Scientific Research Publishing

Advances in Pure Mathematics, 2026, 16(9), 740-764
https://www.scirp.org/journal/apm
ISSN Online: 2160-0384
ISSN Print: 2160-0368

# Multi-Criteria Optimisation: The Environment-Dependent Protein Folding Process Simulation Tool

**Irena Roterman[1*], Katarzyna Stapor[2], Dawid Dulak[3], Leszek Konieczny[4]**

[1]Department of Bioinformatics and Telemedicine, Medical College, Jagiellonian University, Krakow, Poland
[2]Faculty of Automatic, Electronics and Computer Science, Department of Applied Informatics, Silesian University of Technology, Gliwice, Poland
[3]ABB Business Sp z o. o., Warszawa, Poland
[4]Chair of Medical Biochemistry, Medical College, Jagiellonian University, Krakow, Poland
Email: *myroterm@cyf-kr.edu.pl, katarzyna.stapor@polsl.pl, dawid.dulak@gmail.com, mbkoniec@cyf-kr.edu.pl

**How to cite this paper:** Roterman, I., Stapor, K., Dulak, D. and Konieczny, L. (2026) Multi-Criteria Optimisation: The Environment-Dependent Protein Folding Process Simulation Tool. *Advances in Pure Mathematics*, **16**, 740-764.
https://doi.org/10.4236/apm.2026.169036

**Received:** August 20, 2026
**Accepted:** September 21, 2026
**Published:** September 24, 2026



## Abstract

The multiple minima problem in protein structure prediction arises from the large number of structural forms corresponding to the energy states of intramolecule interactions that represent local or global minima that ensure stability. To indicate an appropriate energy minimum and its related spatial structure by simulating the folding process remains a puzzle that has not been solved using existing methods. The proposition presented herein is related to the necessity of treating the folding process as a mechanism based on multi-criteria optimisation that includes both internal force field and external force field optimisation. The protein folding process proved to be dependent on external environmental conditions. This requires treating the protein folding process as multi-criteria optimisation using the technique called the Pareto front. The need of using the multi-criteria optimisation procedure, as well as the propositions to implement this solution to folding process simulation tools are presented in this study.

## Keywords

Multi-Criteria Optimisation, Internal Force Field, External Force Field, Protein Folding, Funnel Model

## 1. Introduction

Artificial intelligence-based techniques have been greatly successful in predicting the spatial structure of proteins for a specific sequence of amino acids. This

method is used in AlphaFold, a tool that provides a qualified description of its prediction capabilities and limitations, especially its inability to directly simulate folding pathways [1]. The deep learning artificial intelligence technique applied uses a rich database of protein structures (currently more than 250,000 structures) available in the Protein Data Bank [2]. The structure is designed based on maps of inter-residual contacts in protein structures [3]. The possibility to predict the correct protein structure is greatly prominent in the field of computer-assisted drug design, which is impossible without knowing the spatial structure of the target protein. The expected correction of protein activity by a drug requires knowing the correct structure of the protein [4]. The techniques of experimental determination of the spatial structure are expensive and time-consuming and sometimes impossible due to the complex environmental conditions for obtaining an isolated protein [5]. Thus, the introduction of numerical techniques is important for progress in pharmacy and pharmacology [6]-[9].

From a scientific point perspective, however, the question of mechanism behind this process, which occurs incessantly in every living organism by replacing proteins subject to degradation with new proteins, remains unanswered. On every occasion, this process provides proper structure to guarantee biological activity. However, it is not failure-free. Misfolded proteins are subject to the system degradation or correction process called the unfolding process response (UPR) [10]-[12].

The existing numerical techniques of protein structure prediction based on a specific amino acid sequence consist in searching for such a conformation that presents a low energy condition that guarantees protein stability. This is related to inter-atomic internal interactions. The difficulty is to identify the proper energy minimum, as they are very numerous. This is the origin of the term multiple minima problem [13] [14].

Specialised protein structure prediction tools are based on the process of minimisation of the energy expressing the internal force field. These are non-binding interactions (electrostatic, vdW and torsional potential) that determine the stable arrangement that eliminates collisions in the form of high energy states. The project that tracks progress in the field, known as CASP (Critical Assessment of Structure Prediction), initiated in 1996 enables monitoring progress in this field by assessing the predicted structures obtained by project participants in the blind prediction system [15]-[17].

The recent editions are dominated by AI techniques (AlphaFold), which however fail to answer the question about the process of proper polypeptide chain folding that produces stable structures, while retaining their form of biological activity—a strictly defined function.

The analysis of the models submitted by competition participants showed a strong dependency of the production of a correct structure on environmental conditions [18] [19]. The models submitted in the project were assessed using the criterion of the fuzzy oil drop model (FOD-M) presented herein [19]. A designed

model correctness criterion was demonstrated that included the match degree of the predicted structure to the environmental conditions of the native protein. It should be noted that the so-called very good structure prediction programs that show high correctness for one protein fail for another protein. This means that the tool is not universal and the end result depends on a specific protein that requires different environmental conditions corresponding to the parameters of the force field in use to form. This applies to programs that only rely on internal force field optimisation (inter-atomic interactions in protein body) [20]-[23].

The purpose of this study is to demonstrate the necessity of replacing energy function optimisation limited to the internal force field with multi-criteria optimisation that includes the dependency of the polypeptide chain structure formation process on environmental conditions—the external force field. This entails the need to consider the folding process simulation model as multi-criteria optimisation. The introduction of state function as a function dependent on two functions with opposing direction (the minimisation of one results in an increased value of another), which requires using the Pareto front model [24]-[28].

## 2. Materials and Methods

### 2.1. Data

The reasoning is conducted based on the analysis of example varied protein structures (protein size) with different biological activity showing varied distributions of hydrophobicity as impact of environmental factors on the folding process. This entails active involvement of the environment in the direction of the protein folding process, which should be included in the simulation of the protein folding process.

Table 1 provides selected examples with the determination of variation from the perspective of the structure assessed according to the hydrophobicity distribution criterion. Water—essential and necessary for life processes, the environment forces its specific characteristics and cooperates with the building blocks of all components of living organisms that seem to be adapted to cooperate with water. The analysis also covered the proteins representing the so-called misfolding, which consists in preserving an entirely different structure without any biological activity, which results in significant pathological disturbances [29]. This refers to proteins referred to as amyloids. The analysis also covered the example of protein assisting the folding process of other proteins by introduction to a locally different environment. Such proteins are referred to as chaperonins. They are large structures in capsule form that generate an environment that locally differs from the aqueous environment inside the capsule. This environment generates specific conditions, in which the folding protein adjusts its structure to obtain the expected form of structure, as well as the form of biological activity formed in it. The presented examples confirm the need to include the representation of altered environmental conditions in the folding process with protein folding process simulation-focused tools.

**Table 1.** Set of proteins analysed in this study. The detailed analysis applies to selected examples given in bold and underlined. The number of amino acids in chains are given according to original files available in PDB. However the analysis is limited to chain fragment of common length and common amino acid sequences for native and amyloid forms.

| PDB ID | Characteristics | Chain length | Ref |
|---|---|---|---|
| **2LX2** | Antifreeze | 67 | [30] |
| 2WXC | Ultra-fast | 47 | [31] |
| 2L42 | DNA binding | 106 | [32] |
| 2L6Q | Ultra-fast | 62 | [33] |
| 1WAE | Enzyme | 336 | [34] |
| 1BBL | Enzyme | 51 | [35] |
| **2PJ6** | Enzyme | 306 | [36] |
| 2LNL | Membrane protein | 309 | [37] |
| 5YWY-A | Signalling | 332 | [38] |
| 1ABZ | Kinase | 109 | [39] |
| 1F4Z | Membrane protein | 227 | [40] |
| **3QDC** | Rhodopsin | 239 | [41] |
| 7SSX | Membrane channel | 856 | [42] |
| **1DVQ** | Transthyretin | 124 | [43] |
| **4BJL-VL** | Immunoglobulin domain | 110 | [44] |
| **6SDZ** | Amyloid | 127 | [45] |
| **6HUD** | Amyloid | 136 | [46] |
| **3IZI** | Chaperonin | 8208 | [47] |
| | Chaperonin | 8708 | |
| 8WUC | GroEL | 14x528 | [48] |
| | GroES | 14x94 | |

The following detailed analysis presents selected examples (2LX2, 2PJ6, 3QDC) that represent the impact of aqueous environment on the folding process with an increasing involvement of changing characteristics of aqueous environment.

Ultra-fast means proteins characterised by a very fast folding process after experimental unfolding and re-folding [49]. Antifreeze proteins are proteins solved in water that occur in organisms that survive hibernation. Their role consists in introducing a water particle arrangement that excludes the formation of ice by imposing an arrangement adjusted to the charge distribution at the antifreeze protein surface [50]. Membrane proteins are proteins anchored in the membrane and thus in a hydrophobic environment opposite to the polar water environment [51]. Enzymes are proteins that show specific activity that requires interaction with a specific substrate and conducting a specific reaction to form a product necessary for other processes occurring in a cell or in extracellular space. Membrane channel is a protein anchored in the membrane with a central channel that enables specific

transport of suitable components to the inside of the cell [52]. Such structure requires the exposure of hydrophobic amino acids outside with centrally positioned polar amino acids used to build channel walls, in which the transport occurs primarily by electrostatic interactions [51]. The set of proteins representing wide spectrum of different relations between expected and observed distribution of hydrophobicicty is presented in [53].

A representative of proteins subject to amyloid transformation was also analysed—a misfolded form that causes significant neurodegenerative disorders (Parkinson's, Alzheimer's and other). This product is stored (not subject to degradation), which results in disorders primarily in brain functioning [29]. A representative of proteins referred to as chaperonins was also analysed. They are molecular chaperones that accompany the protein folding process. Chaperonins are mostly large multi-chain structures in capsule shape, inside which the folding process of specific proteins occurs. The role of chaperonins proves the need to introduce an environmental variable for this process in the form of force field, to which the structure of the folding protein is adjusted.

The set of proteins presented in **Table 1** presents variation by chain length (the aa number), biological activity and placement within a cell or biological fluids. The analysis of these proteins using the fuzzy oil drop model (FOD-M) shows significant importance for the homoeostasis of the protein world, which is preserved with a significant impact of the aqueous environment and its aim-oriented modification.

