J Mol Model (2015) 21: 151
DOI 10.1007/s00894-015-2691-4

ORIGINAL PAPER

Comparative modeling of hypothetical amyloid pores based on cylindrin

Magdalena Zulpo[1] · **Malgorzata Kotulska**[1]

Received: 29 October 2014 / Accepted: 27 April 2015 / Published online: 21 May 2015

Abstract Cylindrin is a six-stranded antiparallel beta barrel obtained from amyloidogenic strands of crystallin. It induces cell toxicity through an unknown mechanism. In this work, the potential use of the structure of cylindrin as a template for modeling amyloid pores—hypothetical transmembrane structures which appear during amyloid diseases—was studied. Using comparative modeling (performed by Modeller), we tested the stability of cylindrin-based pores made from several amyloid-forming and non-amyloid-forming strands deriving from mutated cylindrin and the prion sup35. We showed that cylindrin could be used as a template for modeling pores made from strands of amyloid proteins, but that the cylindrin structure does not result from the amyloidogenicity of these fragments, as fibril non-formers from the prion were also able to form a similar structure. Finally, we tested whether the cellular toxicity of cylindrin and related structures could be due to its incorporation into the cell membrane, leading to the creation of conducting ionic channels. The results of modeling indicate that cylindrin and tandem-repeat cylindrin, mutants of them, and cylindrin-like amyloid pores from prion sequences can only localize on the periphery of the membrane, and are not able to conduct any ions into the cell. These findings explain experimental results obtained for large unilamellar vesicles incubated with cylindrin, where conductance was not observed.

Keywords Amyloid · Amyloid channel · Cylindrin · Structure modeling · Comparative modeling

This paper belongs to Topical Collection 6th Conference on Modeling & Design of Molecular Materials in Kudowa Zdrój (MDMM 2014)

Electronic supplementary material The online version of this article (doi:10.1007/s00894-015-2691-4) contains supplementary material, which is available to authorized users.

✉ Malgorzata Kotulska
Malgorzata.Kotulska@pwr.edu.pl

[1] Department of Biomedical Engineering, Wroclaw University of Technology, ul. Wybrzeze Wyspianskiego 27, 50-370 Wroclaw, Poland

Introduction

In recent years, particularly in developed countries, the occurrence rates of amyloid-based neurodegenerative diseases have increased significantly. Such diseases include Alzheimer's disease (AD), Parkinson's disease, amyotrophic lateral sclerosis (ALS), Huntington's disease, frontotemporal dementia, amyloid polyneuropathy, prion diseases (e.g., Creutzfeldt–Jacob disease), and others [1]. However, amyloids play a role not only in neurodegenerative diseases but also in other conditions, for example in diabetes type 2. These disorders are characterized by an accumulation of protein deposits in the tissue, and they are regarded as civilization diseases. The mechanisms of these diseases are not fully understood. Although studies indicate that amyloid diseases are linked to genetics, the influence of lifestyle cannot be excluded. There is a correlation between amyloidosis and age, which may help explain the high incidence of this disease in countries with a high life expectancy.

In general, amyloid proteins primarily exhibit a non-amyloid structure which is involved in physiological processes that are crucial to cell functioning. The aggregation of proteins and their fragments does not necessarily lead to amyloid formation, and protein aggregation can be promoted by various factors, such as a high protein concentration, a high temperature, a low pH, and binding to metals. Protein aggregation can be a reversible process unless it concerns amyloids. Amyloid beta-sheet aggregates have a specific, highly ordered

structural motif called a steric zipper. This distinguishes them from other protein aggregates of the beta-sheet type, which are not as regular. Due to its structure, an amyloid is very durable and resistant to the activities of proteolytic enzymes, and cannot be easily dissolved. The distinctive structure of a steric zipper permits the selective detection of amyloids using a variety of microscopic techniques or fluorescent probes (such as thioflavin T) with which they form compounds. The tertiary structures of amyloid proteins are very similar, regardless of the degree of similarity between the amino acid sequences.

It is currently believed that short peptide sequences with amyloidogenic properties (called "hotspots") may be responsible for the aggregation of amyloid proteins. These fragments, which are 4–10 residues long (typically hexapeptides), have a high propensity for strong interactions that lead to protein aggregation [2]. Previous studies have suggested that amyloidogenic fragments may show consistency in terms of both the average physicochemical properties of their amino acids and their amino acid sequences. There have been attempts to predict the sequences of such peptides by computational modeling using physics- and chemistry-based models, including FoldAmyloid [3]. This method focuses on the density of protein-contact sites. In other methods, such as Zipper DB, the peptide is threaded onto an amyloid fiber backbone and then its energy and stability are determined [4–6]. Statistical methods include simple frequency profiles, such as the Waltz method [7]. Our group used another approach, employing machine learning methods; these were found to be highly efficient provided a sufficiently large and representative training dataset was available [8–10]. Based on those studies, we published a web tool, Fish Amyloid, which recognizes amyloidogenic sequences (http://www.comprec.pwr.wroc.pl). However, all such computational methods must overcome the issue of the rather small amount of experimental data available to aid their construction and validation.

Recently, it was shown that amyloid fibrils themselves are not cytotoxic to nerve cells and are not the main cause of neurodegenerative diseases [1, 11]. It was revealed that the processes associated with such diseases occur due to the cytotoxicity caused by oligomers, which are intermediate forms between a native protein and its mature amyloid fiber. These oligomers typically have greatly expanded amino acid chains that interact very easily to form numerous hydrogen bonds. It is now believed that this strong reactivity of oligomers underlies amyloid cytotoxicity, although a detailed mechanism has not yet been discovered.

There are several hypotheses that provide feasible mechanisms for the development of fully symptomatic neurodegenerative diseases. For example, some authors state that immunological response is crucial to disease development [11]. Recent studies have also indicated that neurodegenerative processes may correspond to the incorporation of amyloid oligomers into the cell and organelle membranes, creating weakly cation-selective ion channels that allow an uncontrolled influx of calcium into nerve cells [12, 13]. This influx arises due to the high calcium concentration gradient between the extracellular space and the cytoplasm. The excessive influx of calcium into the cytoplasm leads to the disruption of intracellular pathways, membrane depolarization, and ATP depletion following attempts to restore the transmembrane potential. Another observed effect is mitochondrial membrane depolarization and impairment of mitochondrial function connected to the calcium efflux and the release of cytochrome c. This hypothesis of oligomeric amyloid channels is justified by the very high concentration of intracellular calcium observed in amyloid diseases [14], as well as the effects of membrane depolarization and the lack of success of treatments that attempt to eliminate oligomers from cells while leaving the transmembrane peptides in the cell membrane, as they are much more difficult to remove [1]. Channel formation has been experimentally observed for Abeta, SOD1, alpha-synuclein, and prion proteins [1]. Electrophysiological studies of cell lines performed in vitro [15] and model planar lipid membranes [16, 17] have shown conductivity characteristics that are similar to those of amyloid deposits. Furthermore, microscopic studies in patients suffering from amyloid diseases showed the occurrence of unidentified excess transmembrane proteins. These structures were absent from the cells of healthy individuals.

