting in different directions relative to the membrane surface. The N-terminal part of the UL49.5 protein from HSV-1 does not have anchoring residues that keep the entire protein in the membrane, as is the case for the protein from BoHV-1. This is another significant difference in structure, behavior, and biological activity of the proteins encoded by different viruses.

The structural analysis of the UL49.5 protein encoded by both viruses shows significant differences in the transmembrane part. The UL49.5 protein from BoHV-1 has two short α-helices oriented at an angle of about 90°, while the one encoded by HSV-1 has a single long transmembrane helix, although the fluctuations of this part of the protein from both viruses are similar. Interestingly, both proteins have proline residues in the N-terminal part of the transmembrane helix, which, as

shown in previous work (P16), are crucial for the biological activity of the UL49.5 protein from BoHV-1. However, it is not certain how proline residues in this position affect the biological activity of UL49.5 from HSV-1, but it is known from the simulations that they stiffen the first turn of the transmembrane helix.

In the previous work (P16), we conducted molecular dynamics simulations of the UL49.5 protein from BoHV-1 in a simple homogeneous POPC bilayer model. Compared to the results of molecular dynamics simulations in a multi-component ER membrane model, we did not observe significant changes in structure, except for the C-terminal part of the protein (93RGRG96). In the POPC bilayer, the C-terminal end of the UL49.5 protein is hidden in the membrane, while in the multi-component ER model, this part of the protein protrudes from the membrane on its luminal side. Based on previous observations, this C-terminal motif should activate the entire degradasome pathway.

Previous studies emphasize the role of the membrane-anchored N-terminal domain of the UL49.5 protein from BoHV-1 as the most important part in inhibiting TAP activity. We previously showed (P16) that the N-terminal domain contains a PPQ motif, which is necessary for direct binding to TAP or regulating the conformation of UL49.5 required for TAP binding. The UL49.5 protein from HSV-1 has a disordered, elastic N-terminal domain lacking a motif similar to PPQ. Furthermore, during docking of the UL49.5 protein from HSV-1 to TAP using the HDOCK server, it was found that the high elasticity of the N-terminal domain appears to be the main reason for the failure in spatial fitting of the proteins, and the server automatically cut off 28 N-terminal amino-acid residues (this effect was not observed for the UL49.5 protein from BoHV-1).

In summary, our results show that there are several factors contributing to the failure of UL49.5 from HSV-1 to bind to TAP. First, the long and rigid transmembrane helix of the TMD from HSV-1, unlike BoHV-1, fits poorly to the transmembrane part of the TAP protein. In the case of the UL49.5 protein from BoHV-1, the transmembrane fragment is longer, which forces the protein to adopt a structure of two connected shorter helices to fit into the ER membrane, and such an arrangement of helices likely allows better fitting to TAP. Second, the highly mobile N-terminal part of the HSV-1 protein, enriched with Gly and Pro residues, makes it highly elastic. This elasticity, combined with the lack of anchoring motifs such as the PPQ motif present in UL49.5 from BoHV-1, hinders the N-terminal part of the protein from HSV-1 from embedding in the membrane in the same way as the protein from BoHV-1. These structural differences, particularly the elasticity of the N-terminal region and the lack of anchoring motifs in HSV-1 with UL49.5, may contribute to its inability to inhibit the TAP complex, unlike the protein from BoHV-1.

H9, H10, and H11 Stability and Influence of Disulfide Bonds on Selected Proteins Studied Using Static and Dynamic Models

Despite the fact that disulfide bonds are present in more than 20% of proteins in the Protein Data Bank (PDB), their role is often overlooked during experimental and theoretical studies. This stems partly from the common, but not always correct, view that disulfide bonds primarily serve as stabilizers of protein structure. However, it is known that the reduction of a single disulfide bond does not always decrease the stability of the protein, and the addition of a new bond does not always increase it. Disulfide bonds can also play other roles, such as maintaining rigidity, preventing enzymatic proteolysis, or regulating protein function. Unfortunately, classical force fields do not allow for chemical reactions, such as the formation and breaking of disulfide bonds between cysteine side chains, during molecular dynamics simulations. The only possibility is to predefine which cysteines should be bonded by a disulfide bridge.

Since the formation of both native and non-native disulfide bonds is a natural step in the folding of parts of proteins, such as ribonuclease A, I improved the method for dynamically forming and breaking disulfide bonds in the coarse-grained UNRES force field (H9) with an additional

energy barrier to prevent the formation of disulfide bonds between three cysteine resides at once, which was possible in UNRES due to the simplified coarse-grained representation used (H10). The developed potential used the distance and angle between cysteine side chains to determine the possibility of forming and breaking disulfide bonds, allowing for the regulation of the potential well depth to match the conditions of the redox environment.

In the next step, I performed test molecular dynamics simulations for selected proteins to examine their correct folding during simulations, i.e., obtaining the native structure of the protein starting from an extended conformation, in the UNRES force field without disulfide bonds and with static and dynamic treatment of these bonds (H9). The obtained results indicate that simply including disulfide bonds in unfolded proteins does not significantly improve the quality of structure prediction because proteins often get trapped in non-native conformations, energetically blocked by the presence of disulfide bonds, resulting in, for example, entangled conformations. Using a dynamic description of disulfide bonds also does not significantly improve the ability to predict the native structure of proteins due to the longer computational time required for folding, resulting from the formation of non-native disulfide bonds. The conducted studies showed that studying the folding process of proteins is very challenging, and the presence of disulfide bonds further significantly increases this difficulty.

For this reason, I used the developed extension of the dynamic treatment of disulfide bonds to perform a series of molecular dynamics simulations at different temperatures, observing the dynamics and stability of ribonuclease A's disulfide bonds. The studies showed that thermal unfolding of the protein most frequently occurs through the breaking of the disulfide bond between cysteine residues Cys40-Cys95 (Figure 15), which connect two poorly structured parts of the protein. Additionally, fragments of ribonuclease A around residues 40 and 95 exhibit mainly repulsive interactions, leading to significant destabilization of the disulfide bond under reducing conditions. A similar mechanism of reduced stability occurs for the disulfide bond between cysteine residues Cys26-Cys84, although the reduction in stability is smaller because both these fragments are highly structured and do not exhibit large fluctuations in their secondary structures or changes in their orientation.

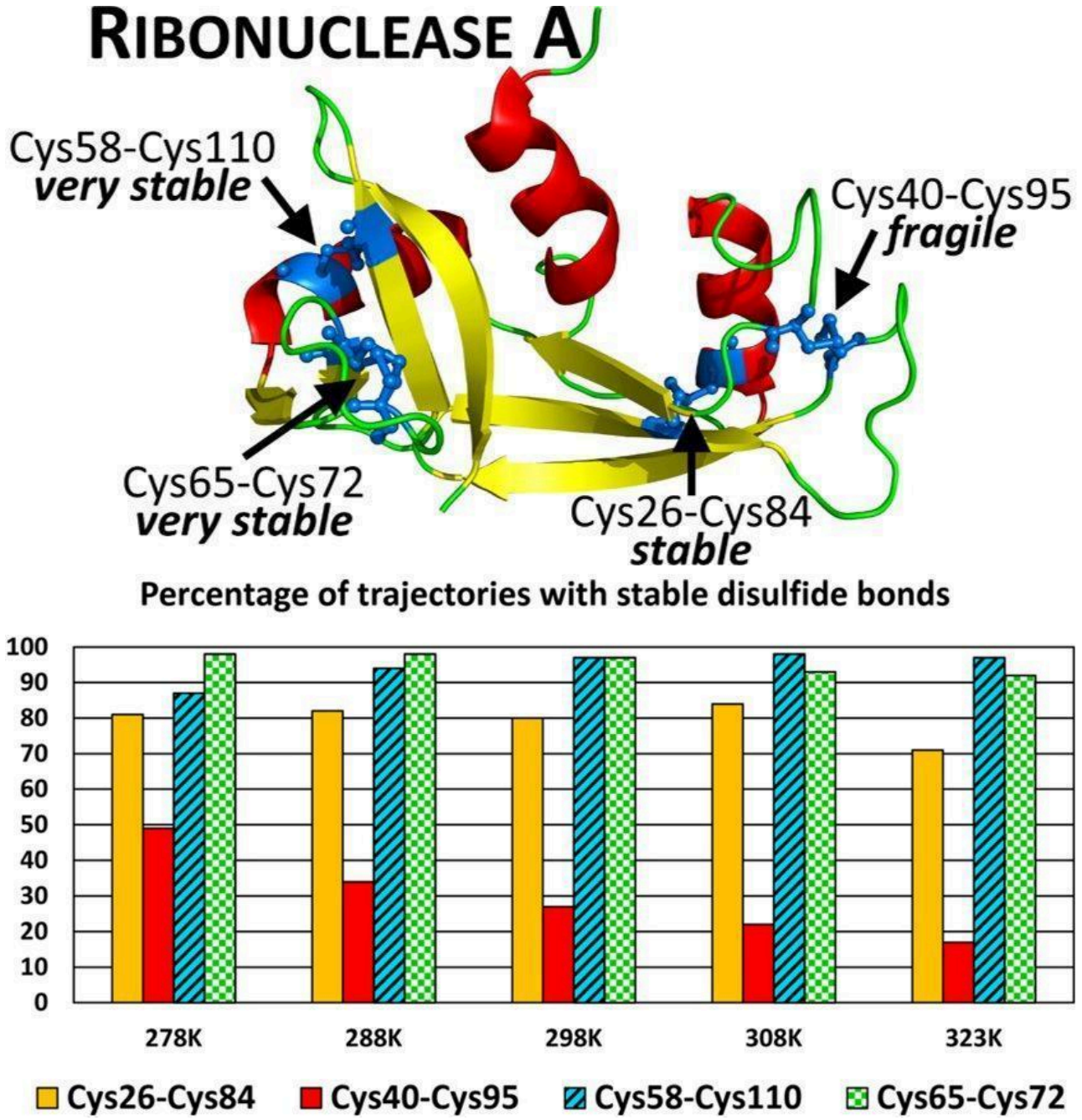


Figure 15. Ribbon representation of ribonuclease A (upper part of the figure) and plot of the percentage of trajectories in which the given disulfide bond was stable (lower part of the figure). Figure from work H10. Reprinted with permission from J. Chem. Theory Comput. 2017, 13, 11, 5721–5730. Copyright 2024 American Chemical Society.