## 2.2. Description of the Model

The internal force field defines a set of non-binding interaction types, for which the optimum spatial arrangement representing advantageous mutual arrangements of interacting components is searched for.

The internal force field is expressed as a function in the form of:

$$E_{\mathrm{INT}} = E_{\mathrm{INT}}[\text{electrostatic } (r_{ij}), \text{ vdw } (r_{ij}), \text{ tp (rotation angle)}] \quad (1)$$

where tp is the torsional potential.

The selection of an appropriate energy minimum of the function $E_{\mathrm{INT}}$ is impeded by the so-called multiple minima problem. Thus, the protein structure can assume multiple stable forms by positioning itself in one of multiple local energy minima. Finding that proper minimum requires indicating a structure that guarantees stability, but also permitting the acceptable specific flexibility required to ensure a biological function.

External force field is a term that directs the environmental impact, which represents specific properties and has a characteristic impact on the optimum arrangement of individual components in relation to the local environment.

The form of the external force field expressed using the fuzzy oil drop model is the following:

$$E_{\mathrm{EXT}} = E_{\mathrm{EXT}}(\text{field-adjusted structure in the form of continuum}).$$

The following description of the form of the external force field explains the principle of writing the environmental impact on the function, which continu-

ously expresses the adjustment of the arrangement of specific components in a protein body meeting the requirements arising from the external force field characteristics.

The proposition is based on an elementary phenomenon, which results in arranging bi-polar molecules (molecules with composition varied relative to the surrounding water particles) in micelle form. The form of this structure is arranged to expose polar parts to the environment (*i.e.* to the surface), focusing non-polar parts in the centre, resulting in isolation of the hydrophobic parts from the polar environment. It should be noted that aqueous environment is a polar environment that conditions the functioning of all living organisms. It is a commonly known notion that life originated in water.

Thus, the starting point is a model arising from the specific characteristics of the polar environment of water. The notation of an idealised distribution of hydrophobicity (entire spectrum from total non-polarity to polar components) is expressed by a 3D Gaussian function spanning on the entire object body (in protein body). This function expresses the highest concentration of hydrophobicity in protein centre, with hydrophobicity levels decreasing away from the centre and reaching zero levels of hydrophobicity on the surface. Parameters $\sigma_x$, $\sigma_y$ and $\sigma_z$ express adjustment to protein size and shape.

$$T_i = \frac{1}{\sum_{i=1}^{N} T_i} \exp\left[\frac{-\left(x_i - \overline{x}\right)^2}{2\sigma_x^2}\right] \exp\left[\frac{-\left(y_i - \overline{y}\right)^2}{2\sigma_y^2}\right] \exp\left[\frac{-\left(z_i - \overline{z}\right)^2}{2\sigma_z^2}\right] \tag{2}$$

The first expression in Equation (2) results in the normalisation of the distribution defined as $T$ (theoretical − idealised). The method of determining appropriate values for parameters $\sigma_x$, $\sigma_y$ and $\sigma_z$ is important. The X-axis (*i.e.* $\sigma_x$) coincides with the longest inter-atomic distance within the molecule. The Y-axis coincides with the greatest molecule "thickness" (the greatest distance between the positions of atoms in the YZ plane—the distance between projections of atom positions in the YZ plane). In this orientation (according to the three-sigma rule), $\sigma_x$, $\sigma_y$ and $\sigma_z$ parameter values are determined. The method of determining these parameter values will prove important for the structural transformation simulation in misfolding proteins (discussed further in this study).

The actual distribution however results from inter-residual hydrophobic interactions. It is expressed by the function proposed by M. Levitt [54].

$$O_i = \frac{1}{\sum_{i=1}^{N} O_i} \sum_j \begin{cases} \left(H_i^r + H_j^r\right)\left(1 - \frac{1}{2}\left(1 - \frac{1}{2}\left(7\left(\frac{r_{ij}}{c}\right)^2 - 9\left(\frac{r_{ij}}{c}\right)^4 + 5\left(\frac{r_{ij}}{c}\right)^6 - \left(\frac{r_{ij}}{c}\right)^8 - \left(\frac{r_{ij}}{c}\right)^8\right)\right)\right), \text{for } r_{ij} \le c \\ 0, \text{for } r_{ij} > c \end{cases} \tag{3}$$

It includes $c$—the cutoff distance (9 Å) and the distance ($r_{ij}$) between the positions of the so-called effective atoms (averaged position of atoms included in the given amino acid). The function defined as $O$ (observed), also normalised (the first expression in Equation (3)), assigns an effective hydrophobicity level $O_i$ to

atom positions according to impact.

Distributions are compared based on divergence entropy proposed by Kullback-Leibler [55].

$$D_{KL}\left(P \mid Q\right) = \sum_{i=1}^{N} P_i \log_2 \frac{P_i}{Q_i} \tag{4}$$

where $P_i$—the distribution analysed (the $O$ distribution in our model), $Q$—the reference distribution (the $T$ distribution in our model).

The value of $D_{KL}(O|T)$ expressing the similarity of the $O$ distribution in relation to reference distribution (idealised in micelle-like form) cannot be interpreted directly (entropy) (Equation (4)). Thus, the measure of the distance between the $O$ distribution and the second reference distribution in the form of $R_i = 1/N$ is introduced, where $N$ is the number of amino acids in the chain. This distribution represents an even distribution of hydrophobicity levels. It is thus uniform distribution opposite to the 3D Gaussian function distribution, in which the hydrophobic nucleus is present in the central part of protein body.

The comparison of $D_{KL}(O|T)$ to $D_{KL}(O|R)$ facilitates approximation—the similarity of the $O$ distribution in relation to one of two opposite reference distributions ($T$ and $R$). For comparison purposes, parameter RD (Relative Distance) (Equation (5)) of the form was introduced

$$\mathrm{RD} = \frac{D_{KL}\left(O \mid T\right)}{D_{KL}\left(O \mid T\right) + D_{KL}\left(O \mid R\right)} \tag{5}$$

The value of RD < 0.5 approximates the $O$ distribution to the $T$ distribution, which indicates the presence of the hydrophobic nucleus in the analysed protein structure.

The existing part of the FOD-M model is related to the basic environment, *i.e.* water. Proteins are also active in other environments. Such environment is a cellular membrane with specific characteristics opposite to water. The membrane is a hydrophobic environment. For the stabilisation of membrane proteins, the exposure of hydrophobic residues with a concentration of polar residues in the centre (to function as the ion channel) (Equation (6)). The notation form of the appropriate idealised distribution of the membrane environment is:

$$T_{MAX} - T_i \tag{6}$$

To generalise the diversity of the environments, in which proteins are active (and thus with structure adjusted to the given environment), the function is proposed to express the consensus between the involvement of aqueous environment and hydrophobic components in the following form:

$$M_i = \left[T_i + K * \left(T_{MAX} - T_i\right)_n\right]_n \tag{7}$$

where the $n$ index denotes normalisation.

An important parameter for the proposed form of external force field is $K$, of which the value indicates the degree of involvement of the non-aqueous factor. Value $K = 0.0$ denotes an aqueous environment directed to fold proteins towards generating a hydrophobicity nucleus and a polar surface. The value of parameter

$K$ does not have a maximum limit. The examples discussed in this study will show the variation of this parameter.

The variation of the external force field results from experimental discoveries in the form of activity of proteins referred to as molecular chaperones (including chaperones and chaperonins) (Equation (7)). They are proteins that accompany the folding process in the form of closed capsule, in which the protein folding process is isolated from the aqueous environment. The degree of diversity and variation of the form of the external force field proves to represent the entire spectrum of variation.

The FOD-M model shows an arrangement close to the micelle-like arrangement preferred by aqueous environment. This model enables the identification of local disorder. This form of local disorder assumes the form of local hydrophobicity excess. If it takes place on the surface, it entails potential readiness to form assemblies with another protein with a similar defect [55].

The local hydrophobicity deficit is very often connected with the presence of cavity prepared for ligand binding or for interaction with substrates in case of enzymes [56]. The biological function is coded in the protein structure specifically in the form of local mismatches of the $O$ distribution in relation to the $T$ distribution. Perfect replication of the hydrophobicity distribution in the form of 3D Gaussian function would result in very good solubility without any interaction potential, except for accidental unstable contacts with ions or other polar particles. The local mismatch of the $O$ distribution in relation to the $T$ distribution is a code for identifying other particles, including proteins with similar irregularity. Consequently, proteins are referred to as an intelligent micelle, in which a local structural defect (according to FOD-M model criteria) is a record of readiness for potential interaction, *i.e.* for proper biological function [57].

The graphical presentation of the procedure is shown in **Figure 1**.

The presentation of **Figure 1(e)** represents the degree of variation of external conditions. Profile $T$ expresses an idealised distribution determined by the conditions of aqueous environment. The introduction of a non-zero component of Equation (6) (value $K \neq 0.0$), representing the involvement of opposite components in the environment, generates the $M$ function, to which the folding protein adjusts its distribution of hydrophobicity within the molecule. Examples of structures are identified, for which the directional coefficient of the function expressing Mi relations in relation to $T_I$ assumes negative values. It entails a structure that not only mitigates extremes by introducing a mild course of the function. A negative directional coefficient entails the introduction of an arrangement inverse of the water-induced arrangement.

The following examples of varied structure show the entire spectrum of the potential to obtain changes in the distribution of hydrophobicity in proteins with varied structure, biological function and activity environments (e.g. membrane proteins anchored in membrane environment).

The $K$ value derives retrospectively from deposited structures. However the description using $K$ parameter may be applied to models submitted in CASP project

[17] [18]. The models deposited in CASP16 were verified regarding the evaluation score revealing accordance and discordance between FOD-M based assessment in respect to CASP criteria [18]. The high accordance between the CASP-system evaluation and RD-dependence has been identified [17]. So far only limited simulation of protein folding influenced by $K$ value has been performed [18].

