Mutating Amyloid Fibrils into Specific Conformations

Prof. Jerome Waldispuhl
Mohamed Smaoui
Yuksel Asli Sari

Winter 2013

Abstract

Amyloids are protein aggregates that arise from inappropriately folded proteins and polypeptides that aggregate a stacking behaviour to form insoluble fibrils and they cause serious diseases such as Alzheimer's disease. Changing amyloid shape may prevent mentioned diseases, as well as showing other uses like forming very strong microtubules. We developed a computer software Mutation Generator, which suggests a number of mutations on the amyloid that will change its structure to a pre-defined shape. It makes use of hydrophobic and polar qualities of the aminoacids in the chain to decide which aminoacid to replace with. It determines at which points a mutation is most needed by calculating the difference between hydrophobicity and polarity of the current aminoacid and the corresponding point at the pre-defined shape. Then it makes a suggestion according to which quality current aminoacid is missing the most.

1 Introduction

Amyloids are protein aggregates that arise from inappropriately folded proteins and polypeptides. These polypeptides tend to interact and aggregate to form insoluble fibrils by stacking on top of each other as can be seen in Figure 1. Amyloid fibrils are long, straight and unbranched filaments of 40-120 Å in diameter and they are resistant to protease digestion. This

malformation of normally soluble proteins to depositions of insoluble proteins cause cell death and tissue degeneration [1]. Amyloids are associated with more than 20 serious diseases, among which, one of the most important is Alzheimers disease that is caused by β-amyloid protein (Aβ) [2].

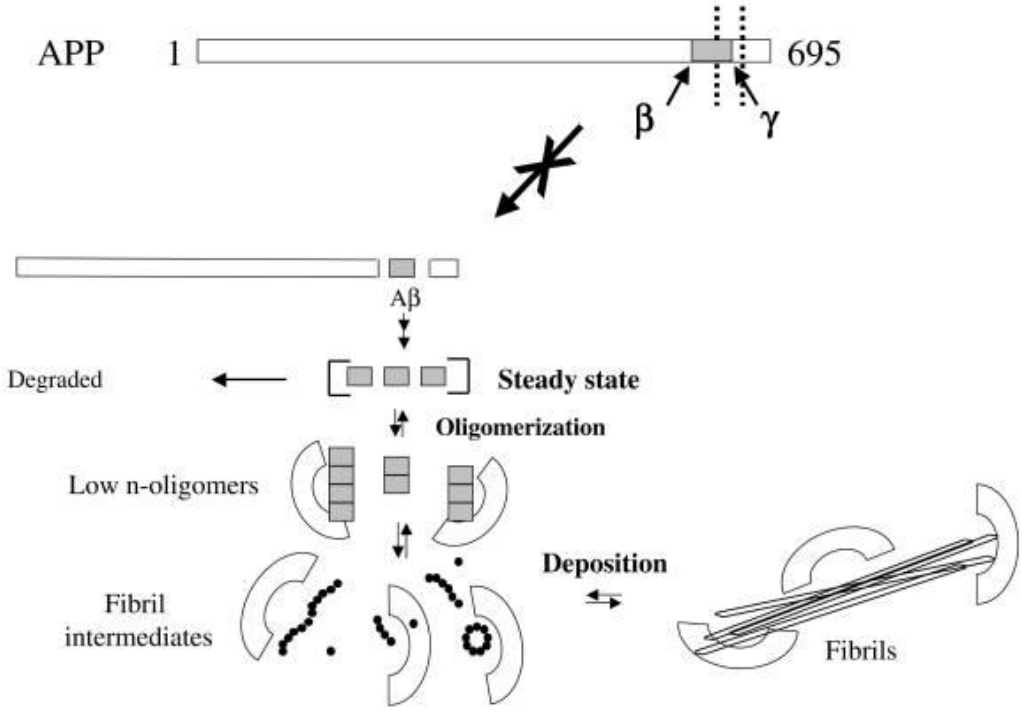

Figure 1: Emergence of Aβ by proteolytic cleavage from amyloid precursor protein (APP) followed by association of Aβ to form oligomers and fibrils in Alzheimers disease [2]

Researching on amyloids is important because if the structure can be possibly changed by mutating the aminoacid of the polypeptides while forming in the cell, this may be a technique to prevent further cell deaths. Another reason to change the structure of the fibrils is to invoke few mutations to turn the fibrils into a more desired shape, such as microfibrils, which are very strong structures and thus demanded by the industry. We developed a new method that aims to generate a structurally described the output fibril from an existing polypeptide chain.