To obtain a complete picture of the stability and function of disulfide bonds in ribonuclease A, we performed coarse-grained and all-atom steered molecular dynamics (SMD) simulations, corresponding to atomic force microscopy (AFM) experiments, by increase of the distance between the N- and C-terminal ends of the protein over time. The first step in this direction was to introduce a potential into the all-atom Amber force field that allows for the breaking and forming of disulfide bonds during simulations, similar to that implemented in the coarse-grained UNRES force field. Due to the use of an all-atom representation, this potential did not need to differentiate based on the valence angle of adjacent cysteines and relied solely on the distance between the sulfur atoms in the thiol groups of the side chains.

The conducted coarse-grained and all-atom stretching simulations of ribonuclease A showed good agreement with each other and demonstrated that the stability of the protein significantly depends on the redox conditions in which it is found. Furthermore, we identified that the most internal disulfide bond, Cys65-Cys72, which connects the most local fragments of the protein, is the most stable and is the last to break during mechanical stretching of the protein. This observation was subsequently confirmed by independent experimental studies performed by another research group.[16]

The obtained force profiles during protein stretching clearly showed that each disulfide bond breaking event during stretching is accompanied by an increase (a peak) in force and significant local and global conformational changes. The conducted studies provided detailed insights into the crucial stabilizing role of disulfide bonds and their dynamics during the unfolding of ribonuclease A, showing good agreement between the methods used and experimental data. The developed models for dynamically treating disulfide bonds during molecular dynamics simulations enable the study of various processes involving proteins, including those difficult to capture experimentally.

H12. Development of the UNRES-Dock Method for Predicting Protein-Protein and Protein-Peptide Complex Structures

Another desired functionality missing in the UNRES package was the ability to predict protein-protein and protein-peptide complex structures. For this reason, I developed the UNRES-Dock protocol for predicting such structures, based on coarse-grained molecular dynamics simulations with multiple replica exchange (MREMD) (H12). The obtained method used the structures of peptides and proteins or optionally their sequences to generate a "cloud" of possibly diverse initial positions of the systems (Figure 16). Subsequently, molecular dynamics simulations with replica exchange were performed to explore the conformational space of the system (72 replicas at 36 temperatures in the conducted simulations), and their converged parts were subjected to weighted histogram analysis and cluster analysis to generate representative structures, which were then converted into all-atom models.

During the development of the docking protocol, I discovered that each element of the protocol was crucial for the quality of the obtained structures and the efficiency of the entire method. For example, generating a cloud of initial positions, in contrast to classical docking, did not ideally fit the surfaces and interactions of the proteins but generated sets of diverse structures to accelerate and improve the exploration of the conformational space of the system during molecular dynamics simulations with replica exchange. The tests conducted by me showed that this approach was significantly more efficient than using a single initial orientation of the proteins or models of protein complexes with a large interaction interface. The latter approach was particularly problematic, often resulting in low convergence of simulations due to the system remaining in one orientation that was a local energy minimum.

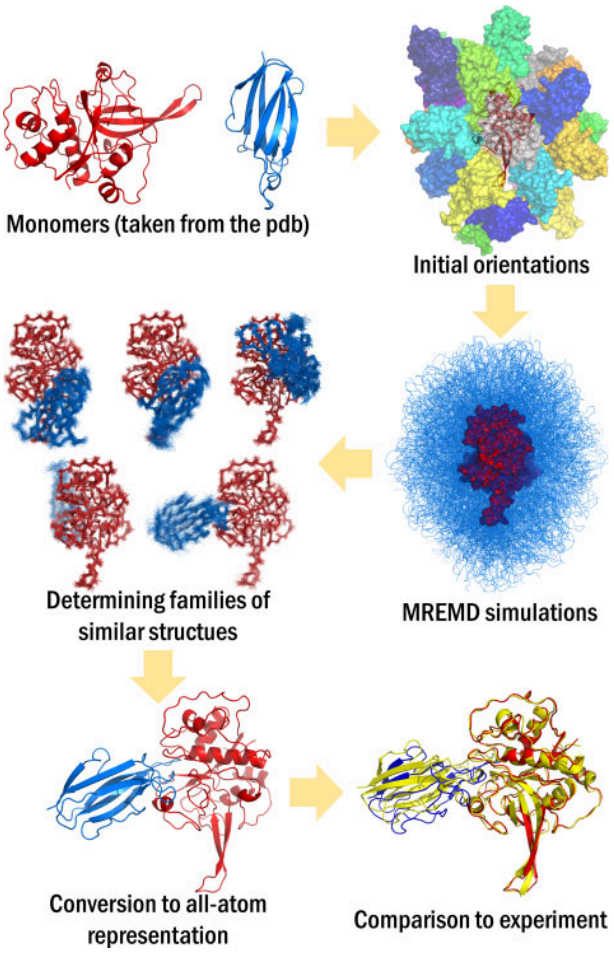


Figure 16. Schematic visualization of the protein-protein docking protocol in the UNRES-Dock method. Figure from work H12. Reprinted with permission from Bioinformatics, Volume 37, Issue 11, June 2021, Pages 1613–1615. Copyright 2024 Oxford University Press, license 5744190046736).

The developed UNRES-Dock protocol allows for adjusting the level of conformational flexibility, enabling semi-rigid and fully flexible docking of molecules. The main element of the conformational space exploration is molecular dynamics simulations with replica exchange, where constraints stabilizing the structures of individual components are applied without affecting their relative orientation. By using these constraints on the distances between interaction centers within a single protein, it is possible to regulate the level of their flexibility during docking simulations. Importantly, the developed docking protocol uses the UNRES force field with periodic boundary conditions and proved to be significantly more efficient than the previously used "tether" or “leash” methods, which did not allow proteins to drift into infinite space while being burdened with the directional effect of keeping proteins close to each other using tether constraints.

The application of molecular dynamics with replica exchange proved crucial for achieving the appropriate degree of conformational space exploration and converging the simulation results due to the ability to overcome energy barriers in high-temperature trajectories and allowing the dissociation of less stable complex forms. The conducted test simulations showed that the proper selection of replica temperatures and the strength of constraints applied to individual proteins was key to achieving convergence and preventing proteins from unfolding at high temperatures. This is important because most proteins during binding do not undergo drastic conformational changes but only minor adjustments, mainly within the interaction interface, and using restraints of appropriate strength allows only limited conformational changes, significantly improving the computational efficiency of the method. What is important, for most systems, the UNRES force field used in the

method showed a good correlation between the free energy of the system and the similarity to the native structure of the complex, presented in the form of the root mean square deviation (RMSD, Figure 17), demonstrating the possibility of routinely obtaining good results, provided that the conformational space has been sufficiently effectively explored.

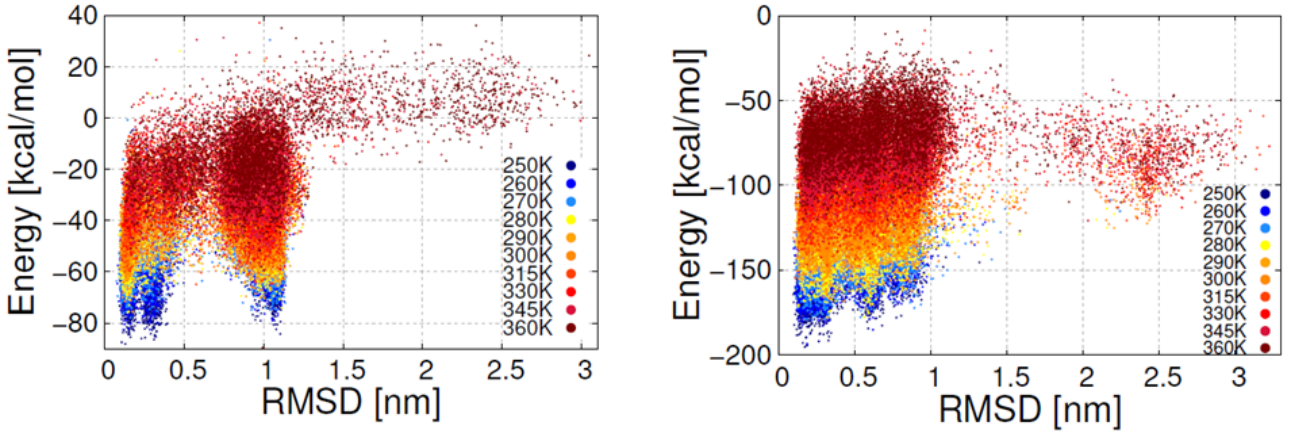


Figure 17. Plot of the free energy of the complex versus RMSD values for replicas at different temperatures for a selected protein-protein complex (PDB code: 2WFU; left panel) and protein-peptide complex (PDB code: 3BFQ, right panel). Figure from work H12. Reprinted with permission from Bioinformatics, Volume 37, Issue 11, June 2021, Pages 1613–1615. Copyright 2024 Oxford University Press, license 5744190046736).

In summary, the developed UNRES-Dock protocol enables the prediction of protein-protein and protein-peptide complex structures with the possibility of adjusting the level of structural flexibility, allowing for the study of semi-rigid and fully flexible binding mechanisms. For protein-peptide systems, the achieved efficiency significantly surpasses that of similar methods, such as CABS-Dock. The docking accuracy for protein-protein complexes reaches a satisfactory level, both for typical proteins commonly used as training and test proteins in molecular docking method development and for non-standard ones, such as homodimeric proteins, which are rarely included in protein complex databases. This behavior is different from other methods used for docking, such as Attract, which strongly rely on statistical data and achieve very high efficiency for typical systems but significantly lose effectiveness for non-standard systems or those not included in the pool of test proteins used during method optimization (Table 2).

Table 2. Number of protein-protein and protein-peptide complexes predicted with a given accuracy according to the CAPRI score scale for the UNRES-Dock method when generating 10 structures (Top10), 50 structures (Top50), and for the Attract and CABS-Dock methods when generating the default number of 10 structures. Table adapted from work H12. Reprinted with permission from Bioinformatics, Volume 37, Issue 11, June 2021, Pages 1613–1615. Copyright 2024 Oxford University Press, license 5744190046736).