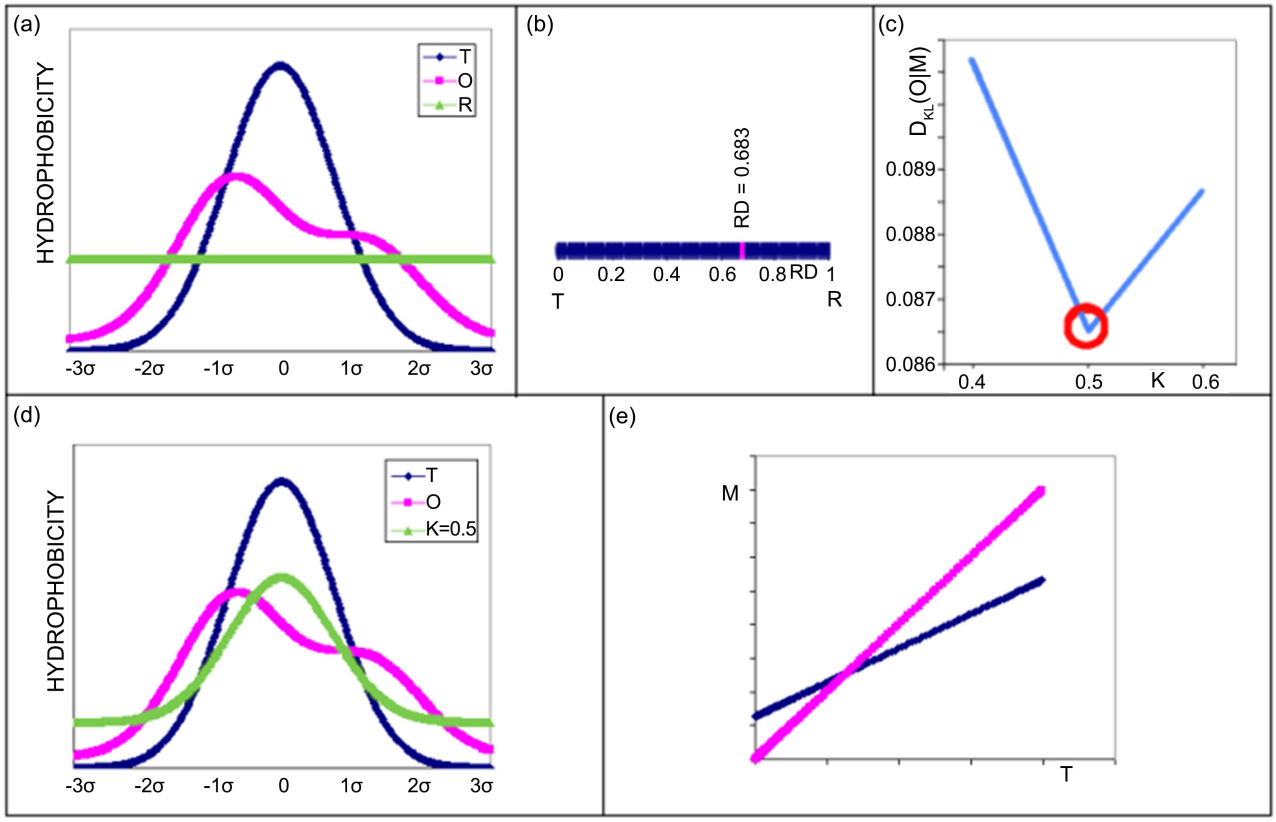


**Figure 1.** Graphical presentation of the procedure under the FOD-M model, reduced to 1D for readability. (a) $T$ (blue), $O$ (pink) and $M$ (light green) distribution; (b) Value of parameter RD = 0.683; (c) Determination of the value of parameter $K$ for the minimum value of $D_{KL}(O|M)$; (d) $T$ (blue), $O$ (pink) and $M$ (light green) distribution for $K = 0.5$; (e) relation of the $M$ distribution to the $T$ distribution.

## 2.3. Programs Used

The programs for determining the $T$, $O$ and $M$ distributions and the values of the RD and $K$ parameters are available in the open access system.

The potential user has two possible access to the program:

The program allowing calculation of RD as well as $T$ and $O$ distribution is accessible upon request on CodeOcean platform:
https://codeocean.com/capsule/3084411/tree. Please contact the corresponding author to get access to your private program instance.

The application—implemented in collaboration with the Sano Centre for Computational Medicine (https://sano.science) and running on resources contributed by ACC Cyfronet AGH (https://www.cyfronet.pl) in the framework of the PL-Grid Infrastructure (https://plgrid.pl)—provides a web wrapper for the above-

mentioned computational component and is freely available at https://hphob.sano.science. The output file presents: columns C—*T* distribution, column D—*O* distribution, column P—*M* distribution. The values of RD and K parameters are available in a separate output file.

The VMD program was used to present the 3D structures (https://www.ks.uiuc.edu/Research/vmd/, accessed Dec. 2025) [58].

## 3. Results

### 3.1. Variation of Protein Structures in FOD-M Model View

In Table 2, proteins representing different forms of activity dependent on the form of mismatch of the *O* distribution to the *T* distribution, show variation according to FOD-M model criteria.

Table 2. Examples of proteins with varied characteristics based on the FOD-M model. The proteins given in bold and underlined—the detailed analysis is presented.

| PDB ID | RD | *K* |
|---|---|---|
| **2LX2** | 0.312 | 0.0 |
| 2WXC | 0.382 | 0.1 |
| 2L42 | 0.387 | 0.2 |
| 2L6Q | 0.473 | 0.3 |
| 1WAE | 0.490 | 0.4 |
| 1BBL | 0.556 | 0.5 |
| **2PJ6** | 0.563 | 0.6 |
| 2LNL | 0.677 | 0.7 |
| 5YWY-A | 0.666 | 0.8 |
| 1ABZ | 0.681 | 0.9 |
| 1F4Z | 0.705 | 1.0 |
| **3QDC** | 0.786 | 1.3 |
| 7SSX | 0.742 | 1.5 |

### 3.2. Proteins Active in Aqueous Environment—Low *K*

For protein characteristics with very low *K*, a representative of antifreeze proteins was selected, type III (PDB ID 2LX2).

Antifreeze proteins are proteins that prevent water freezing. Their operating principle consists of forcing water particles into an arrangement aligned with the distribution of charges on the surface of the protein. This process prevents the formation of ice structure. The expectations for this function are as follows: good solubility, no interaction with other molecules except for water (elimination of complex formation of other proteins) and structural stability. All these objectives ensure protein complex formation with stable hydrophobic nucleus and polar surface. The protein represents a distribution of hydrophobicity according to the mi-

celle-like arrangement to reflect a distribution matching the 3D Gaussian function. The parameter value of $K = 0.0$ denotes perfect match of the O distribution to the T distribution in the given protein. It is illustrated by the summary of $T$ and $O$ profiles and m for $K = 0.0$. The $M$ distribution is identical with the $T$ distribution in that case (Figure 2(a)).

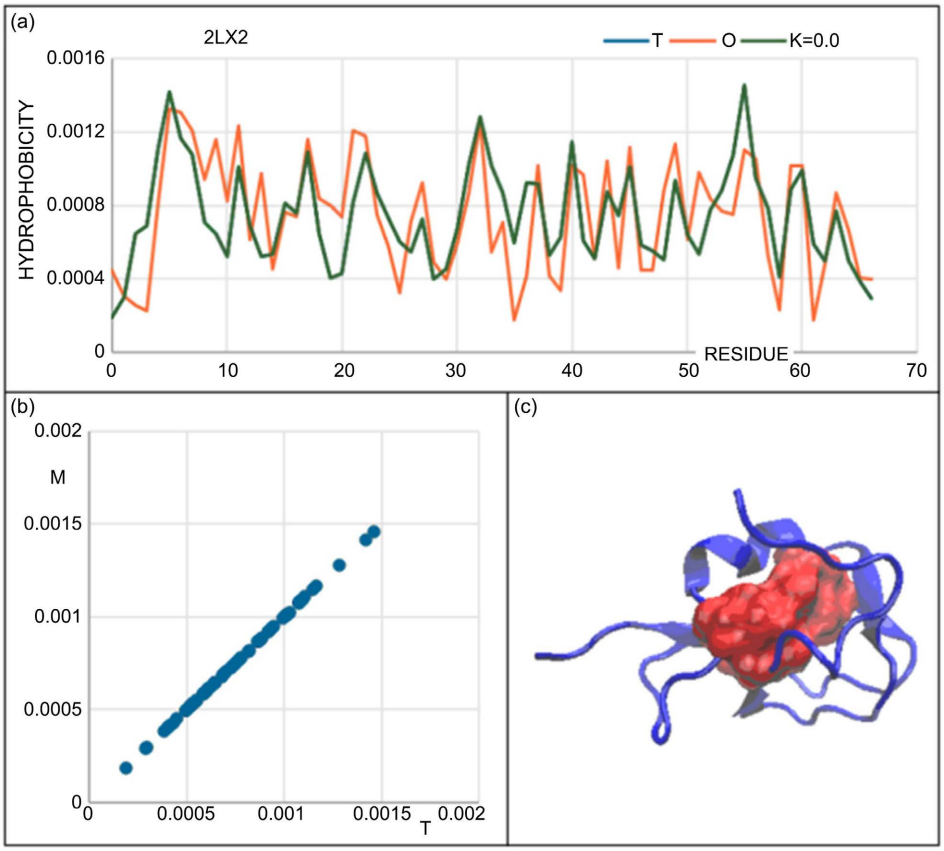


Figure 2. Characteristics of the antifreeze protein, type III. (a) The T distribution (blue), the $O$ distribution (red) and the $M$ distribution matching the $T$ distribution for $K = 0.0$; (b) Relation of value Mi to value $T_i$; (c) The 3D presentation with highlighted residues included in the hydrophobic nucleus—the position of residues, for which the local maxima for the $T$ and $O$ are identical (as shown in (a)).

The characteristic relation of value $M_i$ to $T_i$ (Figure 2(b)) matches the $M_i = T_i$ relation for the discussed protein. The position of residues representing high Oi level of hydropobicity according to theoretical expectations Ti are shown in red space filling representation to distinguish the hydrophobic core present in this protein (Figure 2(c)).