Despite the high number of papers devoted to the formation of amyloid fibrils, the detailed mechanism of this process is still a mystery. It is known that the formation of fibrils immediately precedes the formation of metastable oligomers of various sizes [18, 19]. At the same time, it is not clear which of the resulting oligomers can serve as the nucleus for the subsequent growth of fibrils. The question of the sizes of protofibril (a single strand of a mature fibril) nuclei formed by different proteins and peptides is still the source of debate because of the absence of any firm knowledge of the mechanisms underlying amyloid formation. It is possible that variations in fibril morphology arise due to the presence of similar but non-identical ways of forming mature fibrils. Differences may be established at the beginning of fibril formation, including during the nucleation step. In a recently published study [20] of the process of Abeta peptide amyloid formation, it was shown that the nucleus may be a remodeled oligomer which varies from two to three monomers in size.

It is necessary to determine the structure of the hypothetical amyloid channels together with details of the fibril formation process, their interactions with lipids of the cell membrane, their functional characteristics, and the effects of the activity of amyloid channels on cells in order to understand conditions necessary for their existence and their connection to neurodegenerative diseases [21]. Knowledge of the features of amyloid channel structure could facilitate the development of methods of preventing neurodegenerative diseases.

In the work presented in this article, we attempted to establish whether the cellular toxicity associated with neurodegenerative diseases is mainly due to the incorporation of amyloid oligomers into the cell membrane and the subsequent creation of conducting ionic channels. To model these channels we used cylindrin—the only amyloid oligomer pore with an experimentally resolved structure. Cylindrin was artificially induced in cells by incorporating amyloidogenic peptide fragments from crystallin into a bacterial genome. Importantly, cylindrin activated the cellular cytotoxicity observed in amyloid diseases, although its mechanism was not resolved. In our study, we investigated whether cylindrin could be a good template for other oligomeric forms of amyloids. We then investigated whether cylindrin and structurally related proteins, including amyloidogenic sequences of aminoacids, exhibit a high propensity to form transmembrane channels that are capable of disrupting the membrane barrier and allowing excess ions to enter the cell.

Methods

Modeling structures

A full protein threading modeling method was used to obtain tentative amyloid pore structures. Threading uses template protein structures that are considered to be potentially similar to and thus a good first approximation of the target structure. This tentative model was improved in subsequent modeling steps. We used Modeller [22], a software tool designed for homology and comparative modeling. Comparative modeling allowed us to use a protein of known structure to generate a template for modeling a related protein with a known sequence and an unsolved structure. Further refinement of the model was performed based on spatial features of the template [23]. Modeller allows full protein threading in which sequence homology between the target sequence and the template protein is not necessary and the two molecules are similar in other aspects (e.g., they have a similar function). Our basic template structure was cylindrin, which spontaneously forms from amyloid fibrils derived from a sequence of αβ crystallin (UniProt ID: P02511 for α-crystallin B chain) [24]. Here, it was used to derive hypothetical amyloid pore structures.

Modeller uses a classical method based on the satisfaction of a spatial restraints algorithm [23, 22]. In this algorithm, template structures are compared to obtain spatial features: distances and dihedral angles. Next, template sequences are aligned with the target sequence. Finally, spatial restraints obtained from template structures are transferred to the target sequence and satisfied as accurately as possible [23, 22]. Spatial restraints in Modeller are calculated based on a statistical analysis of the relationships between similar protein structures [23]. Each spatial restraint is expressed as a probability density function. General probability density functions (pdfs) were obtained from dataset analysis [25]. Spatial restraints and CHARMM22 force-field terms enforcing proper stereochemistry were used. Both the CHARMM22 terms and the spatial restraints were combined into a molecular objective function (the molecular pdf) [23]. The molecular pdf was expressed as the negative natural logarithm of the sum of a quadratic function, the harmonic lower and upper bounds, a cosine function, the weighted sum of a few Gaussian functions, Coulomb's law, the Lennard–Jones potential, and cubic splines [23]. Geometric features included distances, angles, dihedral angles, pairs of dihedral angles (for two, three, four, and eight atoms), the shortest distance, solvent accessibility, and atom density (expressed as the number of atoms around the central atom) [23]. The most probable structure was associated with the minimum in the molecular pdf function. In the last modeling step, the molecular objective function was optimized by Modeller. Optimization was divided into two parts: conjugate gradient optimization in Cartesian space and then molecular dynamics (MD) with simulated annealing [22, 23]. The optimized value of the molecular pdf is called the molpdf. The lower the value of the molpdf, the greater the satisfaction of the spatial restraints (i.e., the better the fit).

Sequences

The cylindrin sequence is given in Table 1. This sequence was tested by Eisenberg and coworkers [24] and found to form a beta-barrel structure. All other sequences that may potentially form cylindrin-like structures and have been studied using molecular dynamics [26] are also given in Table 1. Hence, Table 1 includes cylindrin (K11V), cylindrin mutants ($K11V^{V2L}$, $K11V^{V4AV8A}$, $K11V^{V4GV8G}$, $K11V^{V4LV8L}$), as well as tandem-repeat cylindrin (K11V-TR) and its mutant ($K11V^{V4W}$-TR). The structures of K11V and K11V-TR were derived experimentally by Laganowsky and colleagues (PDB codes: 3SGO and 3SGR, respectively) [24]. These structures,

Table 1 Cylindrin (K11V), tandem repeat cylindrin (K11V-TR), and mutant sequences [24, 26]. Positions given in bold are mutated amino acids (for the cylindrin mutants) or added amino acids (for the tandem-repeat cylindrin)

Label	Sequence
K11V	KVKVLGDVIEV
$K11V^{V2L}$	K**L**KVLGDVIEV
$K11V^{V4AV8A}$	KVK**A**LGD**A**IEV
$K11V^{V4GV8G}$	KVK**G**LGD**G**IEV
$K11V^{V4LV8L}$	KVK**L**LGD**L**IEV
K11V-TR	**G**KVKVLGDVIEV**GG**KVKVLGD VIEV
$K11V^{V4W}$-TR	**G**KLKWLGDVIEV**GG**KLKWLGDVIEV

as well as that of K11V^{V4W}-TR, were experimentally tested for amyloidogenicity [24].

Tandem-repeat cylindrin was created from K11V strands, which were combined into pairs with additional glycine residues. There is also an additional glycine as the initial residue of each chain (Table 1) [24]. Similarly to K11V, K11V-TR creates highly toxic oligomers with a beta-barrel shape. However, the K11V mutant (K11V^{V4W}-TR) was almost nontoxic to mammalian cells [24].

The cylindrin mutants were derived by Berhanu and colleagues [26] using computational mutation of selected residues. The starting configuration was the wild-type cylindrin (PDB code: 3SGO). Mutants were created [26] by replacing selected residues with alanine, glycine, and leucine using VMD software, and were then refined with molecular dynamics.