Protein-protein			
Quality	Top10	Top50	Top10 Attract
High	0	0	12
Medium	8	8	1
Acceptable	3	7	1
Incorrect	9	5	6

Protein-peptide			
Quality	Top10	Top50	Top10 CABS-Dock
High	3	4	0
Medium	11	19	6
Acceptable	18	11	22
Incorrect	3	1	7

H13. Modeling the CD28-CD86 Protein Complex Structure

The final test of the UNRES-Dock method was to predict the structure of the CD28-CD86 protein complex, which had not been experimentally resolved at that time (H13). To achieve this, I performed the full protocol for predicting protein complex structures, which was developed and described in work H12, and then conducted all-atom molecular dynamics simulations to verify the stability of the 10 predicted representative models based on cluster analysis of conformational ensembles generated by the weighted histogram analysis method (WHAM) on coarse-grained molecular dynamics trajectories with replica exchange (H13). The all-atom simulations showed that out of the 10 obtained models of the CD28-CD86 complex, only model number 7 was fully stable during 500 ns simulations. To confirm this result, I performed five additional all-atom trajectories for this arrangement of proteins, where the crystal structures of CD28 (PDB code: 1YJD) and CD86 (PDB code: 1I85) were placed in orientations corresponding to model number 7. All these simulations showed high stability of the complex, with RMSD values of the order of 0.4 nm (Figure S3 in work H13). The studies also indicated that even for the stability of a heterodimer with a large interaction interface, the accuracy of the local elements of the structures of its components, undergoing minor deformations during the reconstruction from coarse-grained to all-atom models, is crucial.

The obtained structure of the CD28-CD86 complex using the UNRES-Dock protocol was not only confirmed by all-atom simulations but also showed similarity to the analogous CTLA-4-CD86 complex and consistency with available experimental data regarding potential amino-acid residues involved in the interaction interface. Analysis of the interface in the predicted model indicates the key role of residues Glu32, Arg34, Tyr51, and Met99, Tyr100, and Pro102 from the MYPPPY region on CD28, as well as Phe31, Val39, Glu42, Tyr44, Thr93, and Ile96 on CD86. The designed model not only confirms previous experimental observations but also presents a complete interaction interface (Figure 18) along with the importance of individual amino-acid residues for maintaining its stability. Additionally, simulations of the association kinetics of CD28-CD86 and CTLA-4-CD86 complexes showed about a 4-fold lower stability of the former, which is consistent with experimental findings.

In summary, the developed method, combining simulations at different levels of resolution, allows for reliable modeling of the structure and dynamics of protein complexes. The obtained results not only verify the accuracy of the UNRES-Dock method but also provide valuable information on the structural and thermodynamic basis of immune system protein interactions, which are important for the search for new therapies.

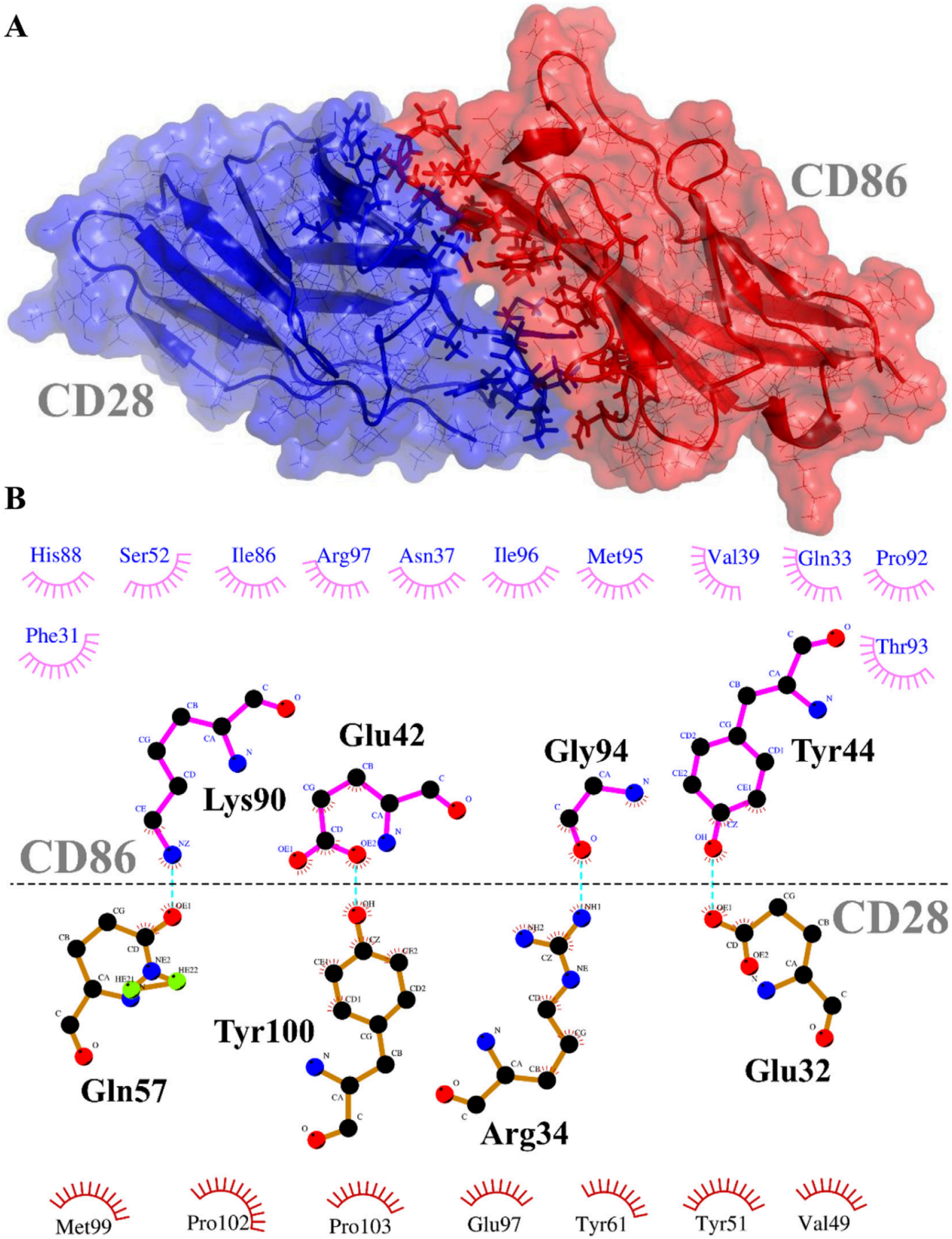


Figure 18. Visualization of the obtained stable conformation of the CD86-CD28 complex with highlighted amino-acid residues involved in the interaction between proteins. Figure from work H13. Reprinted with permission from Journal of Molecular Graphics and Modelling Volume 103, March 2021, 107802. Copyright 2024 Elsevier.

5.3.4 Summary of the Significance of the Publications

The results presented in the cycle of thematically related publications significantly enhanced the capabilities of simulating the structure and dynamics of proteins and their complexes. The developed protocols and methods enabled the performance and analysis of routine, repetitive, and physics-based simulations to predict protein structures, and the optimization performed significantly improved the efficiency of the coarse-grained UNRES force field. Some of these functions were implemented in the UNRES and UNRES-Dock servers, allowing for intuitive execution of simple simulations and generating files for long simulations on supercomputers. The developed methods for predicting protein structures and their complexes were tested and appreciated in the international CASP and CAPRI experiments (P11) and were used to obtain structures of many proteins and their complexes. Additionally, the introduction and optimization of the potential enabling dynamic formation and breaking of disulfide bonds during molecular dynamics simulations expanded the capabilities of modeling systems containing this type of post-translational modification, which accounts for over 20% of all proteins.

My research also contributed to a better understanding of the theoretical foundations underlying the fundamental properties of proteins. The conducted simulations not only determined the efficiency of various methods for predicting the structures and physicochemical properties of peptides and proteins but also showed that, especially for disordered proteins, the dominant part of the free energy comes from interactions between the protein and the solvent (water), rather than interactions within the protein itself. I provided new insights into the interactions between α-synuclein and Aβ42, two key molecules with inherently disordered structures involved in the pathogenesis of neurodegenerative diseases. My results shed new light on the early stages of the aggregation process of these proteins and may contribute to a better understanding of the molecular basis of neurodegenerative diseases and facilitate the design of new therapeutic strategies.

Furthermore, I demonstrated that for the studied disordered peptides and small oligomers, electrostatic interactions dominate over van der Waals interactions, emphasizing the importance of solvent interactions as the primary factor determining their structure and dynamics. Therefore, it is crucial to use the best models not only for the proteins themselves but also for water, ions, and lipids and to accurately reflect physiological conditions. These latter points were particularly important during the study of the dynamics and structure of viral proteins UL49.5, whose behavior in a monocomponent lipid bilayer model showed good agreement with experimental data obtained using the same model but did not explain all physiological effects. I also showed that not always more advanced sampling methods are necessary to achieve high precision in calculations, and sometimes they can even lead to a slight deterioration in the quality of structure prediction.

The analysis of reducing and oxidizing conditions revealed that disulfide bonds significantly differ in stability and function, and their impact on protein structure stability greatly depends on the local character of the fragments in which they occur. Additionally, I demonstrated that the stability and sequence of breaking disulfide bonds significantly depend on external factors such as the reducing-oxidizing environment, temperature, or mechanical stress.

I proved that combining studies at different levels of resolution not only serves as a verification when experimental data are not available but also allows for understanding the properties of molecules at different time scales. Thanks to such observations, I determined the mechanism and pathways of aggregation of four chains of amyloid β in the coarse-grained UNRES force field, showing that the dominant pathway is the formation of dimers from monomers and then tetramers from dimers, while the pathway of forming tetramers from trimers and monomers is much less favorable. The obtained tetrameric models were verified using all-atom simulations, and I identified a model consistent with all non-contradictory literature data. A similar approach enabled me to predict the structure of the CD28-CD86 protein complex, important in the immune response of the organism, using the coarse-grained UNRES-Dock protocol, whose results were verified using all-atom methods, allowing the narrowing down from ten coarse-grained models to one all-atom

model that shows high similarity to the analogous CTLA-4-CD86 complex. The developed models of proteins, their complexes, and aggregates can serve for further research, for example, to develop substances inhibiting the early stages of amyloid β aggregation.

5.4 Results outside the publication cycle published before obtaining a doctoral degree

The research I conducted before obtaining my doctoral degree focused on the utilization and improvement of the UNRES force field for predicting protein structures. To this end, I participated in protein structure prediction as part of the 10th edition of the CASP experiment, demonstrating the high potential of the method while also revealing its weaknesses, such as the arbitrariness of selecting representative structures and the inconsistent prediction performance dependent on the preferences of the force field user (D1). The approach used highlighted the high efficiency in predicting the relative arrangement of protein domains, achieving the best result for one of the studied systems among over 100 research groups participating in the experiment. To improve the predictive capabilities of the UNRES force field within poorly structured elements, I determined and introduced into the UNRES force field potentials based on statistical analysis of amino-acid residues constituting loop-type protein structures, using a simplified alphabet of amino-acid residue types (D2), and then analogously, but for all amino-acid residue types, for potentials determined by semi-empirical methods (D5). The introduction of these correlation potentials significantly improved the quality of predicting loop conformations and other partially unstructured fragments in the test protein set. The improvement in the prediction quality of poorly structured fragments also improved the arrangement of structured elements, such as α-helices and β-sheets, which was particularly evident in the case of the potential based on the physics of interactions (D5).