This enzyme is characterised by the increased value of $K = 0.6$ for the porcine pancreatic carbonoxypeptidase protein with PDB ID—2PJ6. The enzyme, as expected considering its activity, should be ready for specific interaction with the substrate. Consequently, in the enzyme structure the presence of cavity is expected, in which the $T$ and $O$ profiles are revealed by positions or chain fragments with a local hydrophobicity deficit. The analysis of $T$ and $O$ profiles (Figure 3(a))

clearly reveals the positions of residues located near the molecule centre (dominant $T_i$ maxima), for which the $O_i$ levels are significantly decreased. The location of these residues is shown in Figure 3(c)—with residues representing the deficit as pink. Due to the difference of the $O$ profile from the $T$ profile, the parameter value is $K = 0.6$. The force field variation is shown in Figure 3(b), where the dependency characteristic of the aqueous environment (angle of inclination in relation to the x-axis = 45 deg) changes to reveal modification of the field, to which the structure of the discussed protein adjusted. The placement of catalytic residues exactly within the cavity, where the catalysis occurs, is not accidental either. The catalytic reaction requires non-water environment to run.

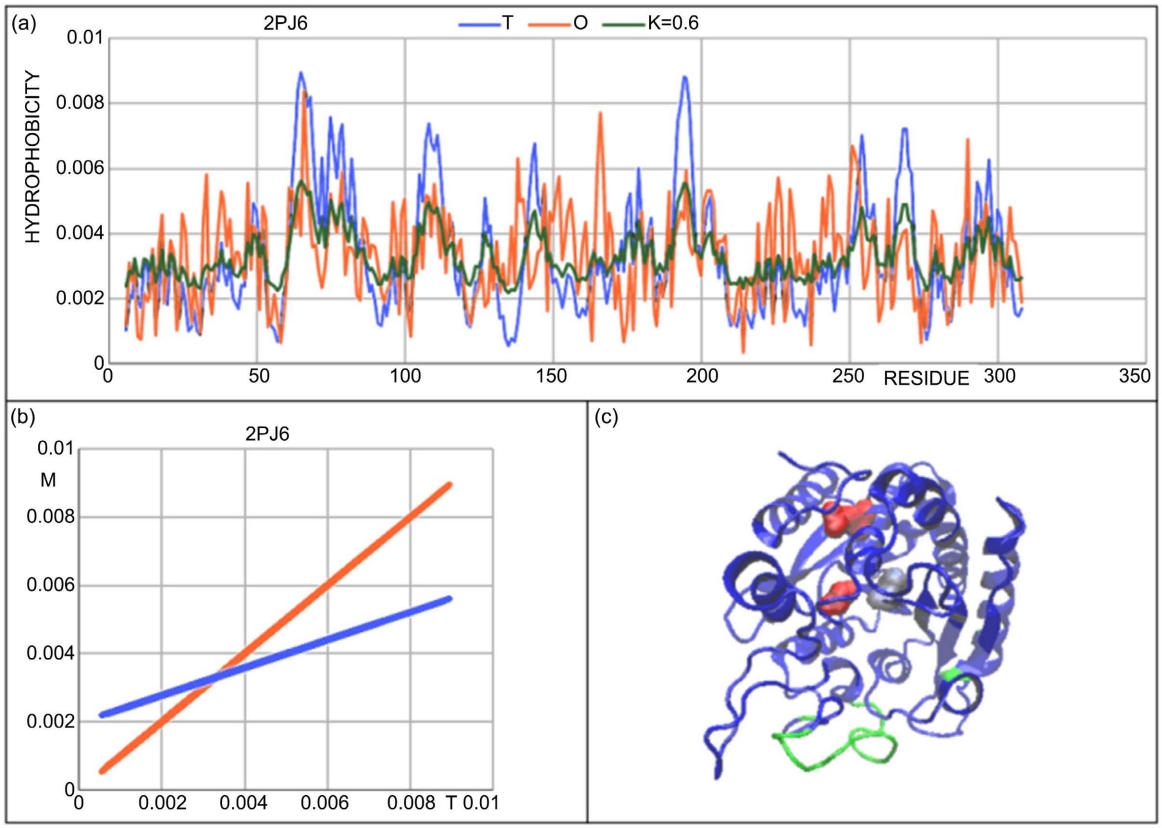


**Figure 3.** Characteristics of example enzyme. (a) The $T$ distribution (blue), the $O$ distribution (red) and the $M$ distribution for the $K$ value given in the legend. Marked with catalytic residue positions—vertical red line in position of catalytic residue; (b) Relation of value $M_i$ to value $T_i$; (c) 3D presentation with highlighted residues showing the deficit of hydrophobicity (pink). Catalytic residue red spacefilling.

### 3.3. Membrane Proteins—$K$ Values in the Range of $0.6 < K < 1.5$

For stability, membrane proteins in the hydrophobic environment of the membrane expose hydrophobic residues with simultaneous concentration of polar residues in the centre part. This is exemplified by rhodopsin (PDB ID—3QDC). The positions and even sections showing a deficit of hydrophobicity are revealed by the set of $T$ and $O$ profiles with clear positions of the deficit and local excess of hydrophobicity. The centring of polar residues is connected with the function of molecule transport

to the inside of the cell. This process is based on electrostatic interactions. Consequently, the concentration of polar residues in the central channel area. The increase of the parameter value of $K = 1.3$ changes the dependency relation of $M_i$ on $T_i$ in relation to the assembly present in aqueous environment (Figure 4(c)).

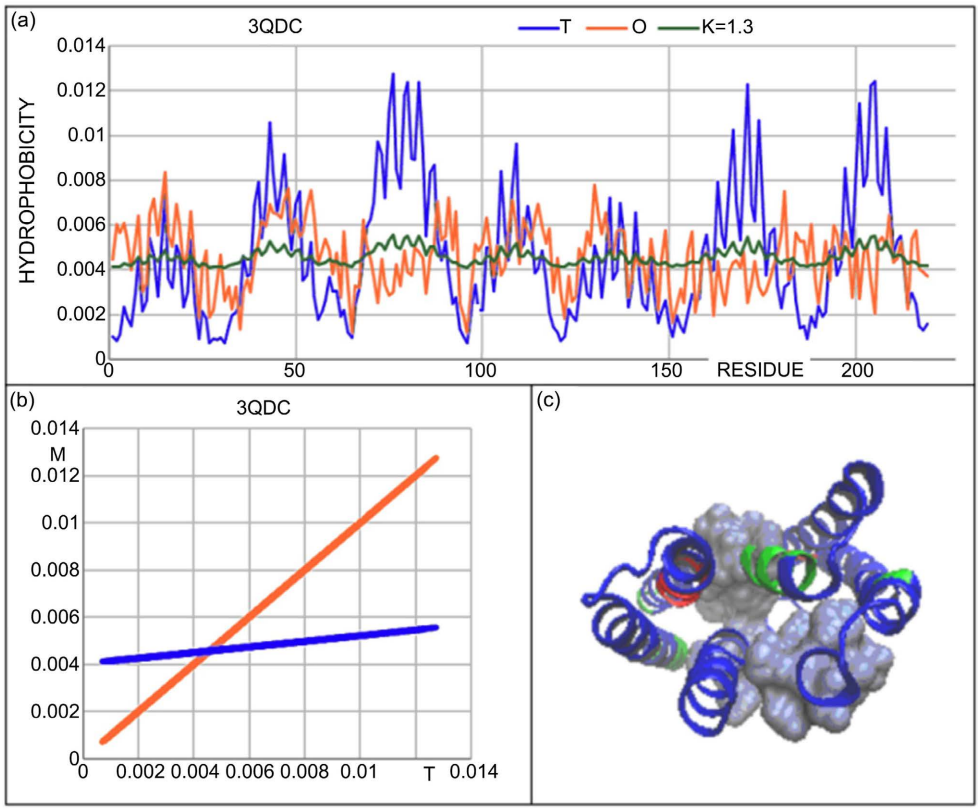


**Figure 4.** Characteristics of the membrane protein—rhodopsin. (a) The $T$ distribution (blue), the $O$ distribution (red) and the $M$ distribution for the $K$ value given in the legend; (b) Relation of value $M_i$ to value $T_i$; (c) 3D presentation with highlighted residues indicating the deficit of hydrophobicity in the molecule centre.

The gradually increasing deviation of $M_i$ from $T_i$ with an increase of the parameter value of $K$ is visible (Figure 4(a) and Figure 4(b)).

### 3.4. Amyloid Protein—Significant Variation of the Parameter Value of *K* Resulting from Structural Transformation

A separate issue is proteins obtaining structures referred to as misfolded. They are proteins that undergo structural transformation (without mutation) leading to the forms that enable the formation of amyloid fibrils [29]. The process occurring under *in vivo* conditions consists in 3D structure (solid, globular) conversion to 2D flat form. This phenomenon also applies to pathological proteins that cause neurodegenerative diseases, like Parkinson's, Alzheimer's etc. [29]. The structures of these proteins are formed under physiological conditions for currently unknown reasons. The structures corresponding to amyloid forms are also obtained under laboratory conditions. This transformation results from the physical process in the form of

shaking (aqueous protein solution subject to shaking) with varied time. There is evidence that practically each protein could be transformed to amyloid fibrils condition in such conditions [29]. Significant changes expressed by characteristics based on the FOD-M model are given in Table 3. The proteins administered are transthyretin and VL (PDB ID—1DVQ) and VL domain of IgG (PDB ID—4BJL).