All modeled sequences from prion protein sup35 are collected in Table 3 of the "Appendix," and were based on data published in [7]. Each sequence had ten residues and they were not identical. It was experimentally confirmed [7] that 12 of the fragments create fibrils, while the remaining 31 do not (Table 3 in the "Appendix"). We divided the fibril non-formers into two groups based on whether they were positioned in the middle or at the end of the sup35 sequence. For easier recognition, we denoted fibril-former sequences A1–A12, while fibril non-formers from the middle of sup35 were labeled B1–B13 and fibril non-formers from the end of sup35 were labeled C1–C18 (Table 3 in the "Appendix"). Using Modeller software, we created structural models based on the cylindrin structure (template) and the sequences of mutated cylindrin and sup35.

Generation of fibril pore models

In our study, we used the structure of cylindrin as a template for modeling unknown amyloid peptide structures deriving from sequences. We used two modeling modes: automodeling mode (employing Modeller's default parameters) and user mode (with user-defined symmetry restraints on Cα). Modeling with symmetry restraints on Cα was performed for four different weights: 0.2, 0.4, 0.6, and 1.0. Each sequence was modeled in automodeling mode and four user modes with different weight values. Ten possible structural models of our sequences were obtained from each modeling mode; that is, 50 models were generated for each sequence.

For each mode, Modeller required three types of input files: a pdb (Protein Data Bank) file with coordinates of the template structure, an alignment file defining the target sequence using a one letter code, and a script file with Modeller commands. As a template, we used the cylindrin sequence in the form of a biological assembly. 3SGO, a biological assembly, was therefore downloaded from PDB. Additionally, its individual chains were joined together to form a six-stranded structure within the same frame. The resulting file (3SGO_assembly.pdb) is provided in the "Electronic supplementary material," ESM. The Modeller script file was run on Linux (Mint Cinnamon 64-bit) via Python (version 2.7).

For each modeling mode, sequences of template and target proteins were provided in the alignment file (*.pir). The level of model refinement was defined in the Modeller script file as "very slow" for all modeling modes. Additionally, in the user mode, a new class was added to the Modeller script file (examples are available in the Modeller manual). This class allows the selection of objects (Cα from each chain here) and the addition of restraints. In the symmetry restraints, two groups of atoms and weights (from 0.2 to 1.0) were defined. In this case, the symmetry restraints were added to the corresponding chains of the target. For instance, Cα of chain A was symmetrical to Cα in B and C. A higher weight means a more rigid fit of the selected objects.

Outputs from Modeller are provided as model coordinate files (*.pdb and *_fit.pdb files) and a *.log output file. Output model files contain the coordinates of the created models (*.pdb). Created model coordinates were often shifted in comparison to the template (*_fit.pdb). Due to this spatial offset, additional output model coordinate and template files that were fitted to each other were created. This fitting facilitated further visual comparison of the generated models to the template structure (3SGO). Information about model parameters was collected in the log output Modeller file, which provided data on restraint violations. In our analysis, we considered calculated molpdf values and root mean square (RMS) deviations from the original 3SGO (Cα atoms and N–O atoms in the main chain).

Transmembrane propensity of all modeled amyloid pores

The transmembrane propensity of the pores was tested using the Positioning of Protein in Membrane (PPM) web server (http://opm.phar.umich.edu/server.php). PPM is based on information from the Orientation of Protein in Membrane (OPM) database [27]. Spatial parameters returned by the PPM web server are calculated based on the PDB file of a model molecule. The tilt angle, depth of membrane insertion, and transfer energy were calculated for each generated model.

The transfer energy (ΔG_{transf}) is defined as the energy required to transfer a protein from water to the DOPC bilayer. It is calculated as the sum of the solvent-accessible surface area (SASA)-dependent and electrostatic terms. The first term accounts for van der Waals and hydrogen bonding, interactions between the solute (protein) and solvent, and the entropy of solvent molecules in the first solvation shell [27]. The second term includes the solvation energy of dipoles and ions and the deionization penalty for ionizable groups in a nonpolar environment [27]. For integral membrane proteins and peptides,

ΔG_{transf} is usually between −400 and −10 kcal/mol. For peripheral proteins, ΔG_{transf} is usually between −15 and −1.5 kcal/mol. The root mean square deviation between experimental and calculated values (based on the OPM database) is about 0.74 kcal/mol for small proteins and 1.13 kcal/mol for peripheral proteins. The orientation of the protein in the bilayer is determined by minimizing the transfer energy. For small transmembrane proteins, the uncertainty in the penetration depth is about 1 Å, and that for the tilt angle is 5°. They are larger for peripheral proteins and may reach 3 Å for the penetration depth and 50° for the tilt angle [27].

Statistical analysis

Statistical analysis was performed to test the compatibility of the results obtained from Modeller with corresponding data derived from experiments [24] or MD studies [26]. It was also used to find differences between the models obtained for amyloid formers and non-formers from sup35 sequence fragments (Table 3 in the "Appendix").

The Kolmogorov–Smirnov and Wilcoxon statistical tests were performed at a confidence level of 0.95, using R software, on the output parameters from Modeller (RMSD Cα, RMSD N-O, molpdf) and the PPM web server (tilt, penetration depth, and ΔG_{transf}).

The Kolmogorov–Smirnov (R function: ks.test{}) test can distinguish between two populations based on their distributions. Hence, it is possible to conclude if two groups of samples come from populations described by the same distribution function. The Wilcoxon test is based on pseudo-median estimator calculations. Using the parameters listed above, a two-sample Wilcoxon test was performed (R function: wilcox.test{}). The two-sample Wilcoxon test estimates the median of the differences between pairs of samples from the first and the second populations. Results were visualized and preliminary analysis was performed with R box plots (ggplot2 package, function: geom_boxplot{}). The upper side of the