Simultaneously, as part of the WeFold initiative, I made an effort to improve the ability of the UNRES force field to predict protein structures by utilizing external data provided by collaborators from other research groups, specifying probable ranges of distances between amino-acid residues in the prediction protocol (D3). This collaboration resulted in a significant improvement in predicting the general structure type for several proteins, but it did not significantly improve local imperfections resulting from the use of a coarse-grained representation.

The improvements introduced by me and other creators of the UNRES force field served to identify further directions for the development of the unified coarse-grained model (UCGM, later renamed to UNICORN) not only for proteins (UNRES) but also for other biomolecules, such as nucleic acids (NARES), carbohydrates (SUGRES), and lipids (D4).

Subsequently, we used the improved UNRES force field to investigate the effect of arginine on its binding protein from Thermotoga maritima and the associated structural changes occurring in the protein. This project, in addition to proposing a mechanism for arginine binding by the protein from Thermotoga maritima, demonstrated that the improved UNRES force field, despite its coarse-grained nature, has sufficient accuracy to predict the impact of the interaction of a single amino-acid residue with a protein and the structural changes occurring under the influence of the interaction (D6). Similarly to this work, the UNRES force field was used to predict the structure and dynamics of the yeast Isu1-Jac1 protein complex, showing that while maintaining the experimental binding interface by three amino-acid residues from Jac1 (L105, L109, and Y163), these proteins have three distinct orientations forming stable complexes (D8), differing in additional amino-acid residues involved in protein binding. These studies showed that these proteins have a high degree of conformational freedom and dynamics of protein orientation, maintaining interactions between at least three amino-acid residues from the Jac1 protein in each of the structures.

The last work I performed before obtaining my doctoral degree focused on using multimodal soft restraints based on structure fragments predicted using other methods based on statistical sequence-structure similarity methods of proteins to improve the predictive capabilities of the

UNRES force field (D7). The conducted research concentrated on difficult systems in which the structures of individual domains were correctly predicted by leading bioinformatics methods, but their relative arrangement was not correctly determined due to its different type compared to proteins with known structures. The performed research showed that the UNRES force field, supported by data on local structures, is able to predict domain arrangements of many proteins better than any other method, especially if they have a large interaction interface, thanks to the use of potentials based on the physics of interactions rather than statistics.

In addition to regular scientific articles, during my doctoral studies, I published several conference papers based on preliminary observations made during the research for my doctorate (DM3-4). As part of a research internship at Jackson State University, I investigated the interaction ability and potential nanotoxicity of carbon nanotubes to Tol receptors (DM1) and the interactions of gold nanoparticles with proteins and DNA (DM2). Additionally, in a Polish-language book chapter (DM5), I synthetically described the basic concepts used in molecular modeling.

List of publications published before obtaining a doctoral degree:

D1. „Lessons from application of the UNRES force field to predictions of structures of CASP10 targets", Yi He, Magdalena A. Mozolewska, **Paweł Krupa**, Adam K. Sieradzan, Tomasz K. Wirecki, Adam Liwo, Khatuna Kachlishvili, Shalom Rackovsky, Dawid Jagieła, Rafał Ślusarz, Cezary R. Czaplewski, Stanisław Ołdziej, Harold A. Scheraga*; Proceedings of the National Academy of Sciences (USA), 2013, 110(37), 14936-14941.

D2. „Improvement of the treatment of loop structures in the UNRES force field by inclusion of coupling between backbone- and side-chain-local conformational states", **Paweł Krupa**, Adam K. Sieradzan,* S. Rackovsky, Maciej Baranowski, Stanisław Ołdziej, Harold A. Scheraga, Adam Liwo, Cezary Czaplewski; Journal of Chemical Theory and Computation, 2013, 9(10), 4620–4632.

D3. „WeFold: A Coopetition for Protein Structure Prediction", George A. Khoury, Adam Liwo, Firas Khatib, Hongyi Zhou, Gaurav Chopra, Jaume Bacardit, Leandro O. Bortot, Rodrigo A. Faccioli, Xin Deng, Yi He, **Pawel Krupa**, Jilong Li, Magdalena A. Mozolewska, Adam K. Sieradzan, James Smadbeck, Tomasz Wirecki, Seth Cooper, Jeff Flatten, Kefan Xu, David Baker, Jianlin Cheng, Alexandre C. B. Delbem, Christodoulos A. Floudas, Chen Keasar, Michael Levitt, Zoran Popović, Harold A. Scheraga, Jeffrey Skolnick, Silvia N. Crivelli* and Foldit Players; Proteins: Structure, Function, and Bioinformatics, 2014, 98(9):1850–1868.

D4. „A Unified Coarse-Grained Model of Biological Macromolecules Based on Mean-Field Multipole-Multipole Interactions", Adam Liwo*, Maciej Baranowski, Cezary Czaplewski, Ewa Gołaś, Yi He, Dawid Jagieła, **Paweł Krupa**, Maciej Maciejczyk, Mariusz Makowski, Magdalena A. Mozolewska, Andrei Niadzvedtski, Stanisław Ołdziej, Harold A. Scheraga, Adam K. Sieradzan, Rafał Ślusarz, Tomasz Wirecki, Yanping Yin, Bartłomiej Zaborowski; Journal of Molecular Modeling, 2014, 20(8):2306.

D5. "Physics-based potentials for the coupling between backbone- and side-chain-local conformational states in the united residue (UNRES) force field for protein simulations", Adam K. Sieradzan, **Paweł Krupa***, Harold A. Scheraga, Adam Liwo, and Cezary Czaplewski; Journal of Chemical Theory and Computation, 2015, 11(2): 817–831.

D6. "Studies of conformational changes of an arginine-binding protein from Thermotoga maritima in the presence and the absence of ligand with use of the molecular dynamics simulations with the coarse-grained UNRES force field.", Agnieszka G. Lipska, Adam K. Sieradzan, **Paweł Krupa**, Magdalena A. Mozolewska, Sabato D'Auria, Adam Liwo*, Journal of Molecular Modeling, 2015, 21(3):64.

D7. “Prediction of protein structure by template-based modeling combined with the UNRES force field”, **Paweł Krupa**, Magdalena A. Mozolewska, Keehyoung Joo, Jooyoung Lee, Cezary Czaplewski, Adam Liwo*, Journal of Chemical Information and Modeling 2015, 55(6):1271-1281.

D8. “Molecular modeling of the binding modes of the Iron-sulfur protein to the Jac1 co-chaperone from Saccharomyces cerevisiae by all-atom and coarse-grained approaches”, Magdalena A. Mozolewska, **Paweł Krupa**, Harold A. Scheraga, Adam Liwo*, Proteins: Structure, Function, and Bioinformatics 2015, 83(8):1414-1426.

Peer-reviewed English-language publications:

DM1. “Preliminary studies of interaction between nanotubes and toll-like receptors”, M.A. Mozolewska,* **P. Krupa**, B. Rasulev, A. Liwo, J. Leszczynski; TASK Quarterly, 2014, 18(4), 351-355.

DM2. “Towards mechanisms of nanotoxicity – interaction of gold nanoparticles with proteins and DNA”, **P. Krupa***, M.A. Mozolewska, B. Rasulev, C. Czaplewski, J. Leszczynski; TASK Quarterly, 2014, 18(4), 337-341.

Peer-reviewed Polish-language publications:

DM3. „Badanie procesu zwijania białek przy użyciu pola gruboziarnistego UNRES na podstawie peptydu 1E0L (domena WW) i jego mutantów”, Paweł Krupa, Magdalena Mozolewska, Gia Maisuradze, Adam Liwo, Harold A. Scheraga; Młodzi naukowcy dla polskiej nauki, pod redakcją dr inż. Marcin Kuczera, Część 11 – Nauki przyrodnicze, 213-219, CREATIVETIME, Kraków 2013.

DM4. „Zastosowanie pola gruboziarnistego UNRES do symulacji dynamiki molekularnej kompleksów białek na przykładzie drożdżowych białek opiekuńczych”, Magdalena Mozolewska, Paweł Krupa; Młodzi naukowcy dla polskiej nauki, pod redakcją dr inż. Marcin Kuczera, Część 11 – Nauki przyrodnicze, 139-143, CREATIVETIME, Kraków 2013.

Peer-reviewed Polish-language book chapter:

DM5. „Podstawy modelowania molekularnego”, Paweł Krupa – rozdział w książce Nowe trendy w naukach przyrodniczych 4 pod redakcją dr inż. Marcin Kuczera, Creative Science – Monografia 2013, tom 1, 154-163, CREATIVETIME, Kraków 2013.

5.5 Results outside the publication cycle published after obtaining a doctoral degree

The scientific research I conducted outside the cycle of thematically related publications focused on various issues relevant from the perspective of physics, biology, and medicine. Below, I present a synthetic description of these studies, divided into five categories based on the research topic.

5.5.1 Predicting the structure and dynamics of proteins and their complexes

Coarse-grained simulations using the UNRES force field allowed us to determine the folding mechanism of the WW protein and the impact of mutations of selected amino-acid residues on this process (P1). By using a coarse-grained model describing this protein, it was possible to perform many (over 1000) trajectories of sufficient length to reach the native conformation. Such a large

number of trajectories was necessary to understand not only the dominant folding pathway but also the less frequent pathways, the statistics of their occurrence, and the impact of mutations on this process. The performed simulations also enabled the determination of the phase transition temperatures of the protein and its mutants, demonstrating high accuracy in relation to experimental data.

One of the areas of my interest was the further development of the UNRES force field to increase its protein structure prediction capabilities. To this end, I developed a method for imposing restraints on protein structure fragments based on multiple available models (P2) using soft multimodal restraints on angles and distances, enabling the automatic selection of the best features of the predicted models during simulations in the coarse-grained UNRES force field. We improved and tested this method as part of the 12th edition of the CASP experiment (P7), achieving top performance on most of the predicted structures. The latest improvement of the method involves enhancing the ability to select common fragments among predicted protein models and utilizing statistical potentials for amino-acid fragments from the DFA method (P12). We applied a similar approach, but within the WeFold initiative, to improve predictions using distance restraints provided by collaborating research groups (P8), improving the quality of the predicted protein structures. Another approach to protein structure prediction involved using data from the SAXS experiment to fit the distributions of modeled structures to experimental observations (P6), significantly improving the distribution of folded and partially unfolded structures in simulations.