Table 3. Structural changes of globular proteins (NATIVE) to amyloid forms described using the FOD-M criteria.

| NATIVE | | | AMYLOID | | |
|---|---|---|---|---|---|
| PDB ID | RD | $K$ | $K$ | RD | PDB ID |
| 1DVQ | 0.563 | 0.5 | 1.7 | 0.717 | 6SDZ |
| 4BJL-VL | 0.555 | 0.5 | 1.2 | 0.773 | 6HUD |

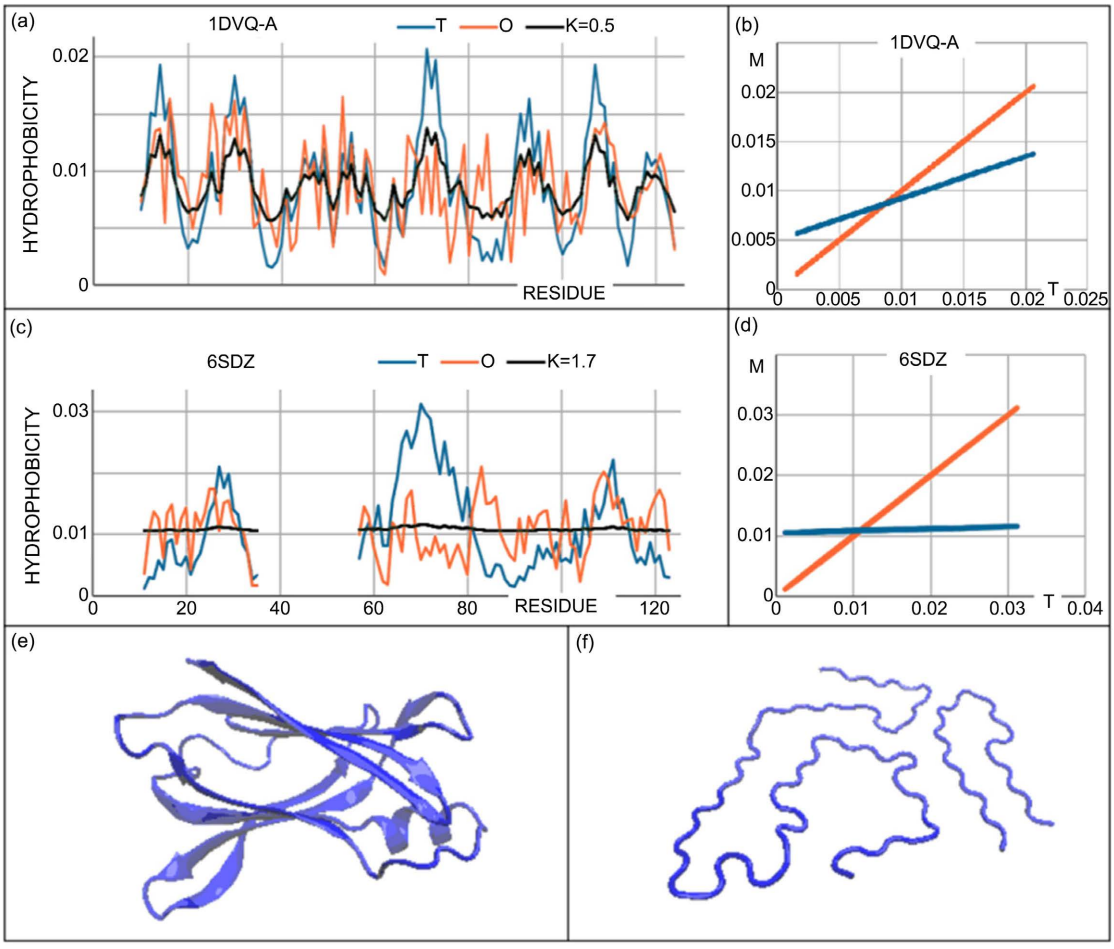


Figure 5. Native protein characteristics subject to amyloid transformation—transthyretin. (a) Profiles $T$ (blue), $O$ (red) and $M$ (black) for the $K$ value given in the legend—native form. (b) Profiles $T$ (blue), $O$ (red) and $M$ (black) for the $K$ value given in the legend—amyloid form; (c) $M_i$ in relation to $T_i$ for the native form—red-line representing the state with $K = 0.0$ (aqueous environment), line blue—relation for the native form of transthyretin; (d) $M_i$ in relation to $T_i$ for the native form—red-line representing the state with $K = 0.0$ (aqueous environment), line blue—relation for the amyloid form of transthyretin; (e) 3D presentation of native transthyretin; (f) 3D presentation of amyloid transthyretin.

A large change of parameters describing the hydrophobicity parameters after amyloid transformation (Table 3) is shown in Figure 5(a) and Figure 5(c). Another large change is shown in Figure 5(e) and Figure 5(f). The solid, globular structure after transformation changes to the 2D form—flat, enabling the construction of fibrils with an unlimited number of chains (stack of multiple chains one on another). Such structures, as an effect of misfolding, are hazards for the proper functioning of the organism. Internal factor variation is also illustrated in Figure 5(b) and Figure 5(d), where the relation of the $M_i$ distribution to $T_i$ assumes a near-constant form. The position of the $M_i$ line in relation to $T_i$ for the native form has a smaller directional coefficient. For the amyloid form this relation expressed by the directional coefficient is near-zero. This entails a significant variation of the environmental conditions inducing the structure observed in the amyloid form.

The given forms of proteins subject to amyloid transformation indicate the significance of environmental conditions. The given protein sequences do not differ—analysis is limited to chain parts common for both structural forms (however the original PDB files report the lengths given in Table 1). The structural change is drastic. Access to the proposed tool would enable simulation of the given transformation [59]. The detailed description of the procedure including the environmental characteristics leading to 2D form (flat) is given in [60].

### 3.5. Molecular Chaperones (Chaperonins and Chaperones)

Except for the extremely different water and membrane environments, an entire series of proteins, *i.e.* molecular chaperones accompanying the folding process. These are chaperones and chaperonins with very extensive structure—often in the form of large capsule. Within, this capsule creates conditions different from the aqueous environment, in which the folding process occurs. It is assumed that the parameter value of $K$ for the given chaperonin structure expresses the characteristics of the local environment. The value of $K$ for the chaperonin is considered under the FOD-M model as the measure of the variation of the local force field inside the capsule in relation to the aqueous environment.

An example of a very complex structure and a representative of chaperonins (molecular chaperones) in this analysis is the protein referred to as chaperonin type II [61]. This is a structure in the capsule form that forms space in the central part, in which the protein folding process occurs. This example is a generation model of a micro-space deviating from aqueous environment characteristics, which introduces its own rules depending on the chain composition (amino acid sequence) and capsule shape and size. Inside this structure, the structure formation process occurs according to the conditions therein. The characteristics of this protein require an extensive analysis, which herein is limited only to the presentation of the specific characteristics of the structure herein, to characterise the type of the external force field. The parameter values of RD and $K$ are on one hand characterised by status of the structure itself in relation to the aqueous envi-

ronment, but at the same time they define the conditions of the changed external force field for the protein folding in the interior representing the field. A broader analysis of other proteins in this group is available ([61], the large set of these proteins—in preparation).

The example of chaperonin discussed here represents Group II chaperonins ATP-dependent ring-shaped complexes (PDB ID—3IZI). The capsule comprises two symmetrically oriented caps (C2 symmetry), consisting of 8 chains each (C8 symmetry) (Figure 6(c)). The status of this capsule is expressed by very high RD and $K$ values (Table 4).

Table 4. Characteristics of chaperonin, type II.

| PDB ID | RD | $K$ | Chain length |
|---|---|---|---|
| 3IZI | 0.779 | 5.5 | 16 × 513 |

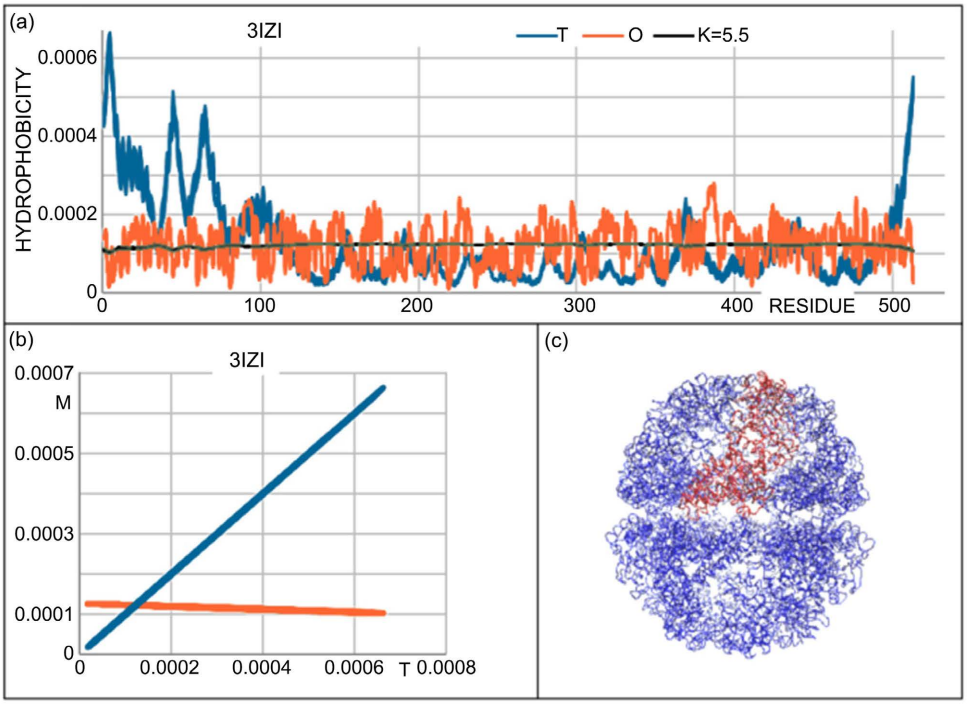


Figure 6. Characteristics of chaperonin, type II (PDB ID—3IZI). (a) Profiles $T$ (blue), $O$ (red) and $M$ (gray) for $K$ = 5.5; (b) $M_i$ relation to $T_i$ revealing a negative directional coefficient—red line. Blue line—the reference water conditions with $K$ = 0.0; (c) Presentation of the 3D capsule with chain A highlighted—red.

The RD and $K$ values describing the capsule status are very high. Profiles $T$ and $O$ generate profile $M$ close to the straight line parallel to x-axis. This entails the presence of organisation generating the force field in constant form of the environment without any direction from the perspective of the distribution of hydrophobicity. Profiles $T$, $O$ and $M$ for the given K reveal a complete omission of the rules imposed by the aqueous environment, while introducing a nearly uniform and constant form of profile $M$ (Figure 6(a)).