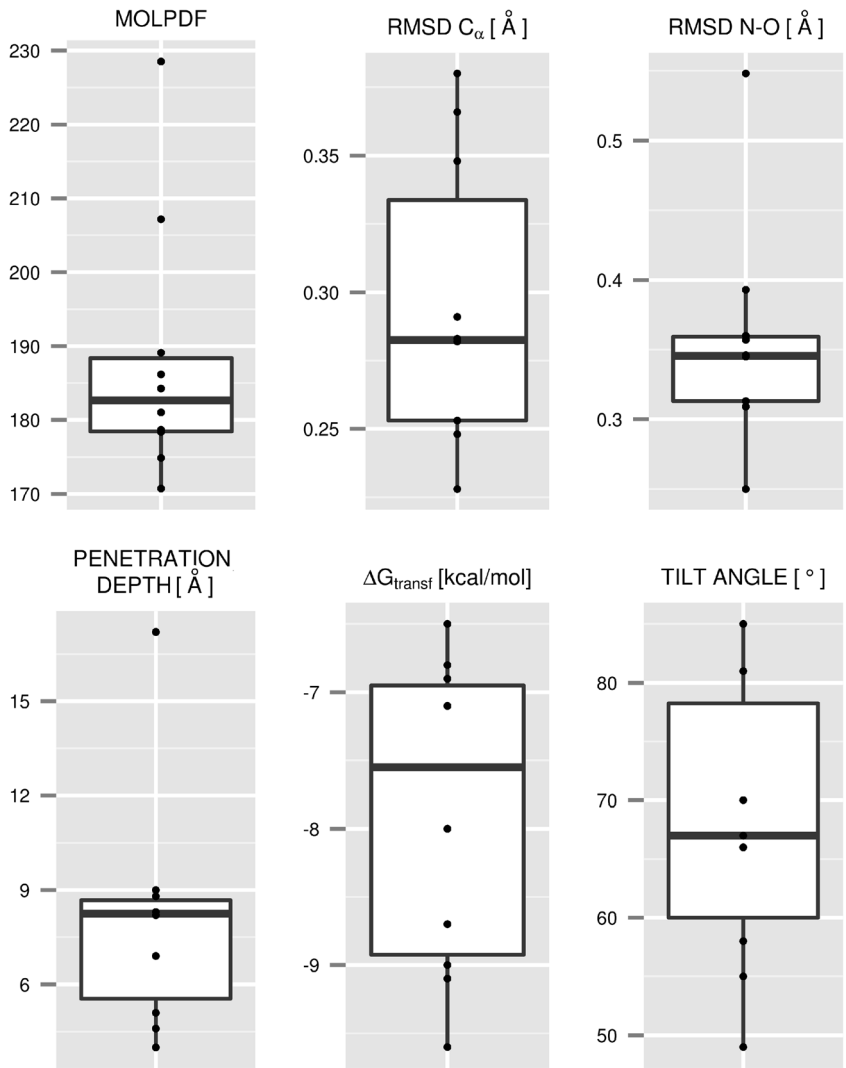

Fig. 1 Distributions of the Modeller and PPM server parameters for the K11V automodel mode. *Points in the plots* represent the values for each of 10 automodels. *Lines inside boxes* indicate median values. Box plots are bounded by the first and third quartiles

box plot expresses the third quartile, while the lower side shows the first quartile. The height of the box represents the interquartile range, which means that 50 % of all observations fall between the upper and lower box sides. The bold line inside each box expresses the median value.

Results and discussion

Cylindrin and its mutants

Fifty models each were generated for wild-type cylindrin and its mutants using different modeling parameters. Modeling cylindrin sequence on cylindrin structure gave the baseline for the range of output parameters from the Modeller and PPM for analyses of other modeling tasks (see "Methods"). The results obtained in this case are presented in Fig. 1. Based on the molpdf criteria, the best wild-type cylindrin model obtained using the automodeling method was selected. The molpdf value for that model equaled 170.74, the RMSD for Cα was 0.283 Å, while the RMSD for N–O was 0.313 Å. The values of the Modeller parameters varied depending on the modeling mode, as shown in Fig. 2. These variations resulted from the introduction of additional user-defined symmetry restraints. Transmembrane features obtained from the PPM server did not present differences in the distribution of parameters between modeling modes. For the best cylindrin model, these parameters were as follows: 4.6 Å for the penetration depth, −7.1 kcal/mol for ΔG_{transf}, and 85° for the tilt angle.

The energy values obtained for cylindrin transfer are typical of peripheral proteins, which implies that cylindrin does

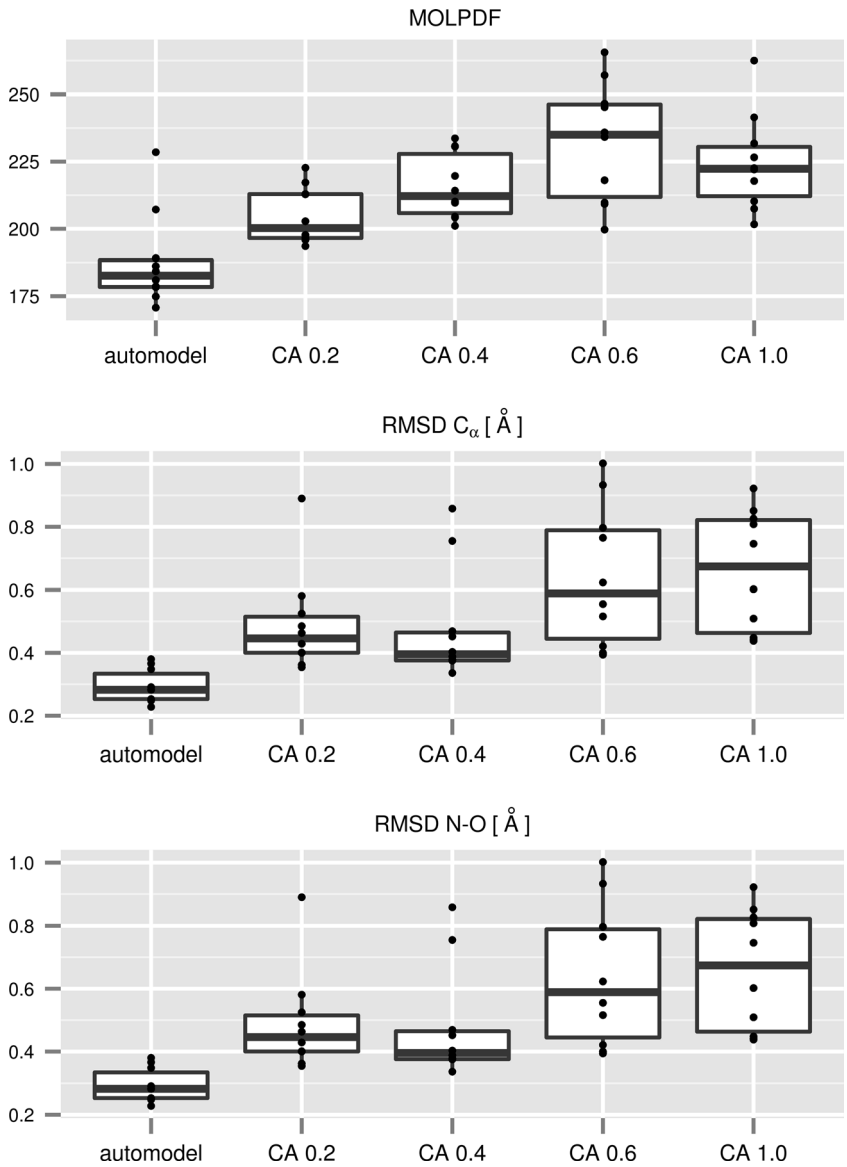

Fig. 2 Distributions of the Modeller parameters for K11V and different modeling methods: automodeling method (*automodel*) and user-defined symmetry restraints for Cα (*CA*) with four weights (0.2, 0.4, 0.6, 1.0). *Points on the plots* represent the values for each model. *Lines inside boxes* indicate median values. Box plots are bounded by the first and third quartiles

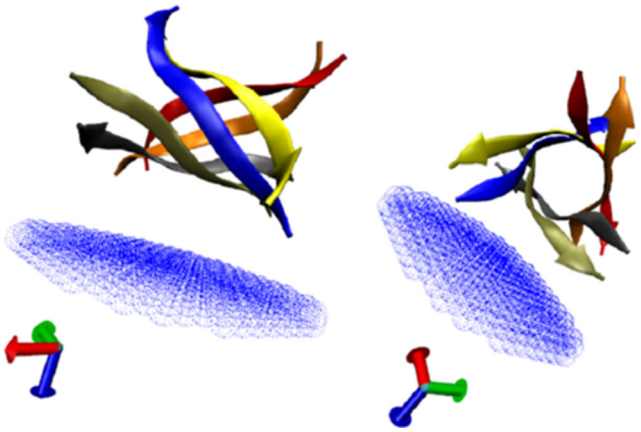

Fig. 3 The orientation of cylindrin (K11V) relative to the membrane surface (which is shown as a *blue mesh*). Picture was generated with VMD

not incorporate into the membrane. This hypothesis is also supported by the penetration depth values, which are too small to allow for pore creation. The tilt angle was also unfavorable to peptide incorporation into the membrane—a high tilt angle value indicates that cylindrin is almost parallel to the membrane surface (Fig. 3). Figure 3 was generated based on the PPM outcome; the corresponding pdb file is provided in the ESM (3SGO_PPM.pdb).