Simultaneously, I participated in a project aimed at improving the predictive capabilities of the UNRES force field without using external information, such as partial experimental data or data from bioinformatics analyses, but only using protein sequences (P10), after performing force field optimization (H2). In the 13th edition of the CASP-CAPRI experiment, we predicted protein-protein structures (P11) using a preliminary version of the UNRES-Dock method (H11), demonstrating the high efficiency of the method and identifying the strengths and weaknesses of the test versions of the docking protocol. In the 14th edition of the CASP experiment, we additionally used the improved UNRES force field to predict the structures of proteins and protein complexes of various sizes (even up to an entire viral capsid containing nearly 70,000 amino-acid residues) (P14). It should be noted that the above achievements would not have been possible without the work on the problem of ergodicity of molecular dynamics simulations with classical and Hamiltonian replica exchange in the UNRES force field using distance restraints (P5). These studies showed that, especially with low-quality initial structures, using a larger number of trajectories with different restraint strengths significantly improves the ergodicity of simulations and the quality of the obtained statistical ensembles. This effect becomes less significant when the initial structure is close to the native structure and the applied restraints are highly consistent with it, which, however, occurs mainly in the case of proteins with trivial topology. These observations enabled, among other things, much better planning of research for the optimization of the UNRES force field using the maximum likelihood method (H2).

Due to its unique nature, based on physical interactions rather than statistics, the energy function from the coarse-grained UNRES force field, after modifications allowing approximate determination of the energy of proteins with missing fragments and the use of machine learning methods to optimize the weights of individual contributions, proved to be a good tool for assessing the quality of predicted protein models (protein threading), being on par with or even surpassing available methods such as DFIRE (P9). It is worth noting that the observations obtained during the research within this project (P9) also facilitated the design of distance and angle restraints based on selected protein models (P7 and P12).

An interesting example of the use of disulfide bonds is the study of the aggregation process of amyloid β with two amino-acid residues substituted with cysteine residues (L17C/L34C, P22). The modified amyloid β is able to accelerate aggregation while maintaining the ability to aggregate with the wild-type form, simultaneously showing a greater tendency to form folded structures resembling protofibrils even in monomeric form. Interestingly, this change reduced, rather than increased,

which is usually associated with the rate of aggregation, the tendency of the monomeric form of amyloid β to adopt β-type secondary structures. Importantly, this result was obtained using three different force fields (all-atom Amber and coarse-grained Martini 3 and SIRAH) in two software packages (Amber and Gromacs), providing higher reliability of the observed phenomena.

Another example of using multiple methods in one study and the importance of considering the presence of disulfide bonds is the prediction of an efficient inhibitor of the HVEM-LIGHT complex (P23). These studies were based on all-atom molecular dynamics simulations in various variants to find the shortest possible peptide based on the HVEM protein that strongly interacts with the LIGHT protein. By combining classical molecular dynamics simulations with free energy analysis using the MM-GBSA method with entropy estimation, the work of transition between the free and bound form, steered molecular dynamics simulations to dissociate the designed peptides from the LIGHT protein, and the use of the umbrella sampling method to determine the potentials of mean force, we determined the dynamic and thermodynamic properties and the mechanism of binding of these peptides to the LIGHT protein. The obtained results unequivocally showed that the peptide, which is a variant of the HVEM protein with a shortened sequence and lacking one of the disulfide bonds present in the physiological form of the protein, allows the best fit to the LIGHT protein and the greatest interaction strength.

5.5.2 Determining the molecular basis of potential nanotoxicity of molecules

Thurincin H is one of the few exceptional peptides containing covalent bonds analogous to disulfide bonds, which, however, do not connect the side chains of cysteines but the main chain of one amino-acid residue with the side chain of cysteine (P3). The presence of such unusual bridges is one of the reasons for the strong antibacterial properties of this molecule. Our coarse-grained molecular dynamics simulations for all possible variants of the presence and absence of these bonds showed that the presence of all thioether bridges destabilizes the structure of this peptide and serves a function of ensuring enzymatic, rather than structural, stability, directly influencing the bacteriocin character of the peptide.

In the next project, we investigated the possibility of the occurrence of nanotoxicity induced by C60 fullerene derivatives in the vicinity of proteins. To this end, we conducted extensive studies of proteins possessing regions, so-called binding pockets, capable of binding various fullerene derivatives. Analysis of the performed screening dockings indicated potential systems that, in the subsequent molecular dynamics simulations, showed high stability of fullerene derivatives in the binding pockets (P4). These studies proved that at least some fullerene derivatives can be efficient protein inhibitors and cause a nanotoxicity effect for organisms. I also participated in the introduction of a continuous and molecular model of C60 fullerene into the coarse-grained UNRES force field, not only expanding the capabilities of the method but also demonstrating the inhibitory properties of fullerenes on selected proteins (P20). The above works showed that carbon nanoparticles can significantly affect the structure and dynamics of proteins, causing local toxicity effects.

5.5.3 Studies focusing on the role and impact of lipid membranes on proteins and the influence of various molecules on lipid bilayers

Performed all-atom molecular dynamics simulations showed that both copper ions and amyloid β peptides can bind to and destabilize lipid membranes, but their combination negates this effect (P13). According to new literature reports, amyloid β oligomers are more toxic than fibrils. Therefore, we decided to test how forms containing twelve chains (dodecamers) and short fibrils (protofibrils) interact with lipid membranes resembling the lipid composition of neuron membranes. To this end, I built possible models of amyloid β dodecamers using the most stable tetrameric structures obtained in previous studies (H3 and H7). Our molecular dynamics simulations

unequivocally proved that dodecamers have a much stronger effect on membranes, causing their rupture, than protofibrils (P15), for which this effect was absent.

In turn, molecular dynamics simulations performed at different resolution levels and experimental measurements for the UL49.5 protein from the BoHV-1 herpesvirus showed that the RRR motif is crucial for stabilizing the protein arrangement in the lipid membrane, and its mutation to alanines results in a lack of activity (P16). These studies provided a basis for further research using a heterogeneous lipid membrane model, with a lipid composition similar to the endoplasmic reticulum membrane, and comparing the conformational dynamics of active and inactive protein variants from different herpesviruses (H7).

One of the most interesting studies I conducted were molecular dynamics simulations explaining the experimental observations of why serotonin has different effects on different homogeneous lipid membranes (P17). Classical molecular dynamics simulations showed that serotonin molecules are less organized in terms of both penetration depth and angle in uncharged (POPC) than charged (POPS) membranes, which resulted from weaker interactions of hydrophilic parts with the heads of uncharged lipids. In turn, steered molecular dynamics simulations explained the mechanism of increased resistance to rupture of charged membranes compared to uncharged ones. This effect is caused by the formation of a layer of three oppositely charged molecules in the case of charged lipid heads: lipid heads, serotonin, and ions present in the solution. In the case of uncharged membranes, this effect was not observed, and the slight decrease in membrane resistance in the presence of serotonin resulted from the weakening of interactions between lipid tails. In the continuation of research on the impact of serotonin on cells, we showed how the presence of serotonin affects membrane hydrophilicity and increased ability of water to penetrate its interior, resulting in exocytosis of old cells containing higher serotonin concentrations (P21).

5.5.4 Establishing the molecular basis of mechanical stability of nucleic acids with specific sequences

A series of steered molecular dynamics simulations in the coarse-grained NARES force field for nucleic acids with different sequences showed that only telomeric sequences exhibit stronger interactions with each other within the double-stranded DNA and maintain stability during mechanical tearing by stepwise restoration of connections between repeating nucleotide sequences (P18). Similar simulations performed for nucleic acids with single-strand breaks (SSBs) showed a strong dependence of stability on chain length and sequence, and high agreement with simple models describing these dependencies (P19).

5.5.5 Other scientific works

In several book chapters (R1-R7) and one conference publication (M1), I described the current capabilities of the UNRES force field, including new functions enabling the study of breaking and forming disulfide bonds, simulating carbon nanoparticles, or performing molecular docking, as well as a description of other computational methods for simulating proteins, their structure, aggregation ability, and nucleic acids.

List of publications published after obtaining a doctoral degree, not included in the cycle, divided into publications in journals indexed in the Web of Science Core Collection (PX), chapters in books with international reach (RX), and other works (MX):

P1. “Preventing fibril formation of a protein by selective mutation”, Gia G. Maisuradze, Jordi Medina, Khatuna Kachlishvili, **Paweł Krupa**, Magdalena A. Mozolewska, Pau Martin-Malpartida, Luka Maisuradze, Maria J. Macias, and Harold A. Scheraga*, Proceedings of the National Academy of Sciences (USA), 2015, 112(44):13549-13554.

P2. “Use of Restraints from Consensus Fragments of Multiple Server Models to Enhance Protein-Structure Prediction Capability of the UNRES Force Field”, M.A. Mozolewska, P. Krupa, B. Zaborowski, A. Liwo*, J. Lee, K. Lee, C. Czaplewski, Journal of Chemical Information and Modeling 2016, 56(11), 2263–2279.

P3. “Role of the sulfur to α-carbon thioether bridges in thurincin H”, M.A. Mozolewska, A.K. Sieradzan*, A. Niadzvedstki, C. Czaplewski, A. Liwo, **P. Krupa**, Journal of Biomolecular Structure & Dynamics 2017, 35 (13), 2868-2879.

P4. “Inhibitors or Toxins? Large Library Target-Specific Screening of Fullerene-based Nanoparticles for Drug Design Purpose”, Lucky Ahmed, Bakhtiyor Rasulev, Supratik Kar, **Paweł Krupa**, Magdalena A. Mozolewska, Jerzy Leszczynski*, Nanoscale 2017, 9 (29), 10263-10276.

P5. “Ergodicity and model quality in template-restrained canonical and temperature/Hamiltonian replica exchange coarse-grained molecular dynamics simulations of proteins”, A.S. Karczyńska, C. Czaplewski, **P. Krupa**, M.A. Mozolewska, K. Joo, J. Lee, A. Liwo*, Journal of Computational Chemistry 2017, 38 (31), 2730-2746.

P6. “Prediction of protein structure with the coarse-grained UNRES force field assisted by small X-ray scattering data and knowledge-based information”, Agnieszka S. Karczyńska, Magdalena A. Mozolewska, **Paweł Krupa**, Artur Giełdoń, Adam Liwo, Cezary Czaplewski, Proteins: Structure, Function, and Bioinformatics 2018, 86 (S1), 228–239.

P7. “Use of the UNRES force field in template-assisted prediction of proteins structures and the refinement of server models: test with CASP12 targets”, Agnieszka Karczyńska, Magdalena A. Mozolewska, **Paweł Krupa**, Artur Giełdoń, Krzysztof K. Bojarski, Bartłomiej Zaborowski, Adam Liwo, Rafał Ślusarz, Magdalena Ślusarz, Jooyoung Lee, Keehyoung Joo, and Cezary Czaplewski*, Journal of Molecular Graphics and Modelling 2018, 83, 92-99.