2 Methods

The collection of hydrophobicities and polarities of aminoacids has a substantial effect on the overall structure of the protein [3] [4]. For example, if a hydrophobic an aminoacid located inside a fold of the protein structure is changed to a hydrophilic aminoacid, the new aminoacid will tend to be attracted by water and it will probably influence the shape by distorting the fold. Likewise, a polar aminoacid will form hydrogen bonds with other aminoacids in the protein and if this polar aminoacid is changed into a non-polar aminoacid, this bond will most likely be broken.

Following this approach, we developed a computer software called Mutation Generator (source code can be found at `https://github.com/yaslisari/MutationGenerator`, that takes the structural description of the desired shape of a protein and an aminoacid chain as inputs, the input can be a PDB file or a chain. The desired shape is specified as the foreseen hydrophobic (or hydrophilic) regions and polar (or non-polar) regions. The input aminoacid chain is then processed aminoacid by aminoacid calculating the tendency of change. This tendency is calculated using hydrophobicity and polarity tables. The general principle is the further away from the desired output, the more tendency to change. For example if aminoacid 12 is foreseen to be hydrophobic, and aminoacid 12 is already hydrophobic, the tendency to change will be low and if hydrophilic the tendency will be high. A simple summary of how Mutation Generator works is as illustrated in Algorihm 1.

For calculating tendencies, two tendencies are taken into account; these are hydrophobic tendency and polar tendency. To calculate each of these tendencies, one of the seven different hydrophobic tables can be used (Kyte-Doolittle [3], Hopp-Woods [5], Cornette [6], Eisenberg [7], Rose [8], Janin [9] and Engelman GES [10]) and a polarity table is used [4].

Overall tendency is calculated as in the following formula:

Overall tendency = hydrophobic (or hydrophilic) tendency + polar (or non-polar) tendency

Then the tendencies are ranked and first n (a predefined number) aminoacids are suggested to be changed. For each change, Mutation Generator suggest two changes, one of them is most hydrophobic or polar aminoacid (depends on which tendency caused that aminoacid to be chosen) and the other is the median aminoacid in between the initial and the most hydrophobic or polar aminoacid.

The output consists of all the suggested changes. This output is then fed into SCWRL [11] which adds the suggested changes into the PDB file. CreateFibril takes the PDB file to build polymorphic brils of amyloid proteins [12] and AQUASOL explores their stability by means of stability landscapes by using dipolar solvent model that captures the effect of dipole-dipole interactions and computes the hydration shell that forms around proteins. Both the initial chain and the results of Mutation Generator is fed into AQUASOL [13], and in the determination step, overall energies of these chain are compared. The chain with the lowest energy is selected because that will be the most stable model. The overall schema of the method can be seen in Figure 2.

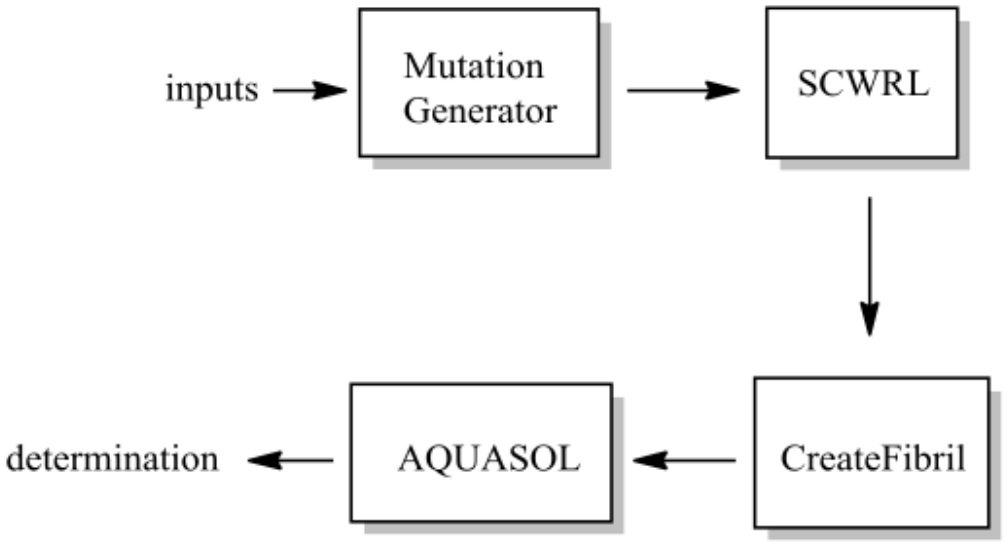

Figure 2: Overall structure of the method

Specifically for A protein, we have a suggested foreseen changes table that can be used with Mutation Generator (Table 1). These suggestions are derived from the shape of the protein (Figure 3).