Our modeling results are consistent with experimental data obtained by Laganowsky and colleagues [24]. Their studies confirmed that cylindrin is toxic to mammalian cells however indicated that there was no increase in liposome permeability in the presence of cylindrin [24].

Wild-type cylindrin had a lower molpdf value than its single and double mutants, as shown in Fig. 4. The significantly higher molpdf values of the mutants of cylindrin allowed

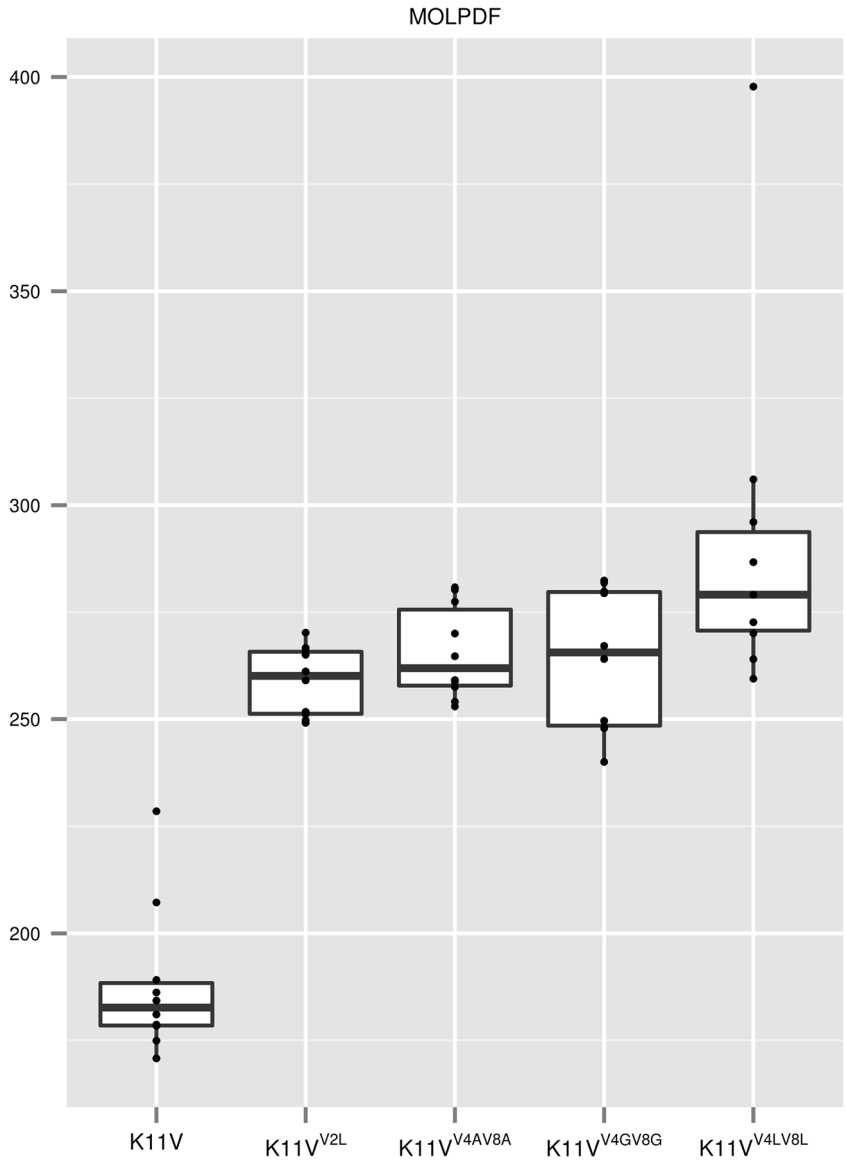

Fig. 4 Box plots with molpdf values for cylindrin and its mutants (automodeling mode). *Points on the plots* represent values for each of 10 models. *Lines inside boxes* indicate median values. Box plots are bounded by the first and third quartiles

wild-type cylindrin to be statistically distinguished from its mutants. We also found that our method is sensitive enough to permit single to be distinguished from double mutations based on the molpdf criteria.

If we consider the transmembrane parameters, the cylindrin mutants exhibited lower values of the penetration depth and wider ranges of the transfer energy and tilt angle. However, there was no transfer energy as low as −12 kcal/mol, the value for the peripheral proteins. The large tilt angles, at least 45°, also do not permit the incorporation of the molecule into the membrane. The distribution of transmembrane parameters was similar for wild-type cylindrin and its mutants. As expected, the values of these parameters did not differ statistically significantly between the wild-type cylindrin and its mutants, meaning that transmembrane parameters could not be used as distinguishing criteria. Additionally, the consistency of the mutant group was tested. Statistical comparisons of $K11V^{V2L}$ with $K11V^{V4AV8A}$, $K11V^{V4GV8G}$, and $K11V^{V4LV8L}$ were performed. No clear statistical differences or similarities were found.

In the molecular dynamics studies [26], wild-type cylindrin showed stability whereas all of the mutants were identified as unstable during the simulations. Thus, the mutants could be distinguished from wild-type cylindrin based on the simulations. Our results obtained from homology modeling were consistent with MD studies and permitted a tentative assessment of structural stability. However, the key question is whether the toxicity of the modeled peptide fragment can be evaluated.

Table 2 p values obtained using the Kolmogorov–Smirnov ($p_{K\text{-}S}$) and Wilcoxon (p_{Wx}) tests for comparisons of $K11V^{V4W}$-TR and K11V in terms of the molpdf parameter. The significance level was 0.05

Method		$p_{K\text{-}S}$	p_{Wx}
Automodel		1.08×10^{-5}	1.08×10^{-5}
User-defined symmetry restraints	CA 0.2	2.17×10^{-4}	4.33×10^{-5}
	CA 0.4	1.08×10^{-5}	1.08×10^{-5}
	CA 0.6	1.08×10^{-5}	1.08×10^{-5}
	CA 1.0	2.06×10^{-3}	7.25×10^{-4}

K11V-TR and its mutant

The tandem-repeat cylindrin oligomer is highly toxic to mammalian cells, while the $K11V^{V4W}$-TR mutant shows little or no cell toxicity [24]. Furthermore, $K11V^{V4W}$-TR exhibits only very weak binding with the conformational antibody A11 [24]. In contrast, both cylindrin and tandem-repeat cylindrin show high affinity for A11 [24]. Accordingly, it may be assumed that K11V-TR and $K11V^{V4W}$-TR exhibit conformational changes at the oligomeric state. Therefore, these two peptides were used as positive and negative controls in our studies.