P8. “Formative assessment of WeFold: a framework of international collaborative pipelines for protein structure prediction”, Chen Keasar, Liam J. McGuffin, Björn Wallner, Gaurav Chopra, Badri Adhikari, Debswapna Bhattacharya, Lauren Blake, Leandro Oliveira Bortot, Renzhi Cao, B.K. Dhanasekaran, Itzhel Dimas, Rodrigo Antonio Faccioli, Eshel Faraggi, Robert Ganzynkowicz, Sambit Ghosh, Soma Ghosh, Artur Giełdoń, Lukasz Golon, Yi He, Lim Heo, Jie Hou, Main Khan, Firas Khatib, George A. Khoury, Chris Kieslich, David E. Kim, **Pawel Krupa**, Gyu Rie Lee, Hongbo Li, Jilong Li, Agnieszka Lipska, Adam Liwo, Ali Hassan A. Maghrabi, Milot Mirdita, Shokoufeh Mirzaei, Magdalena A. Mozolewska, Melis Onel, Sergey Ovchinnikov, Anand Shah, Utkarsh Shah, Tomer Sidi, Adam K. Sieradzan, Magdalena Ślusarz, Rafal Ślusarz, James Smadbeck, Phanourios Tamamis, Nicholas Trieber, Tomasz Wirecki, Yanping Yin, Yang Zhang, Jaume Bacardit, Maciej Baranowski, Nicholas Chapman, Seth Cooper, Alexandre Defelicibus, Jeff Flatten, Brian Koepnick, Zoran Popović, Bartlomiej Zaborowski, David Baker, Jianlin Cheng, Cezary Czaplewski, Alexandre Cláudio Botazzo Delbem, Christodoulos Floudas, Andrzej Kloczkowski, Stanislaw Ołdziej, Michael Levitt, Harold Scheraga, Chaok Seok, Johannes Söding, Saraswathi Vishveshwara, Dong Xu, Foldit Players, and Silvia N. Crivelli, Scientific Reports 2018, 8, Article number: 9939.

P9. “Re-optimized UNRES Potential for Protein Model Quality Assessment”, Eshel Faraggi, **Pawel Krupa**, Magdalena A. Mozolewska, Adam Liwo, Andrzej Kloczkowski*, Genes 2018, 9, 601.

P10. “Evaluation of the scale-consistent UNRES force field in template-free prediction of protein structures in the CASP13 experiment”, Jozef Liwo, Adam Sieradzan, Agnieszka Karczyńska, Agnieszka Lipska, Artur Giełdoń, Celina Sikorska, Cezary Czaplewski, Emilia Lubecka, Karolina Zięba, Magdalena Mozolewska, Magdalena Ślusarz, **Pawel Krupa**, Rafal Slusarz, Sergey Samsonov, Silvia N. Crivelli, Urszula Uciechowska, Łukasz Golon, Journal of Molecular Graphics and Modelling 2019, 92, 154-166.

P11. „Blind prediction of homo- and hetero-protein complexes: The CASP13-CAPRI experiment", Marc F. Lensink*, Guillaume Brysbaert, Nurul Nadzirin, Sameer Velankar, Raphaël A.G. Chaleil, Tereza Gerguri, Paul A. Bates, Elodie Laine, Alessandra Carbone, Sergei Grudinin, Ren Kong, Ran-Ran Liu, Xi-Ming Xu, Hang Shi, Shan Chang, Miriam Eisenstein, Agnieszka Karczynska, Cezary Czaplewski, Emilia Lubecka, Agnieszka Lipska, **Paweł Krupa**, Magdalena Mozolewska, Łukasz Golon, Sergey Samsonov, Adam Liwo, Silvia Crivelli, Guillaume Pagès, Mikhail Karasikov, Maria Kadukova, Yumeng Yan, Sheng-You Huang, Mireia Rosell, Luis Angel Rodríguez-Lumbreras, Miguel Romero-Durana, Lucía Díaz-Bueno, Juan Fernandez-Recio, Charles Christoffer, Genki Terashi, Woong-Hee Shin, Tunde Aderinwale, Sai Raghavendra Maddhuri Venkata Subraman, Daisuke Kihara, Dima Kozakov, Sandor Vajda, Kathyn Porter, Dzmitry Padhorny, Israel Desta, Dmitri Beglov, Mikhail Ignatov, Sergey Kotelnikov, Iain H. Moal, David W. Ritchie, Isaure Chauvot de Beauchêne, Bernard Maigret, Marie-Dominique Devignes, Maria Elisa Ruiz Echartea, Didier Barradas-Bautista, Zhen Cao, Luigi Cavallo, Romina Oliva, Yue Cao, Yang Shen, Minkyung Baek, Taeyong Park, Hyeonuk Woo, Chaok Seok, Merav Braitbard, Lirane Bitton, Dina Scheidman-Duhovny, Justas Dapkūnas, Kliment Olechnovič, Česlovas Venclovas, Petras J. Kundrotas, Saveliy Belkin, Devlina Chakravarty, Varsha D. Badal, Ilya A. Vakser, Thom Vreven, Sweta Vangaveti, Tyler Borrman, Zhiping Weng, Johnathan D. Guest, Ragul Gowthaman, Brian G. Pierce, Xianjin Xu, Rui Duan, Liming Qiu, Jie Hou, Benjamin Ryan Merideth, Zhiwei Ma, Jianlin Cheng, Xiaoqin Zou, Panos I. Koukos, Jorge Roel-Touris, Francesco Ambrosetti, Cunliang Geng, Jörg Schaarschmidt, Mikael E. Trellet, Adrien S.J. Melquiond, Li Xue, Brian Jiménez-García, Charlotte W. van Noort, Rodrigo V. Honorato, Alexandre M.J.J. Bonvin, Shoshana J. Wodak, Proteins: Structure, Function, and Bioinformatics 2019, 87, 1200-1221.

P12. "Improved consensus-fragment selection in template-assisted prediction of protein structures with the UNRES force field in CASP13", Agnieszka S. Karczyńska, Karolina Zięba, Urszula Uciechowska, Magdalena A. Mozolewska, **Paweł Krupa**, Emilia A. Lubecka, Agnieszka G. Lipska, Celina Sikorska, Sergey A. Samsonov, Adam K. Sieradzan, Artur Giełdoń, Adam Liwo, Rafał Ślusarz, Magdalena Ślusarz, Jooyoung Lee, Keehyoung Joo, and Cezary Czaplewski*, Journal of Chemical Information and Modeling 2020, 60, 3, 1844-1864.

P13. "A Computational Model to Unravel the Function of Amyloid-β Peptide in Contact with a Phospholipid Membrane", Pham Dinh Quoc Huy, **Pawel Krupa**, Hoang Linh Nguyen, Giovanni La Penna, and Mai Suan Li*, The Journal of Physical Chemistry B 2020, 124 (16), 3300-3314.

P14. "Modeling protein structures with the coarse-grained UNRES force field in the CASP14 experiment", Anna Antoniak, Iga Biskupek, Krzysztof K. Bojarski, Cezary Czaplewski, Artur Giełdoń, Mateusz Kogut, Małgorzata M. Kogut, **Paweł Krupa**, Agnieszka G. Lipska, Adam Liwo, Emilia A. Lubecka, Mateusz Marcisza, Martyna Maszota-Zieleniak, Sergey A. Samsonov, Adam K.Sieradzan, Magdalena J. Ślusarza, Rafał Ślusarz, Patryk A. Wesołowski, Karolina Zięba, Journal of Molecular Graphics and Modelling 2021, 108, 108008.

P15. "Amyloid beta dodecamer disrupts the neuronal membrane more strongly than mature fibril: Understanding the role of oligomers in neurotoxicity", Hoang Linh Nguyen, Huynh Quang Linh, **Pawel Krupa**, Giovanni La Penna, Mai Suan Li*, The Journal of Physical Chemistry B 2022, 126, 20, 3659–3672.

P16. "The N-terminal proline hinge motif controls the structure of bovine herpesvirus 1-encoded inhibitor of the transporter associated with antigen processing required for its immunomodulatory function", Małgorzata Graul, Natalia Karska, Magda Wąchalska, **Paweł Krupa**, Magdalena J Ślusarz, Marcin Lubocki, Krystyna Bieńkowska-Szewczyk, Sylwia Rodziewicz-Motowidło, Adam K Sieradzan, Andrea D Lipińska, Journal of Molecular Biology 2023, 435 (5), 167964.

P17. “The unusual robustness of neurotransmitter vesicle membranes against serotonin-induced perturbations”, Ankur Gupta, **Pawel Krupa**, Oskar Engberg, Magdalena Krupa, Ankur Chaudhary, Mai Suan Li, Daniel Huster*, Sudipta Maiti*, The Journal of Physical Chemistry B 2023, 127 (9), 1947-1955.

P18. “What Makes Telomeres Unique?”, Adam K. Sieradzan*, **Paweł Krupa**, David Wales, The Journal of Physical Chemistry B 2017, 121(10), 2207-2219.

P19. “Computational Studies of the Mechanical Stability for Single-Strand Break DNA”, **Paweł Krupa***, David Wales, Adam K. Sieradzan, Journal of Physical Chemistry B 2018, 122, 8166-8173.

P20. “Integrating explicit and implicit fullerene models into UNRES force field for protein interaction studies”, Natalia H Rogoża, Magdalena A Krupa, **Pawel Krupa***, Adam K Sieradzan, Molecules 2024, 29(9), 1919.

P21. “Serotonin Promotes Vesicular Association and Fusion by Modifying Lipid Bilayers”, Debsankar Saha Roy, Ankur Gupta, Vicky Vishvakarma, **Pawel Krupa**, Mai Suan Li, and Sudipta Maiti*, Journal of Physical Chemistry B 2024, 128, 20, 4975–4985.

P22. “A Practical Guide to All-Atom and Coarse-Grained Molecular Dynamics Simulations Using Amber and Gromacs: A Case Study of Disulfide-Bond Impact on the Intrinsically Disordered Amyloid Beta”, Pamela Smardz, Midhun Anila, Paweł Rogowski, Bartosz Różycki, **Paweł Krupa***, International Journal of Molecular Sciences 2024, 25(12), 6698.

P23. “Multi-Layered Computational Framework for Designing Peptide Inhibitors of HVEM-LIGHT Interaction”, Piotr Ciura, Pamela Smardz, Marta Spodzieja, Adam K. Sieradzan*, **Paweł Krupa**, Journal of Physical Chemistry B 2024, 128, 28, 6770–6785.