A detailed analysis of the relation of the M distribution to the $T$ distribution (Figure 6(b)) reveals, however, a negative direction coefficient value for this relation. It means that local minima in profile M are located at the points, at which the T distribution (*i.e.* aqueous environment) expects a local maximum of hydrophobicity. In such environment, a folding protein is adjusted to the environment with characteristics inverse of expected by the aqueous environment. This is what the introduction of an external force field (limited to the inside of the capsule) consists in for the folding process occurring inside the capsule. This is an interpretation arising from FOD-M model assumptions.

### 3.6. Mathematical Model—The Proposition to Replicate the Protein Folding Process by Introducing Environmental Conditions

The analysis of the examples given shows an entire spectrum of the potential for external force field involvement in the formation of a native structure with defined activity. The variation of environmental conditions (external force field) seems to be a mechanism of introduction of structural diversity of proteins in addition to the variation arising from the varied sequence records in proteins.

Table 2 reveals very significant variation of protein structures depending on the mismatch degree of the $O$ distribution to the $T$ distribution (parameter RD) and variation from the perspective of different environmental conditions of protein formation. The involvement of the external force field is expressed by the value of parameter $K$.

The summary of listed parameters proves the necessity to include the external force field in folding process simulations. This is proved by the variation of the degree of mismatch with the direction derived from the aqueous environment. It is also evidenced by the characteristics of the proteins of the chaperonin group, in which the force field generated inside the capsule has certain specific characteristics. The folding protein adjusts an appropriate distribution of hydrophobicity to a specifically formed local external force field (the inside of the capsule acts as such).

The proposed function subject to the optimisation process depends on two functions with opposite tendencies. The minimisation of one function (internal force field) leads to the increase of another (external force field). The Pareto front is a technique for finding consensus between the two opposite tendencies.

$$E = E[E_{\mathrm{INT}}\ (\text{electrostics, vdW, tp}),\ E_{\mathrm{EXT}}(K)] \tag{8}$$

where tp—torsional potential.

The $E_{\mathrm{INT}}$ notation is simplified and reduced to electrostatic and vdW interaction and the torsional potential connected with structural changes arising from the rotation in specific bindings (single bindings). The extensive force fields used in many packages (ROBETA [20], CHARMM [21] AMBER [22], GROMACS [23]) include a far greater number of structural variations affecting the energy state of the protein. These are energy contributions related to binding length variation

(stretching vibrations), change of flat angles between bindings (bending vibrations), changes of angles resulting in the spatial construction of flat assemblies (aromatic rings). This level of detail is not expected in the proposed model due to the significant simplification of representation of the structure at the stage of $E_{EXT}$ optimisation, in which effective atom positions are used (Equation (7)).

Due to the opposite effect, since the optimisation of one function results in an increase of the value arising from the other function, it is required to use the technique known as the Pareto front [24]-[28]. This method searches the consensus between the two opposite tendencies of variability of functions $E_{INT}$ and $E_{EXT}$.

The $E_{EXT}$ does not seem to be unified with $E_{INT}$ due to lack of experiences with simulation expressed by Equation (7). It is assumed to be possible after getting experiences in *In Silico* computations according to proposed model.

It is planned to perform certain iterations of the optimisation of $E_{INT}$ followed by certain iterations oriented on $E_{EXT}$. The very limited experience in this matter suggests the possible convergence of both procedures [19]. The more precise procedure is expected after extensive experience in this matter.

The $K$ parameter in currently applied program is found as the result of iteration procedure with increasing $K$ value as long as the minimum value of $D_{KL}(O|M)$ is found (**Figure 1(c)**). The range for $K$ value may be approximately estimated on the basis of RD parameter defined for certain structure (as temporarily delivered by the $E_{INT}$ optimisation procedure).

The current vision of the Pareto procedure applied to discussed issue is presented in details in [59].

For amyloid transformation simulation purposes, another factor is introduced (in addition to manipulation of the value of parameter $K$), *i.e.* step-wise reduction of the value of parameter SigmaZ. The purpose of this operation is to gradually change the 3D form to the 2D form. This change is also justified by the specific environmental characteristics during the experimental process of protein transformation from the native form to the amyloid form. Shaking, which is a physical process only, consists in introducing an increasing presence of the water/air interphase. The specific nature of this interphase consists in two-dimensionality (the contact surface is flat). In this interphase, water structuring presumably supports chain formation in the 2D form as well. Gradually reducing the value of parameter $\sigma_z$ to determine the form of the 3D Gaussian function, to which the folding protein adjusts (structural transformation) during the amyloid transformation is justified by the real condition of the experiment conducted.

### 3.7. Funnel Model in View of the FOD-M Model

Using an environment-dependent course of the protein folding process also provides some modification of the funnel model [62]-[64]. This model visualises the difficult task of selecting an appropriate energy minimum for the specific protein (**Figure 7(a)**). Due to the multiplicity of locally representing forms, combined with the global energy minimum, predicting the protein structure based on the

folding simulation process is virtually impossible to solve. The introduced dependency of environment-directed folding, as assumed based on the FOD-M model, will point towards the identification of the proper stable structure to ensure the expected biological activity for the given protein (Figure 7(b)).

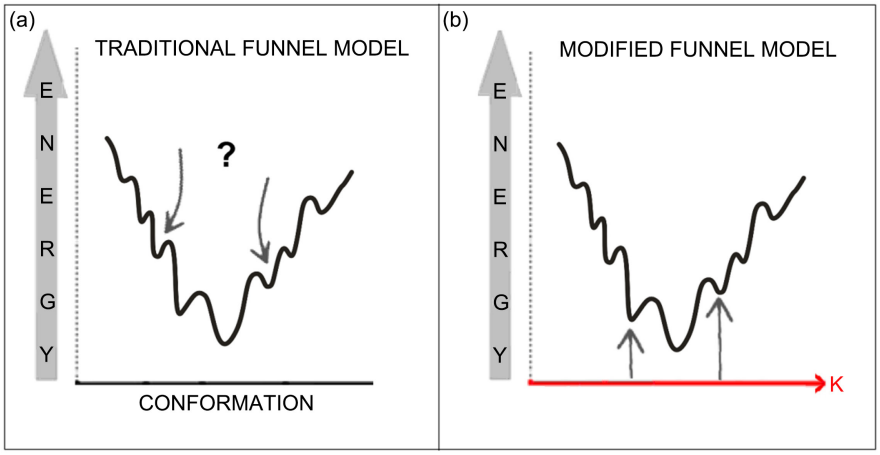


Figure 7. Funnel model. (a) Traditional form; (b) Modified based on the FOD-M model.

## 4. Discussion

The access to this tool based on multi-criteria optimisation, including the presence of two functions subject to optimisation in the optimisation procedure, is of great importance to the studies of protein function, which may change depending on the variation of environmental conditions. Here, a classical example is proteins defined as misfolding process effects. In this group, an important place is held by proteins of the amyloid group that cause neurodegenerative diseases, including Alzheimer's, Parkinson's and other [29]. Using the FOD-M model, it was demonstrated that an environmental factor was involved that dramatically changes the protein structure, thus completely eliminating the activity of these proteins [65].

The access to the tool including the presence of a variable environment would also allow addressing many phenomena, such as e.g. the impact of alcohol consumption. Alcohol, as an environmental conditions variation factor, may be analysed using the proposed tool including the involvement of the environment in the formation of structures serving a specific biological function. This factor has a very high potential for interaction, both with polar and hydrophobic components of proteins and other biological structures (e.g. membrane). This means easy and quick penetration to the inside of the cell. For alcohol, the entire environment of every living organism is available in a very short time. The impact of this factor is not radical and only visible at high concentrations. Low concentrations do not show the effects of presence, however they do reduce the activity of proteins. Not exactly properly folded protein could function seemingly properly. However, with one additional hazard (e.g. a COVID-type virus), the organism functioning at half capacity cannot handle such a condition, which is distant from full and proper activity of the organism. Not fully properly formed proteins under

the environmental disruption by the presence of alcohol cannot function properly in face of an additional hazard in the form of the virus [66]. The mortality rate of the COVID-19 pandemics was proved dependent on the alcohol consumption at the societal level [66].

Simulating the folding process in a changed environment could enable obtaining an adequately varied structure of the broadly understood misfolded proteins. The misfolded process is not limited to the formation of amyloid forms. Knowing the incorrect structure, it will be possible to predict the functional disorder of such protein. It will also become possible to design drugs to correct proteins in the misfolded form. The identification of the value of parameter $K$ will define the type of environmental disorder. This will create a chance to "cure" and correct the environment causing the misfolding.

At this point, an appeal could be expressed to develop a folding process simulation-based protein structure prediction tool to include the variation of environmental conditions. The FOD-M model offers such a solution. This appeal is addressed to computer science specialists specialised in software development. Such cooperation is expected.

## 5. Conclusion

The available protein folding process simulation tools, based on the optimisation of $E_{\text{INT}}$ (internal force field) only, fail to reflect the real conditions, in which the folding process occurs. The inclusion of the impact of the local environment directing the folding process is possible by introducing an additional module to also optimise the function expressing the environmental status, $E_{\text{EXT}}$ (external force field). The FOD-M model using the continuum function to express protein interaction with a specific environment expressed with parameter $K$ in the model enables broadening the representation of the force field to the environmental factor. Due to the opposite direction of the optimisation of $E_{\text{INT}}$ in relation to $E_{\text{EXT}}$ (reducing the value of function $E_{\text{INT}}$ increases the value of $E_{\text{EXT}}$), it is necessary to apply the procedure known as the Pareto front. The access to the tool enabling folding process simulations under varied environmental conditions would enable simulation of the protein transformation process with 3D structure to 2D form, which takes place in amyloid transformation. The use of a model including the impact of the local environment on the folding process would allow obtaining structures referred to as misfolded, not necessarily limited to amyloid forms. It would be possible to obtain protein structures e.g. with increased (varied) alcohol presence in the environment by showing its involvement in processes leading to pathological disorders, as postulated for COVID-19 [67].

## Acknowledgements

Many thanks to Anna Smietanska-Zaremba and Zdzislaw Wisniowski for technical support. This research was carried out within the MSHE project "Support for the activity of Centers of Excellence established in Poland under Horizon

2020" on the basis of contract number MEiN/2023/DIR/3796. This project has received funding from the EU's Horizon 2020 research and innovation programme under grant agreement No. 857533. This publication is supported by the Sano ("Sano"—the name of project—in Latin means "Health") project carried out within the International Research Agendas programme of the FNP, co-financed by the EU under the European Regional Development Fund.