The main aim here was to distinguish the highly toxic K11V-TR from the almost nontoxic mutant $K11V^{V4W}$-TR based on the generated model parameters.

The best models derived using the automodeling method are shown in Fig. 5. It can be observed that the mutated residues do not explicitly disturb the

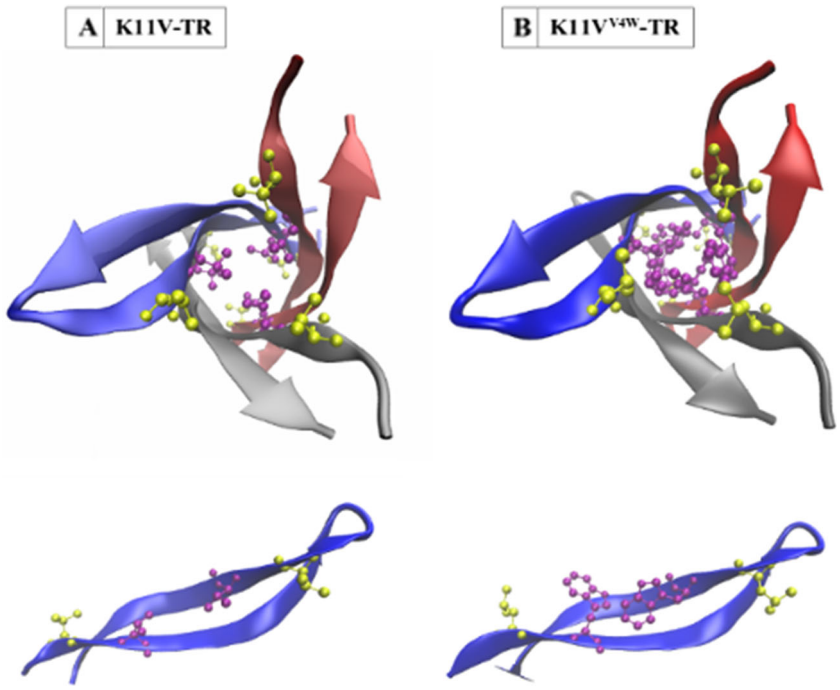

Fig. 5a–b The best models of K11V-TR and $K11V^{V4W}$-TR obtained using the automodeling method, based on the molpdf. **a** *Top*: structure of K11V-TR with V2 (*yellow*) and V4 (*purple*). *Bottom*: chain A with V2 (*yellow*) and V4 (*purple*). **b** *Top*: structure of $K11V^{V4W}$-TR with L2 (*yellow*) and W4 (*purple*). *Bottom*: chain A with L2 (*yellow*) and W4 (*purple*). Pictures were generated with VMD

structure of the tandem-repeat cylindrin. However, tryptophan appears to block the center of the structure of K11V^{V4W}-TR (Fig. 5b).

Table 2 shows the p values obtained from the statistical analysis (with respect to K11V) of the molpdf using the Kolmogorov–Smirnov and Wilcoxon tests. Statistical analysis clearly distinguished between K11V-TR and K11V^{V4W}-TR on the basis of the molpdf. Other parameters, obtained from either Modeller or the PPM server, did not allow for a clear distinction.

These results support the claim that the toxicity of cylindrin does not arise through its creation of transmembrane structure. Also, the difference between the molpdf values of K11V-TR and K11V^{V4W}-TR may allow toxic peptide models to be distinguished from nontoxic ones. The changes in toxicity may result from the disrupted barrel-like structure of the mutants. Therefore, the cylindrin structure appears to be a good template for modeling small amyloid oligomers.

Sup35

In order to verify whether cylindrin is a good template for modeling other small amyloid oligomers, calculations were performed on fragments of the prion protein sup35.

As in the previous cases, 50 models were generated for each sequence. The parameters (molpdf, RMSD Cα, RMSD N–O, ΔG_{transf}, tilt angle, and penetration depth) obtained from Modeller and PPM web server were checked for each model. The main aim here was to attempt to distinguish between three fragment groups (A, B, C; see "Methods") based on those parameters.

Due to the minor sequence adjustment between the prion fragments and the cylindrin template, wide distributions of Modeller parameters were observed. As shown in Fig. 6, the differences between the best and worst models for a given modeling type were negligible.