R1. “Microscopic Physics-Based Models of Proteins and Nucleic Acids: UNRES and NARES”, M. Baranowski, C. Czaplewski, E. Gołaś, Y. He, D. Jagieła, **P. Krupa**, A. Liwo, G.G. Maisuradze, M. Makowski, M.A. Mozolewska, A. Niadzvedtski, A.J. Niemi, S. Rackovsky, R. Ślusarz, A.K. Sieradzan, S. Ołdziej, T. Wirecki, Y. Yin, B. Zaborowski, H.A. Scheraga*, w książce pt. Coarse-Grained Modeling of Biomolecules (Series in Computational Biophysics), Edytor: Garegin A. Papoian, ISBN 978-1466576063, Crc Pr I Llc 2017.

R2. „Use of coarse-grained and all-atom molecular dynamics to study Hsp70 and Hsp40 chaperone action”, E.I. Gołaś, M.A. Mozolewska, **P. Krupa**, C. Czaplewski, H.A. Scheraga, A. Liwo, w książce pt. Frontiers in Structural Biology, 2018, Vol. 1, 22-45 (Bentham Science Publishers), Edytor: Mario D. Galigniana.

R3. „Scale-consistent approach to the derivation of coarse-grained force fields for simulating structure, dynamics, and thermodynamics of biopolymers”, Adam Liwo, Cezary Czaplewski, Adam K. Sieradzan, Emilia A. Lubecka, Agnieszka G. Lipska, Łukasz Golon, Agnieszka Karczyńska, **Paweł Krupa**, Magdalena A. Mozolewska, Mariusz Makowski, Robert Ganzynkowicz, Artur Giełdoń, Maciej Maciejczyk, w książce pt. Computational Approaches for Understanding Dynamical Systems: Protein Folding and Assembly (część serii książek Progress in Molecular Biology and Translational Science) 2020, 170, 73-122. Edytorzy: Bogdan Barz, Brigit Strodel.

R4. “Physics-based coarse-grained modeling in bio- and nanochemistry”, Adam Liwo, Adam K. Sieradzan, Agnieszka S. Karczyńska, Emilia A. Lubecka, Sergey A. Samsonov, Cezary Czaplewski, **Paweł Krupa**, Magdalena Mozolewska, w książce pt. Practical Aspects of Computational Chemistry V (PACC), 2021, V, 31-69. Edytorzy: Jerzy Leszczyński, Manoj Shukla.

R5. “Modeling the Structure, Dynamics, and Transformations of Proteins with the UNRES Force Field”, Adam K. Sieradzan, Cezary Czaplewski, **Paweł Krupa**, Magdalena A. Mozolewska, Agnieszka S. Karczyńska, Agnieszka G. Lipska, Emilia A. Lubecka, Ewa Gołaś, Tomasz Wirecki, Mariusz Makowski, Stanisław Ołdziej, Adam Liwo, w książce pt. Protein Folding (część serii książek Methods in Molecular Biology (MIMB, volume 2376)), 2021, 399-416. Edytor: Victor Muñoz.

R6. “Computational models for study of protein aggregation”, TC Co, MS Li, **P. Krupa***, w książce pt. Computer Simulations of Aggregation of Proteins and Peptides, (część serii książek Methods in Molecular Biology (MIMB, volume 2340)), 2022, 51-78. Edytorzy: Mai Suan Li, Andrzej Kloczkowski, Marek Cieplak, and Maksim Kouza.

R7. “Free docking and template-based docking (physics versus knowledge-based docking)”, Magdalena A. Krupa, **Paweł Krupa***, w książce pt. Protein-protein docking: methods and protocols (część serii książek Methods in Molecular Biology (MIMB, volume 2780)), 2024, 27-41. Edytor: Agnieszka Kaczor.

Peer-reviewed English-language publication

M1. “High Performance Computing with Coarse Grained Model of Biological Macromolecules”, Emilia Lubecka, Adam Sieradzan, Cezary Czaplewski, **Paweł Krupa**, Adam Liwo; Journal of Supercomputing Frontiers and Innovations, 2018, 5(2), 63-75.

6. Presentation of significant scientific or artistic activity carried out at more than one university, scientific or cultural institution, especially at foreign institutions

Most of the scientific research I conducted after obtaining my doctoral degree was carried out at the Institute of Physics of the Polish Academy of Sciences, where I have been employed as an assistant professor since September 1, 2016. Earlier scientific research was conducted mainly at the Department of Theoretical Chemistry at the Faculty of Chemistry of the University of Gdańsk, with whose employees I continue to collaborate to this day.

In 2021, I also established cooperation with employees of the Department of Biomedical Chemistry at the Faculty of Chemistry of the University of Gdańsk, the Environmental Laboratory of Biological NMR at the Institute of Biochemistry and Biophysics of the Polish Academy of Sciences, and the Department of Molecular Biology of Viruses at the Intercollegiate Faculty of Biotechnology of the University of Gdańsk and the Medical University of Gdańsk, with whom I actively collaborate in the field of combined experimental and theoretical research. In 2023, we obtained funding for this research as a project consortium within the NCN Sonata project, and we are currently seeking to extend the funding for scientific research and expand the scope of joint projects as part of the submitted NCN Opus grant application.

My scientific activity is also conducted with many foreign centers as part of collaborations, internships, exchanges, and scientific visits, focusing on the experimental-theoretical combination of methods to obtain a complete picture of the studied physicochemical phenomena.

6.1 Scientific internships in foreign institutions in the years 2012-2016

- August 1, 2015 - July 31, 2016:

  Postdoctoral internship in Harold Scheraga's group at the Baker Laboratory of Chemistry and Chemical Biology, Cornell University, Ithaca, NY, USA.

  As part of the postdoctoral internship, I learned various computational techniques and acquired other scientific skills by participating in group meetings and lectures. I also worked on improving the predictive capabilities of the UNRES force field with reduced representation and adding new functionalities, such as updated dynamic treatment of disulfide bonds in proteins. I then used the developed methods to investigate the role of disulfide bonds in RNase A. Some of the results were written up in the form of scientific publications: Bioinformatics 32 (21), 3270-3278, 2016, Journal of Chemical Theory and Computation 13 (11), 5721-5730, 2017, J. Chem. Inf. Model. 57 (9), 2364-2377, 2017.

- November 17, 2014 - May 1, 2015:

  Research Assistant in Andrzej Kłoczkowski's group at the Battelle Center for Mathematical Medicine, The Research Institute at Nationwide Children's Hospital and Ohio State University in Columbus, OH, USA.

  The scientific internship focused on developing a tool for assessing the quality of protein models and scoring functions, allowing the selection of the most probable protein structures from large ensembles of structures using various machine learning techniques to optimize energy functions based on the physics of interactions. The results were published in Genes 9 (12), 601, 2018.

- May 29, 2014 - August 17, 2014, May 15, 2013 - August 15, 2013, and May 15, 2012 - August 14, 2012:

  Scientific internships in Harold Scheraga's group at the Baker Laboratory of Chemistry and Chemical Biology, Cornell University in Ithaca, NY, USA.

  As part of the scientific internships, I learned various experimental and theoretical techniques and conducted activities popularizing science among children, students, and teachers. The scientific part of the internship also included the development of the coarse-grained UNRES model by adding new potentials based on statistics and physics of interactions to improve the treatment of loops and other poorly structured protein elements during simulations. The internships resulted in numerous results, some of which were published in international journals: Proceedings of the National Academy of Sciences 110 (37), 14936-14941, 2013, Journal of Chemical Theory and Computation 9 (10), 4620-4632 2013, Proteins: Structure, Function, and Bioinformatics 82 (9), 1850-1868, 2014, Journal of Chemical Theory and Computation 11 (2), 817-831, 2015, Proteins: Structure, Function, and Bioinformatics 83 (8), 1414-1426, 2015, Proceedings of the National Academy of Sciences 112 (44), 13549-13554, 2015.

- January 17, 2014 - May 19, 2014:

  Scientific internship in Jerzy Leszczyński's group at the Department of Chemistry and Biochemistry, Jackson State University, MS, USA.

  During the scientific internship, I investigated the potential nanotoxicity of various molecules, such as fullerenes, carbon nanotubes, and gold nanoparticles in contact with proteins and nucleic acids, using approaches based on database searching, docking, DFT, molecular modeling, and molecular dynamics in an all-atom force field. The internship resulted in scientific collaboration, some of the results of which were published in journals: Nanoscale 9 (29) 2017, 10263-10276, Task Quarterly 18 (4), 337-341 2014.

6.2 Short-term scientific visits to foreign institutions in the years 2015-2019

- October 26, 2019 - November 1, 2019, September 30, 2018 - October 7, 2018, and September 10, 2017 - September 16, 2017:

  Scientific visits to Giovanni La Penna's group at the Consiglio Nazionale delle Ricerche (CNR), Istituto di Chimica dei Composti Organo Metallici (ICCOM), Florence, Italy.

  During three scientific visits conducted as part of a Polish-Italian bilateral project, I carried out and discussed with project participants a wide range of experimental and theoretical methods for studying the influence of metal ions on various biomacromolecules, especially proteins, peptides, and their aggregates, such as Aβ42 oligomers. Additionally, I investigated the role of divalent ions and the influence of amyloid β peptides and oligomers on lipid membranes of different compositions. Some of the obtained results were summarized in the form of scientific publications: J. Phys. Chem. B 124 (16), 3300-3314, 2020, J. Phys. Chem. B 2022, 126, 20, 3659–3672, and Int. J. Mol. Sci. 2023, 24(16), 12698.

- October 21-28, 2018:

  Participation in a scientific conference and a scientific visit to the Institute for Computational Science and Technology, Ho Chi Minh City, Vietnam, which resulted in scientific collaboration on multiscale studies of amyloid β oligomers and the following

publications: J. Phys. Chem. B 2019, 123, 34, 7253–7269 and J. Phys. Chem. B 2022, 126, 20, 3659–3672.

- May 22-29, 2015:

  Participation in a scientific conference and a scientific visit to the Korea Institute for Advanced Study, Seoul, Republic of Korea, in Professor Jooyoung Lee's group, which resulted in scientific collaboration and publication of combined bioinformatics-biophysical methods to improve the predictive capabilities of computational methods: J. Chem. Inf. Model. 2015, 55, 6, 1271–1281, J. Chem. Inf. Model. 2016, 56, 11, 2263–2279, and J Comput Chem. 2017 Dec 5;38(31):2730-2746.