## Author Contributions

Conceptualisation, methodology, formal analysis, writing Irena Roterman, software Katarzyna Stapor, data curation Dawid Dulak, senior advisor Leszek Konieczny. All authors have read and agreed to the published version of the manuscript.

## Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

## References

[1] Abramson, J., Adler, J., Dunger, J., Evans, R., Green, T., Pritzel, A., *et al.* (2024) Accurate Structure Prediction of Biomolecular Interactions with *α*Fold 3. *Nature*, **630**, 493-500. https://doi.org/10.1038/s41586-024-07487-w

[2] https://www.rcsb.org/

[3] Fakhoury, Z., Sosso, G.C. and Habershon, S. (2023) Generating Protein Folding Trajectories Using Contact-Map-Driven Directed Walks. *Journal of Chemical Information and Modeling*, **63**, 2181-2195. https://doi.org/10.1021/acs.jcim.3c00023

[4] Rawat, S., Subramaniam, K., Subramanian, S.K., Subbarayan, S., Dhanabalan, S., Chidambaram, S.K.M., *et al.* (2024) Drug Repositioning Using Computer-Aided Drug Design (CADD). *Current Pharmaceutical Biotechnology*, **25**, 301-312. https://doi.org/10.2174/1389201024666230821103601

[5] Cho, H.S., Schotte, F., Stadnytskyi, V. and Anfinrud, P. (2021) Time-Resolved X-Ray Scattering Studies of Proteins. *Current Opinion in Structural Biology*, **70**, 99-107. https://doi.org/10.1016/j.sbi.2021.05.002

[6] Agarwal, U., Tonk, R.K. and Paliwal, S. (2025) Importance of Computer-Aided Drug Design in Modern Pharmaceutical Research. *Current Drug Discovery Technologies*, **22**, e15701638361318. https://doi.org/10.2174/0115701638361318241230073123

[7] Momin, Y. and Beloshe, V. (2025) Pharmacophore Modeling in Drug Design. *Advances in Pharmacology*, **103**, 313-324. https://doi.org/10.1016/bs.apha.2025.01.010

[8] Tur Razia, I., Kanwal, A., Riaz, H.F., Malik, A., Ahsan, M., Saleem Khan, M., *et al.* (2023) Recent Trends in Computer-Aided Drug Design for Anti-Cancer Drug Discovery. *Current Topics in Medicinal Chemistry*, **23**, 2844-2862. https://doi.org/10.2174/0115680266258467231107102643

[9] Takaya, D. (2024) Computer-Aided Drug Design Using the Fragment Molecular Orbital Method: Current Status and Future Applications for SBDD. *Chemical and Pharmaceutical Bulletin*, **72**, 781-786. https://doi.org/10.1248/cpb.c23-00839

[10] Acosta-Alvear, D., Harnoss, J.M., Walter, P. and Ashkenazi, A. (2024) Homeostasis Control in Health and Disease by the Unfolded Protein Response. *Nature Reviews Molecular Cell Biology*, **26**, 193-212. https://doi.org/10.1038/s41580-024-00794-0

[11] Pontisso, I., Ornelas-Guevara, R., Combettes, L. and Dupont, G. (2023) A Journey in UPR Modelling. *Biology of the Cell*, **115**, e2200111. https://doi.org/10.1111/boc.202200111

[12] Shpilka, T. and Haynes, C.M. (2017) The Mitochondrial UPR: Mechanisms, Physiological Functions and Implications in Ageing. *Nature Reviews Molecular Cell Biology*, **19**, 109-120. https://doi.org/10.1038/nrm.2017.110

[13] Kumar, M., Sharma, S. and Mazumder, S. (2023) Role of UPR$^{mt}$ and Mitochondrial Dynamics in Host Immunity: It Takes Two to Tango. *Frontiers in Cellular and Infection Microbiology*, **13**, Article 1135203. https://doi.org/10.3389/fcimb.2023.1135203

[14] Purisima, E.O. and Scheraga, H.A. (1987) An Approach to the Multiple-Minima Problem in Protein Folding by Relaxing Dimensionality. *Journal of Molecular Biology*, **196**, 697-709. https://doi.org/10.1016/0022-2836(87)90041-6

[15] Li, Z. and Scheraga, H.A. (1987) Monte Carlo-Minimization Approach to the Multiple-Minima Problem in Protein Folding. *Proceedings of the National Academy of Sciences of the United States of America*, **84**, 6611-6615. https://doi.org/10.1073/pnas.84.19.6611

[16] https://predictioncenter.org/

[17] Fadini, A., Studer, G. and Read, R.J. (2026) Model Quality Assessment for Casp16. *Proteins*: *Structure*, *Function*, *and Bioinformatics*, **94**, 302-313. https://doi.org/10.1002/prot.70037

[18] Roterman, I., Stapor, K. and Konieczny, L. (2023) Role of Environmental Specificity in CASP Results. *BMC Bioinformatics*, **24**, Article No. 425. https://doi.org/10.1186/s12859-023-05559-8

[19] Gadzała, M., Kalinowska, B., Banach, M., Konieczny, L. and Roterman, I. (2017) Determining Protein Similarity by Comparing Hydrophobic Core Structure. *Heliyon*, **3**, e00235. https://doi.org/10.1016/j.heliyon.2017.e00235

[20] Roterman, I., Sieradzan, A., Stapor, K., Fabian, P., Wesołowski, P. and Konieczny, L. (2022) On the Need to Introduce Environmental Characteristics in Ab Initio Protein Structure Prediction Using a Coarse-Grained UNRES Force Field. *Journal of Molecular Graphics and Modelling*, **114**, Article ID: 108166. https://doi.org/10.1016/j.jmgm.2022.108166

[21] Kim, D.E., Chivian, D. and Baker, D. (2004) Protein Structure Prediction and Analysis Using the Robetta Server. *Nucleic Acids Research*, **32**, W526-W531. https://doi.org/10.1093/nar/gkh468

[22] Hwang, W., Austin, S.L., Blondel, A., Boittier, E.D., Boresch, S., Buck, M., *et al.* (2024) CHARMM at 45: Enhancements in Accessibility, Functionality, and Speed. *The Journal of Physical Chemistry B*, **128**, 9976-10042. https://doi.org/10.1021/acs.jpcb.4c04100

[23] Wang, J., Wolf, R.M., Caldwell, J.W., Kollman, P.A. and Case, D.A. (2004) Development and Testing of a General Amber Force Field. *Journal of Computational Chemistry*, **25**, 1157-1174. https://doi.org/10.1002/jcc.20035

[24] Van Der Spoel, D., Lindahl, E., Hess, B., Groenhof, G., Mark, A.E. and Berendsen, H.J.C. (2005) GROMACS: Fast, Flexible, and Free. *Journal of Computational Chemistry*, **26**, 1701-1718. https://doi.org/10.1002/jcc.20291

[25] Adham, A.M., Mohd-Ghazali, N. and Ahmad, R. (2015) Performance Optimization of a Microchannel Heat Sink Using the Improved Strength Pareto Evolutionary Algorithm (SPEA2). *Journal of Engineering Thermophysics*, **24**, 86-100. https://doi.org/10.1134/s1810232815010087

[26] Deb, K. and Jain, H. (2014) An Evolutionary Many-Objective Optimization Algo-

rithm Using Reference-Point-Based Nondominated Sorting Approach, Part I: Solving Problems with Box Constraints. *IEEE Transactions on Evolutionary Computation*, **18**, 577-601. https://doi.org/10.1109/tevc.2013.2281535

[27] Ishibuchi, H., Imada, R., Setoguchi, Y. and Nojima, Y. (2016) Performance Comparison of NSGA-II and NSGA-III on Various Many-Objective Test Problems. 2016 *IEEE Congress on Evolutionary Computation* (*CEC*), Vancouver, 24-29 July 2016, 3045-3052. https://doi.org/10.1109/cec.2016.7744174

[28] Zitzler, E., Laumanns, M., and Thiele, L. (2001) SPEA2: Improving the Strength Pareto Evolutionary Algorithm. *Evolutionary Methods for Design Optimization and Control with Applications to Industrial Problems*, Greece, 19-21 September 2001, 95-100. https://sop.tik.ee.ethz.ch/publicationListFiles/zlt2001a.pdf

[29] Zitzler, E. and Thiele, L. (1999) Multiobjective Evolutionary Algorithms: A Comparative Case Study and the Strength Pareto Approach. *IEEE Transactions on Evolutionary Computation*, **3**, 257-271. https://doi.org/10.1109/4235.797969

[30] Chiti, F. and Dobson, C.M. (2017) Protein Misfolding, Amyloid Formation, and Human Disease: A Summary of Progress over the Last Decade. *Annual Review of Biochemistry*, **86**, 27-68. https://doi.org/10.1146/annurev-biochem-061516-045115

[31] Kumeta, H., Ogura, K., Nishimiya, Y., Miura, A., Inagaki, F. and Tsuda, S. (2013) NMR Structure Note: A Defective Isoform and Its Activity-Improved Variant of a Type III Antifreeze Protein from Zoarces Elongates Kner. *Journal of Biomolecular NMR*, **55**, 225-230. https://doi.org/10.1007/s10858-012-9703-9

[32] Neuweiler, H., Sharpe, T.D., Rutherford, T.J., Johnson, C.M., Allen, M.D., Ferguson, N., *et al.* (2009) The Folding Mechanism of BBL: Plasticity of Transition-State Structure Observed within an Ultrafast Folding Protein Family. *Journal of Molecular Biology*, **390**, 1060-1073. https://doi.org/10.1016/j.jmb.2009.05.011

[33] Zhang, W., Zhang, J., Zhang, X., Xu, C. and Tu, X. (2011) Solution Structure of Rap1 BRCT Domain from Saccharomyces Cerevisiae Reveals a Novel Fold. *Biochemical and Biophysical Research Communications*, **404**, 1055-1059. https://doi.org/10.1016/j.bbrc.2010.12.109