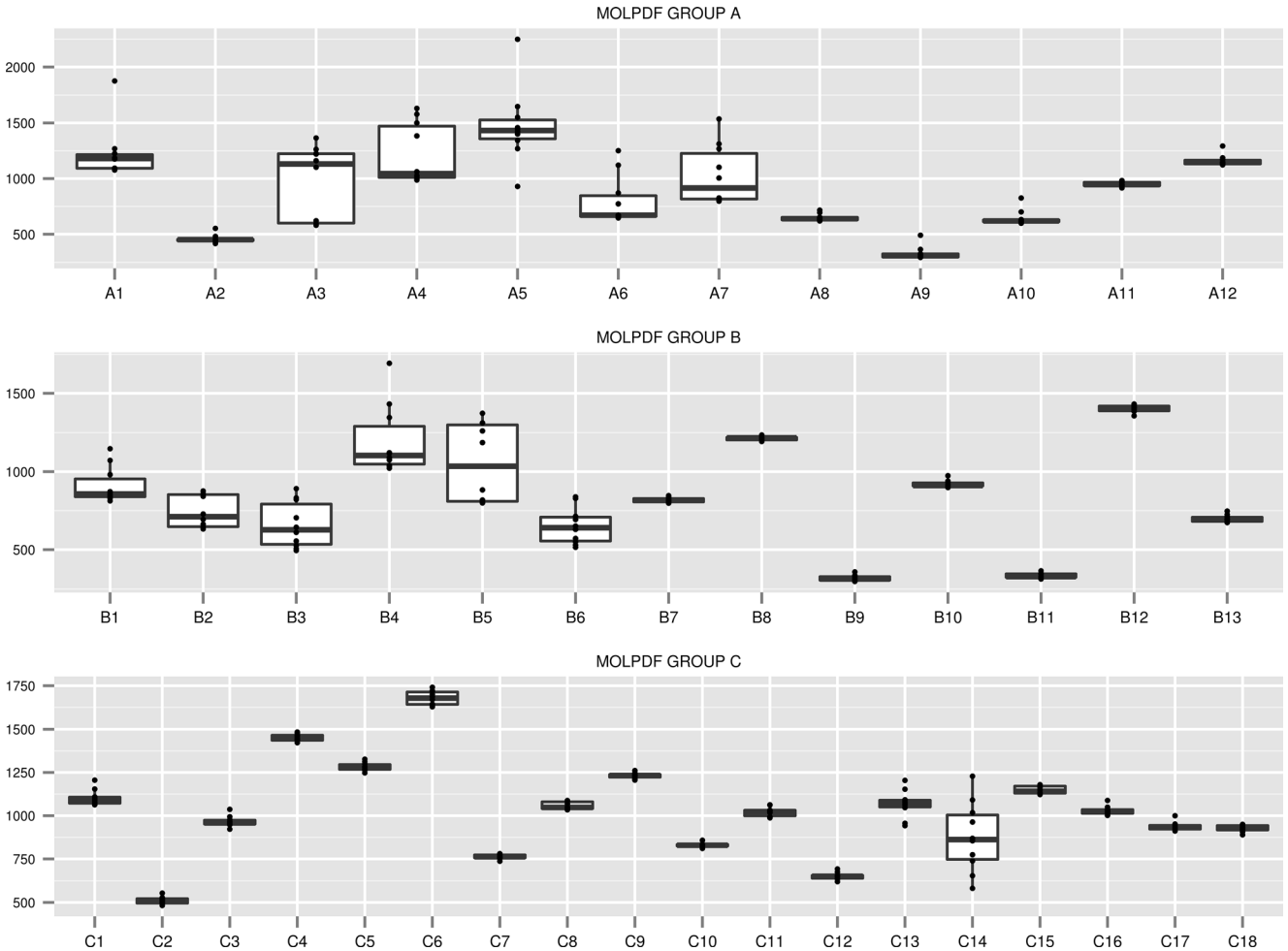

Fig. 6a–c Distribution of the molpdf parameter (automodeling method) for fragments of the prion protein sup35, where the fragments were divided into three groups: **a** fibril formers, **b** fibril non-formers from the middle of the protein, and **c** fibril non-formers from the terminal part of the protein. *Points on the plots* represent values for each of 10 models

There was no clear distinction between groups A, B, and C for any of the Modeller parameters. However, as shown in Fig. 6 for the automodel molpdf, differences were observed along the sequence. At the beginning of the sequence (A1–A7 and B1–B6), the distributions of the values were wider. The remaining sequence was mainly characterized by small distributions of values.

For the transmembrane parameters obtained from the PPM web server, there was no difference between the parameter distributions, regardless of the modeling type employed. For all of the sup35 fragments, the tilt angle was in the range 50–80°, the penetration depth was 4–6 Å, and the transfer energy ranged from −10 to −5 kcal/mol. This means that all of the models generated were peripheral proteins that were oriented parallel to the membrane surface.

Kolmogorov–Smirnov and Wilcoxon tests were performed to check for significant differences between the modeled groups. The Modeller parameters did not present any clear statistical differences between the modeled sequences from different groups. Hence, the models from group A could not be distinguished from the models of other groups. As expected, tests of the PPM web server parameters showed no significant differences among the groups.

In the case of the prion protein sup35, modeled with cylindrin as a template, statistically significant differences between fibril formers and fibril non-formers were not found.

Conclusions

We studied the potential use of the structure of cylindrin as a template for modeling amyloid pores—hypothetical transmembrane structures that are considered to appear in cells during amyloid diseases. The cylindrin protein, obtained from the amyloid segment K11V (KVKVLGDVIEV) artificially induced in a bacterial genome, is the only pore structure of amyloid oligomers that has been experimentally resolved. Its shape—a six-stranded antiparallel beta barrel—may potentially allow the uncontrolled movement of ions between the intra- and extracellular spaces, leading to cell toxicity. Importantly, as expected, cylindrin and tandem-repeat cylindrin induce cell toxicity while some of their mutants do not. However, the mechanism of this toxicity is still unknown. In our study, we threaded short-fibril-forming and non-forming sequences into the cylindrin pore (the template), using comparative modeling provided by Modeller. This approach, although not as accurate as modeling with molecular dynamics, greatly reduced the computational cost of the task, which allowed a large number of models to be tested within a reasonable timeframe. Moreover, our modeling results on the stability of cylindrin and its mutants were in accord with previously published results from molecular dynamics studies. The results from Modeller showed that wild-type cylindrin was a more stable form than its mutants threaded onto the cylindrin structure, and the difference was statistically significant. This confirmed that the modeling method used here is sensitive enough to distinguish the subtleties necessary to achieve the aim of our modeling.

Using this methodology, we addressed the question of whether cylindrin could be a good template for other oligomeric amyloid forms. We found that all mutant forms of crystallin amyloid-forming sequences, as well as prion sup35 sequences, were reasonably compatible with the cylindrin structure and could be modeled based on this template. However, when using prion sup35 sequences, we showed that cylindrin-like pores could also be formed from non-amyloidogenic sequences. Hence, the cylindrin-like structure is not limited to amyloid forms of peptides. We also showed, by comparing the molpdf values of K11V-TR and K11V^{V4W}-TR, that toxic and nontoxic mutants of tandem-repeat cylindrin could be statistically distinguished. However, this statistical distinctness may be attributed to the disrupted barrel-like structure of the mutants.

Finally, we investigated whether the cellular toxicity that occurs after introducing cylindrin or tandem-repeat cylindrin into a cell results from the incorporation of these molecules into the cell membrane, creating conducting ionic channels. A similar study was also conducted on other hypothetical pore structures threaded onto the cylindrin structure. Our results contradicted the hypothesis that cylindrin, tandem-repeat cylindrin, their mutants, or amyloid cylindrin-like pores from prion sequences could form transmembrane channels that are capable of disrupting the membrane barrier. Based on our results obtained from Modeller and PPM, cylindrins can be considered peripheral membrane proteins that position themselves at a small angle with respect to the membrane plane, meaning that they are unable to conduct any ions into the cell. This result is in accord with experimental findings published in the supplement to [24]; that study found that there was no ion leakage from large unilamellar vesicles (LUV) that were incubated with tandem-repeat cylindrin. The results of that study may indicate that a nonconductive transmembrane channel was formed by tandem-repeat cylindrin, or they could indicate that tandem-repeat cylindrin does not insert into the membrane—which our modeling results also suggest.

Acknowledgments This work was partially supported by the grant N N519 643540 from the National Science Centre of Poland.