6.3 Some of the more important scientific collaborations with foreign scientific centers in the years 2019-2024

- Since 2023: Dr. Kien Xuan Ngo, WPI-Nano Life Science Institute, Kanazawa University, Japan - research on combining experimental results (high-speed atomic force microscopy, HS-AFM) and computer simulations to determine the mechanism of α-hemolysin association on the surface and in lipid membranes - publication manuscript in preparation.
- Since 2022: Prof. Sudipta Maiti, Tata Institute of Fundamental Research, India - research on the role of neurotransmitters and the influence of their presence on lipid membranes. Some of the results were published in: J. Phys. Chem. B 2023, 127, 9, 1947–1955 and J. Phys. Chem. B 2024, doi: 10.1021/acs.jpcb.4c00115.
- Since 2021: Prof. Zuzana Gazova, Department of Biophysics, Institute of Experimental Physics, Slovak Academy of Sciences - research on the influence of various substances of natural origin on the aggregation process of amyloid β and lysozyme - a paper titled "Coumarin derivatives – the role of dimerization and linker length on the anti-amyloid activity" is under review in the Journal of Biological Chemistry.
- Since 2019: Dr. Giovanni Roviello, The Institute of Biostructures and Bioimaging (IBB) of the National Research Council (CNR), Naples, Italy - research on the influence of substances of natural origin on amyloid β aggregation. The results were published in Chemico-Biological Interactions 2021, 334, 109300.

7. Presentation of teaching and organizational achievements as well as achievements in popularization of science or art

7.1 Teaching courses for students

- Scientific supervision over the invitation of a lecturer, Prof. David Wales, to deliver a series of prestigious lectures as part of the "Spotlight Talk" at the Warsaw Doctoral School of Exact and BioMedical Sciences, implemented and financed under the "STER programme of the Polish National Agency for Academic Exchange, grant no. BPI/STE/2021/1/00034/U/00001".
- Lecturer (independently and in cooperation with other scientists) at the Warsaw Doctoral School of Exact and BioMedical Sciences:
  - "Practical Aspects in Computational Biophysics", 30 hours, 2023.
  - "Introduction to theoretical and computational biophysics", 30 hours, 2022.
  - "Computational biophysics & Materials Science", 30 hours, 2019.

- "Computational Methods in Materials Science and Biology", 30 hours, 2017/2018. Assistant during Programming classes, 30 hours, 2014/2015.

- Lecturer in Faculty of Chemistry, University of Gdańsk
  - Instructor of three groups of Information Technology, 90 hours, 2013/2014.
  - Assistant during Information Technology classes, 30 hours, 2012/2013.
  - Assistant during Theoretical Chemistry classes, 45 hours, 2011/2012.

7.2 Promoting students

- Assistant supervisor of M.Sc. Pamela Smardz, conducting research within the NCN Sonata project, which I lead at the Institute of Physics of the Polish Academy of Sciences.
- Assistant supervisor of M.Sc. Agnieszka Karczyńska, a doctoral student at the Faculty of Chemistry of the University of Gdańsk from December 16, 2015 - title of the doctoral dissertation: "Prediction of the structure of protein complexes using the coarse-grained UNRES force field with the use of information from databases". Doctoral defense: September 11, 2019.

7.3 Organizational activities and popularization of science

- Planning, developing, organizing, and managing a local computing cluster at the Institute of Physics of the Polish Academy of Sciences, currently consisting of, among others, several specialized graphics cards allowing for the execution of molecular dynamics calculations two orders of magnitude faster than in the case of typical processors (CPUs).
- Organization of workshops with physical and chemical experiments for children, youth, and high school teachers during scientific internships in the group of Prof. Harold Scheraga at Cornell University, as part of activities popularizing science (outreach): August 8, 2015, August 2, 2014, August 3, 2013, and July 12, 2012.
- Co-organization of the From Computational Biophysics to Systems Biology (CBSB14) conference in Gdańsk on July 25-27, 2014.
- Co-organization of the Baltic Science Festival at the Faculty of Chemistry of the University of Gdańsk on May 23-24, 2014, and May 26, 2011.
- Co-organization of the open day at the Faculty of Chemistry of the University of Gdańsk in 2012.

8. Other information

8.1 Scientific projects granted at the national level, in which the habilitation candidate acts as the principal investigator

- Research task leader on behalf of IF PAN in the NCN Sonata project consortium 2022/47/D/NZ7/02399, titled "Identification of key structural elements of immunomodulatory proteins - inhibitors of antigenic peptide transporter TAP encoded by selected alphaherpesviruses", August 1, 2023 - July 31, 2026. Project in progress.
- Principal investigator of the NCN Sonata project UMO-2019/35/D/ST4/03156 titled "Computational studies of the disulfide-bond role in proteins", June 10, 2020 - June 9, 2024. Project completed.

- Principal investigator of the NCN Preludium project UMO-2015/17/N/ST4/03937 titled "Global protein docking algorithms based on the UNRES coarse-grained model", January 28, 2016 - January 27, 2018. Project completed and positively settled.

8.2 Review activities

Author of 58 reviews of scientific manuscripts in international journals: Biophysical Journal (1), Chemical Papers (16), Journal of Molecular Graphics and Modelling (3), Computational and Structural Biotechnology Journal (3), International Journal of Molecular Sciences (7), Molecules (4), Biomolecules (2), Axioms (4), J — Multidisciplinary Scientific Journal (2), The Journal of Physical Chemistry B (2), Journal of Chemical Theory and Computation (2), ACS Chemical Neuroscience (2), Frontiers in Molecular Biosciences Biophysics (1), Journal of Physics D: Applied Physics (1), Journal of Biomolecular Structure & Dynamics (3), Physchem (2), Proteins: Structure, Function and Bioinformatics (3).

Reviewer of grant projects for Teagasc Walsh Scholarship - the Agriculture and Food Development Authority, an organization associated with the Irish state apparatus responsible for the evaluation and granting of funding for research and development projects.

8.3 Invited lectures delivered at conferences of international character

- 26-28.02.2024 – Ninth Korean-Polish Conference "Protein Folding: Theoretical and Experimental Approaches", Seoul, Republic of Korea.
- 24-28.09.2023 – Eighth Polish-Korean Conference "Protein Folding: Theoretical and Experimental Approaches", Jastrzębia Góra, Poland.
- 14-16.09.2022 – Symposium of Polish Bioinformatics Society, Warsaw, Poland.
- 07-08.2019 – The Fourth Workshop of Vietnamese Students in Poland, Warsaw, Poland.
- 25.10.2018 – The First ICST Workshop on Computational Biophysics and Medicine, Ho Chi Minh city, Vietnam.
- 14-18.08.2017 – 2nd International Conference on Computational Genomics and Proteomics, Playa Blanca, Panama.
- 25-29.07.2016 – Coarse-Grained Modeling of Structure and Dynamics of Biomacromolecules, Telluride, CO, USA.
- 01-04.12.2014 – 2nd Mathematical and Computational Medicine Conference, Cancun, Mexico.

8.4 Other lectures delivered at conferences of international character

- 27-31.10.2018 - 14th Rencontres du Vietnam, Computational Biophysics at the Molecular and MesoScales, Quy Nhon, Vietnam.
- 07.10.2017 - Coarse graining of biomolecules and beyond: theory and applications, Warsaw, Poland.
- 12-16.06.2017 - Physics and Biology of Proteins, Natal, Brazil.
- 28.05-01.06.2016 - Second Polish-Korean Conference on Protein Folding Theoretical and Experimental Approaches, Gdańsk, Poland.

- 24-28.05.2015 - The First Korean-Polish Conference on Protein Folding: Theoretical and Experimental Approaches, Seoul, Republic of Korea.

9. References

1. Case, D. A. *et al.* AmberTools. *J. Chem. Inf. Model.* **63**, 6183–6191 (2023).
2. Anfinsen, C. B. Principles that govern the folding of protein chains. *Science* **181**, 223–230 (1973).
3. Hegyi H., Gerstein M., The relationship between protein structure and function: a comprehensive survey with application to the yeast genome. *J. Mol. Biol.* **288**, 147–164 (1999).
4. Harrington, D. J., Adachi, K. & Royer, W. E. The high resolution crystal structure of deoxyhemoglobin S. *J. Mol. Biol.* **272**, (1997).
5. Trivedi, R. & Nagarajaram, H. A. Intrinsically Disordered Proteins: An Overview. *Int. J. Mol. Sci.* **23**, (2022).
6. Yang, J. *et al.* The I-TASSER Suite: protein structure and function prediction. *Nat. Methods* **12**, 7–8 (2015).
7. Baek, M. *et al.* Accurate prediction of protein structures and interactions using a three-track neural network. *Science* **373**, 871–876 (2021).
8. Abramson, J. *et al.* Accurate structure prediction of biomolecular interactions with AlphaFold 3. *Nature* (2024) doi:10.1038/s41586-024-07487-w.
9. Ghag, G. *et al.* Disulfide bonds and disorder in granulin-3: An unusual handshake between structural stability and plasticity. *Protein Sci.* **26**, 1759–1772 (2017).
10. Sugita, Y. & Okamoto, Y. Replica-exchange molecular dynamics method for protein folding. *Chem. Phys. Lett.* **314**, 141–151 (1999).
11. Kumar, S., Rosenberg, J. M., Bouzida, D., Swendsen, R. H. & Kollman, P. A. THE weighted histogram analysis method for free‐energy calculations on biomolecules. I. The method. *J. Comput. Chem.* **13**, 1011–1021 (1992).
12. Kryshtafovych, A., Monastyrskyy, B. & Fidelis, K. CASP11 statistics and the prediction center evaluation system. *Proteins* **84 Suppl 1**, 15–19 (2016).
13. Thu, T. T. M., Co, N. T., Tu, L. A. & Li, M. S. Aggregation rate of amyloid beta peptide is controlled by beta-content in monomeric state. *J. Chem. Phys.* **150**, 225101 (2019).
14. Khaled, M. *et al.* A Hairpin Motif in the Amyloid-β Peptide Is Important for Formation of Disease-Related Oligomers. *J. Am. Chem. Soc.* **145**, 18340–18354 (2023).
15. Huang, F. *et al.* Co-aggregation of α-synuclein with amyloid-β stabilizes β-sheet-rich oligomers and enhances the formation of β-barrels. *Phys. Chem. Chem. Phys.* **25**, 31604–31614 (2023).
16. Lee, J. *et al.* Distinguishing protein chemical topologies using supercharging ion mobility spectrometry‐mass spectrometry. *Angew. Chem. Weinheim Bergstr. Ger.* (2023) doi:10.1002/ange.202314980.

Paweł Krupa