[34] Sborgi, L., Verma, A., Muñoz, V. and de Alba, E. (2011) Revisiting the NMR Structure of the Ultrafast Downhill Folding Protein GPW from Bacteriophage $\lambda$. *PLOS ONE*, **6**, e26409. https://doi.org/10.1371/journal.pone.0026409

[35] Ellis, M.J., Antonyuk, S.V., Strange, R.W., Sawers, G., Eady, R.R. and Hasnain, S.S. (2004) Observation of an Unprecedented Cu Bis-His Site: Crystal Structure of the H129V Mutant of Nitrite Reductase. *Inorganic Chemistry*, **43**, 7591-7593. https://doi.org/10.1021/ic048966p

[36] Robien, M.A., Clore, G.M., Omichinski, J.G., Perham, R.N., Appella, E., Sakaguchi, K., *et al.* (1992) Three-Dimensional Solution Structure of the E3-Binding Domain of the Dihydrolipoamide Succinyltransferase Core from the 2-Oxoglutarate Dehydrogenase Multienzyme Complex of *Escherichia coli*. *Biochemistry*, **31**, 3463-3471. https://doi.org/10.1021/bi00128a021

[37] Adler, M., Buckman, B., Bryant, J., Chang, Z., Chu, K., Emayan, K., *et al.* (2008) Structures of Potent Selective Peptide Mimetics Bound to Carboxypeptidase B. *Acta Crystallographica Section D Biological Crystallography*, **64**, 149-157. https://doi.org/10.1107/s0907444907057228

[38] Park, S.H., Das, B.B., Casagrande, F., Tian, Y., Nothnagel, H.J., Chu, M., *et al.* (2012) Structure of the Chemokine Receptor CXCR1 in Phospholipid Bilayers. *Nature*, **491**, 779-783. https://doi.org/10.1038/nature11580

[39] Toyoda, Y., Morimoto, K., Suno, R., Horita, S., Yamashita, K., Hirata, K., *et al.* (2018)

Ligand Binding to Human Prostaglandin E Receptor EP4 at the Lipid-Bilayer Interface. *Nature Chemical Biology*, **15**, 18-26. https://doi.org/10.1038/s41589-018-0131-3

[40] Fezoui, Y., Connolly, P.J. and Osterhout, J.J. (1997) Solution Structure of *αtα*, a Helical Hairpin Peptide of De Novo Design. *Protein Science*, **6**, 1869-1877. https://doi.org/10.1002/pro.5560060907

[41] Luecke, H., Schobert, B., Cartailler, J., Richter, H., Rosengarth, A., Needleman, R., *et al.* (2000) Coupling Photoisomerization of Retinal to Directional Transport in Bacteriorhodopsin. *Journal of Molecular Biology*, **300**, 1237-1255. https://doi.org/10.1006/jmbi.2000.3884

[42] Gushchin, I., Reshetnyak, A., Borshchevskiy, V., Ishchenko, A., Round, E., Grudinin, S., *et al.* (2011) Active State of Sensory Rhodopsin II: Structural Determinants for Signal Transfer and Proton Pumping. *Journal of Molecular Biology*, **412**, 591-600. https://doi.org/10.1016/j.jmb.2011.07.022

[43] Selvakumar, P., Fernández-Mariño, A.I., Khanra, N., He, C., Paquette, A.J., Wang, B., *et al.* (2022) Structures of the T Cell Potassium Channel Kv1.3 with Immunoglobulin Modulators. *Nature Communications*, **13**, Article No. 3854. https://doi.org/10.1038/s41467-022-31285-5

[44] Sacchettini, J.C., Klabunde, T., Petrassi, H.M., Oza, V.B., Raman, P. and Kelly, J.W. (2000) Rational Design of Potent Human Transthyretin Amyloid Disease Inhibitors. *Nature Structural Biology*, **7**, 312-321. https://doi.org/10.1038/74082

[45] Huang, D.B., Ainsworth, C.F., Stevens, F.J. and Schiffer, M. (1996) Three Quaternary Structures for a Single Protein. *Proceedings of the National Academy of Sciences of the United States of America*, **93**, 7017-7021. https://doi.org/10.1073/pnas.93.14.7017

[46] Schmidt, M., Wiese, S., Adak, V., Engler, J., Agarwal, S., Fritz, G., *et al.* (2019) Cryo-EM Structure of a Transthyretin-Derived Amyloid Fibril from a Patient with Hereditary ATTR Amyloidosis. *Nature Communications*, **10**, Article No. 5008. https://doi.org/10.1038/s41467-019-13038-z

[47] Swuec, P., Lavatelli, F., Tasaki, M., Paissoni, C., Rognoni, P., Maritan, M., *et al.* (2019) Cryo-EM Structure of Cardiac Amyloid Fibrils from an Immunoglobulin Light Chain AL Amyloidosis Patient. *Nature Communications*, **10**, Article No. 1269. https://doi.org/10.1038/s41467-019-09133-w

[48] Douglas, N.R., Reissmann, S., Zhang, J., Chen, B., Jakana, J., Kumar, R., *et al.* (2011) Dual Action of ATP Hydrolysis Couples Lid Closure to Substrate Release into the Group II Chaperonin Chamber. *Cell*, **144**, 240-252. https://doi.org/10.1016/j.cell.2010.12.017

[49] Liao, Z., Gopalasingam, C.C., Kameya, M., Gerle, C., Shigematsu, H., Ishii, M., *et al.* (2024) Structural Insights into Thermophilic Chaperonin Complexes. *Structure*, **32**, 679-689.e4. https://doi.org/10.1016/j.str.2024.02.012

[50] Burton, R.E., Huang, G.S., Daugherty, M.A., Fullbright, P.W. and Oas, T.G. (1996) Microsecond Protein Folding through a Compact Transition State. *Journal of Molecular Biology*, **263**, 311-322. https://doi.org/10.1006/jmbi.1996.0577

[51] Roterman, I., Banach, M. and Konieczny, L. (2017) Antifreeze Proteins. *Bioinformation*, **13**, 400-401. https://doi.org/10.6026/97320630013400

[52] Roterman, I., Stapor, K. and Konieczny, L. (2024) Transmembrane Proteins—Different Anchoring Systems. *Proteins: Structure, Function, and Bioinformatics*, **92**, 593-609. https://doi.org/10.1002/prot.26646

[53] Roterman, I., Konieczny, L. and Stapor, K. (2026) Environment-Dependent Struc-

tural Factors. In: Roterman-Konieczna, I., Ed., *Simulating Protein Folding in Variable Environmental Conditions*, Elsevier, 47-83. https://doi.org/10.1016/b978-0-443-40475-7.00003-1

[54] Levitt, M. (1976) A Simplified Representation of Protein Conformations for Rapid Simulation of Protein Folding. *Journal of Molecular Biology*, **104**, 59-107. https://doi.org/10.1016/0022-2836(76)90004-8

[55] Kullback, S. and Leibler, R.A. (1951) On Information and Sufficiency. *The Annals of Mathematical Statistics*, **22**, 79-86. https://doi.org/10.1214/aoms/1177729694

[56] Banach, M., Konieczny, L. and Roterman, I. (2020) Protein-Protein Interaction Encoded as an Exposure of Hydrophobic Residues on the Surface. In: Roterman-Konieczna, I., Ed., *From Globular Proteins to Amyloids*, Elsevier, 79-89. https://doi.org/10.1016/b978-0-08-102981-7.00008-7

[57] Roterman, I., Konieczny, L., Stapor, K. and Słupina, M. (2024) Hydrophobicity-Based Force Field in Enzymes. *ACS Omega*, **9**, 8188-8203. https://doi.org/10.1021/acsomega.3c08728

[58] Roterman, I. and Konieczny, L. (2023) Protein Is an Intelligent Micelle. *Entropy*, **25**, Article 850. https://doi.org/10.3390/e25060850

[59] Humphrey, W., Dalke, A. and Schulten, K. (1996) VMD: Visual Molecular Dynamics. *Journal of Molecular Graphics*, **14**, 33-38. https://doi.org/10.1016/0263-7855(96)00018-5

[60] Roterman, I., Konieczny, L., Dulak, D. and Stapor, K. (2026) Proposed Computational Procedure for Protein and Amyloid Structure Prediction. In: Roterman-Konieczna, I., Ed., *Simulating Protein Folding in Variable Environmental Conditions*, Elsevier, 107-130.

[61] Banach, M. and Roterman, I. (2020) Amyloid as a Ribbon-Like Micelle. In: Roterman-Konieczna, I., Ed., *From Globular Proteins to Amyloids*, Elsevier, 177-191.

[62] Roterman, I., Stapor, K., Dułak, D. and Konieczny, L. (2024) External Force Field for Protein Folding in Chaperonins-Potential Application in *in Silico* Protein Folding. *ACS Omega*, **9**, 18412-18428. https://doi.org/10.1021/acsomega.4c00409

[63] Roterman, I., Slupina, M. and Konieczny, L. (2024) Protein Folding: Funnel Model Revised. *Computational and Structural Biotechnology Journal*, **23**, 3827-3838. https://doi.org/10.1016/j.csbj.2024.10.030

[64] Reynaud, E. (2010) Protein Misfolding and Degenerative Diseases. *Nature Education*, **3**, Article 2. https://www.nature.com/scitable/topicpage/protein-misfolding-and-degenerative-diseases-14434929/

[65] Huang, Z., Cui, X., Xia, Y., Zhao, K. and Zhang, G. (2023) Pathfinder: Protein Folding Pathway Prediction Based on Conformational Sampling. *PLOS Computational Biology*, **19**, e1011438. https://doi.org/10.1371/journal.pcbi.1011438

[66] Roterman, I., Słupina, M., Dułak, D. and Konieczny, L. (2025) Three Scenarios for Amyloid Transformation in the Context of the Funnel Model. *Computational and Structural Biotechnology Journal*, **27**, 1648-1659. https://doi.org/10.1016/j.csbj.2025.04.016

[67] Roterman, I., Stapor, K., Fabian, P. and Konieczny, L. (2022) *In Silico* Modeling of COVID-19 Pandemic Course Differentiation Using the FOD Model. *Coronaviruses*, **3**, e020622205580. https://doi.org/10.2174/2666796703666220602155421