Appendix

Table 3 Fibril formers and fibril non-formers from sup35 (UniProt ID: P05453) [7] (A: fibril formers, B: fibril non-formers from the middle of sup35, C: fibril non-formers from the end of sup35)

Label	Sequence	Residues	
		From	To
A1	GNNQQNYQQY	7	16
A2	YSQNGNQQQG	16	25
A3	RYQGYQAYNA	28	37
A4	GGYYQNYQGY	43	52
A5	YQNYQGYSGY	46	55
A6	YSGYQQGGYQ	52	61
A7	YQQGGYQQYN	55	64
A8	PQGGRGNYKN	94	103
A9	NFNYNNNLQG	103	112
A10	YNNNLQGYQA	106	115
A11	NLQGYQAGFQ	109	118
A12	NDFQKQQKQA	127	136
B1	AGYQQQYNPQ	67	76
B2	QQQYNPQGGY	70	79
B3	YNPQGGYQQY	73	82
B4	QGGYQQYNPQ	76	85
B5	YQQYNPQGGY	79	88
B6	YNPQGGYQQQ	82	91
B7	KPKKTLKLVS	139	148
B8	KTLKLVSSSG	142	151
B9	KLVSSSGIKL	145	154
B10	SSSGIKLANA	148	157
B11	GIKLANATKK	151	160
B12	LANATKKVGT	154	163
B13	ATKKVGTKPA	157	166
C1	KVGTKPAESD	160	169
C2	TKPAESDKKE	163	172
C3	AESDKKEEEK	166	175
C4	DKKEEEKSAE	169	178
C5	EEEKSAETKE	172	181
C6	KSAETKEPTK	175	184
C7	ETKEPTKEPT	178	187
C8	EPTKEPTKVE	181	190
C9	KEPTKVEEPV	184	193
C10	TKVEEPVKKE	187	196
C11	EEPVKKEEKP	190	199
C12	VKKEEKPVQT	193	202
C13	EEKPVQTEEK	196	205
C14	PVQTEEKTEE	199	208
C15	TEEKTEEKSE	202	211
C16	KTEEKSELPK	205	214
C17	EKSELPKVED	208	217
C18	ELPKVEDLKI	211	220

References

1. Kagan BL (2012) Pores in the pathogenesis of neurodegenerative diseases. Prog Mol Biol Transl Sci 107:296–320
2. Morris KL, Rodger A, Hicks MR, Debulpaep M, Schymkowitz J, Rousseau F, Serpell LC (2010) Exploring the sequence–structure relationship for amyloid peptides. Biochem J 450(2):275–283
3. Garbuzynskiy SO, Lobanov MY, Galzitskaya OV (2010) FoldAmyloid: a method of prediction of amyloidogenic regions from protein sequence. Bioinformatics 26(3):326–332
4. Goldschmidt L, Tenga PK, Riek R, Eisenberg D (2010) Identifying the amylome, proteins capable of forming amyloid-like fibrils. Proc Natl Acad Sci USA 107:3487–3492
5. Bryan AW, O'Donnell CW, Menke M, Cowen LJ, Lindquist S, Berger B (2011) STITCHER: dynamic assembly of likely amyloid and prion β-structures from secondary structure predictions. Proteins 80:410–420
6. O'Donnell CW, Waldispühl J, Lis M, Halfmann R, Devadas S, Lindquist S, Berger B (2011) A method for probing the mutational landscape of amyloid structure. Bioinformatics 27:i34–i42
7. Maurer-Stroh S, Debulpaep M, Kuemmerer N, Lopez de la Paz M, Martins IC, Reumers J, Morris KL, Copland A, Serpell L, Serrano L, Schymkowitz JW, Rousseau F (2010) Exploring the sequence determinants of amyloid structure using position-specific scoring matrices. Nat Methods 7:237–242
8. Stanislawski J, Kotulska M, Unold O (2013) Machine learning methods can replace 3D profile method in classification of amyloidogenic hexapeptides. BMC Bioinforma 14:21
9. Kotulska M, Unold O (2013) On the amyloid datasets used for training PAFIG-how (not) to extend the experimental dataset of hexapeptides. BMC Bioinforma 14:351
10. Gasior P, Kotulska M (2014) FISH Amyloid—a new method for finding amyloidogenic segments in proteins based on site specific co-occurence of aminoacids. BMC Bioinforma 15:54
11. Wojda U, Salinska E, Kuznicki J (2008) Calcium ions in neuronal degeneration. IUBMB Life 60(9):575–590
12. Shafrir Y, Durell S, Arispe N, Guy HR (2010) Models of membrane-bound Alzheimer's Abeta peptide assemblies. Proteins 78(16):3473–3487
13. Kagan BL, Thundimadathil J (2010) Amyloid peptide pores and the beta sheet conformation. Adv Exp Med Biol 677:150–167
14. Wojda U, Kuznicki J (2013) Alzheimer's disease modeling: ups, downs, and perspectives for human induced pluripotent stem cells. J Alzheimers Dis 34(3):563–588
15. Arispe N, Pollard HB, Rojas E (1993) Giant multilevel cation channels formed by Alzheimer disease amyloid beta-protein (A beta P-(1–40)) in bilayer membranes. Proc Natl Acad Sci USA 90(22): 10573–10577
16. Allen MJ, Lacroix JJ, Ramachandran S, Capone R, Whitlock JL, Ghadge GD, Arnsdorf MF, Roos RP, Lal R (2012) Mutant SOD1 forms ion channel: implications for ALS pathophysiology. Neurobiol Dis 45(3):831–838
17. Meleleo D, Galliani A, Notarachille G (2013) AβP1-42 incorporation and channel formation in planar lipid membranes: the role of cholesterol and its oxidation products. J Bioenerg Biomembr 45(4): 369–381
18. Bitan G, Tarus B, Vollers SS, Lashuel HA, Condron MM, Straub JE, Teplow DB (2003) A molecular switch in amyloid assembly: Met35 and amyloid beta-protein oligomerization. J Am Chem Soc 125(50):15359–15365
19. Bernstein SL, Dupuis NF, Lazo ND, Wyttenbach T, Condron MM, Bitan G, Teplow DB, Shea JE, Ruotolo BT, Robinson CV, Bowers MT (2009) Amyloid-β protein oligomerization and the importance of tetramers and dodecamers in the aetiology of Alzheimer's disease. Nat Chem 1(4):326–331

20. Garai K, Frieden C (2013) Quantitative analysis of the time course of Aβ oligomerization and subsequent growth steps using tetramethylrhodamine-labeled Aβ. Proc Natl Acad Sci USA 110(9):3321–3326
21. Kim S, Klimov DK (2013) Binding to the lipid monolayer induces conformational transition in Aβ monomer. J Mol Model 19(2):737–750
22. Sali A, Blundell T (1993) Comparative protein modeling by satisfaction of spatial restraints. J Mol Biol 234(3):779–815
23. Eswar N, Webb B, Marti-Renom MA, Madhusudhan MS, Eramian D, Shen MY, Pieper U, Sali A (2006) Comparative protein structure modeling using Modeller. Curr Protoc Bioinformatics 15:5.6.1–5.6.30
24. Laganowsky A, Liu C, Sawaya MR, Whitelegge JP, Park J, Zhao M, Pensalfini A, Soriaga AB, Landau M, Teng PK, Cascio D, Glabe C, Eisenberg D (2012) Atomic view of a toxic amyloid small oligomer. Science 335(6073):1228–1231
25. Sali A, Overington JP (1994) Derivation of rules for comparative protein modeling from a database of protein structure alignments. Protein Sci 3(9):1582–1596
26. Berhanu WM, Hansmann UH (2013) The stability of cylindrin β-barrel amyloid oligomer models—a molecular dynamics study. Proteins 81(9):1542–1555
27. Lomize MA, Pogozheva ID, Joo H, Mosberg HI, Lomize AL (2012) OPM database and PPM web server: resources for positioning of proteins in membranes. Nucleic Acids Res 40:D370–D376