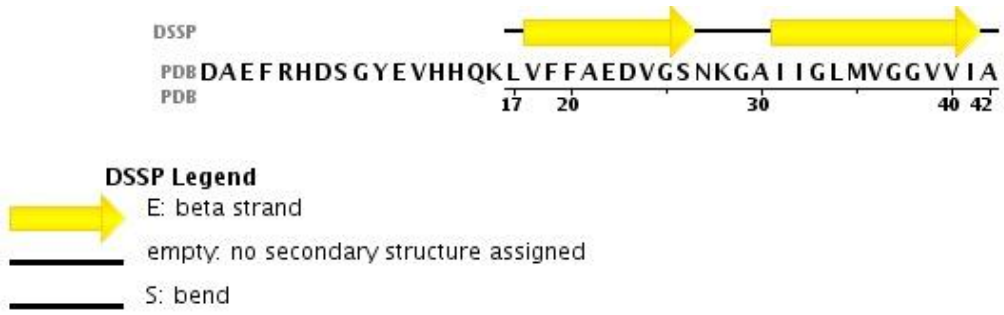

Figure 3: Structure of Aβ protein (RCSB.org)

3 Future Work

After receiving the first results, we aim to ameliorate overall tendency formula by including a learning algorithm such that it finds the best scalars α and β for the following formula:

Overall tendency = α hydrophobic (or hydrophilic) tendency + β polar (or non-polar) tendency

With this modification, the algorithm will be able to decide which of polarity and hydrophobicity is more important. Following this, the ranking of tendencies will be different, resulting in different predictions.

References

[1] M. Ramírez-Alvarado, J. S. Merkel, and L. Regan, "A systematic exploration of the influence of the protein stability on amyloid fibril formation in vitro," *Proceedings of the National Academy of Sciences*, vol. 97, no. 16, pp. 8979–8984, 2000.

[2] G. B. Irvine, O. M. El-Agnaf, G. M. Shankar, and D. M. Walsh, "Protein aggregation in the brain: the molecular basis for alzheimers and parkinsons diseases," *Molecular Medicine*, vol. 14, no. 7-8, p. 451, 2008.

[3] J. Kyte and R. F. Doolittle, "A simple method for displaying the hydropathic character of a protein," *Journal of molecular biology*, vol. 157, no. 1, pp. 105–132, 1982.

[4] J. Zimmerman, N. Eliezer, and R. Simha, "The characterization of amino acid sequences in proteins by statistical methods," *Journal of Theoretical Biology*, vol. 21, no. 2, pp. 170–201, 1968.

[5] T. P. Hopp and K. R. Woods, "Prediction of protein antigenic determinants from amino acid sequences," *Proceedings of the National Academy of Sciences*, vol. 78, no. 6, pp. 3824–3828, 1981.

[6] J. L. Cornette, K. B. Cease, H. Margalit, J. L. Spouge, J. A. Berzofsky, and C. DeLisi, "Hydrophobicity scales and computational techniques for detecting amphipathic structures in proteins," *Journal of molecular biology*, vol. 195, no. 3, pp. 659–685, 1987.

[7] D. Eisenberg, E. Schwarz, M. Komaromy, and R. Wall, "Analysis of membrane and surface protein sequences with the hydrophobic moment plot," *Journal of molecular biology*, vol. 179, no. 1, pp. 125–142, 1984.

[8] G. D. Rose, A. R. Geselowitz, G. J. Lesser, R. H. Lee, M. H. Zehfus, *et al.*, "Hydrophobicity of amino acid residues in globular proteins," *Science*, vol. 229, no. 4716, pp. 834–838, 1985.

[9] J. Janin and C. Chothia, "Role of hydrophobicity in the binding of coenzymes," *Biochemistry*, vol. 17, no. 15, pp. 2943–2948, 1978.

[10] D. Engelman, T. Steitz, and A. Goldman, "Identifying nonpolar transbilayer helices in amino acid sequences of membrane proteins," *Annual review of biophysics and biophysical chemistry*, vol. 15, no. 1, pp. 321–353, 1986.

[11] Q. Wang, A. A. Canutescu, and R. L. Dunbrack, "Scwrl and molide: computer programs for side-chain conformation prediction and homology modeling," *Nature protocols*, vol. 3, no. 12, pp. 1832–1847, 2008.

[12] M. R. Smaoui, *A Computational Framework to Create an Ensemble of Stable Amyloid Fibrils*. McGill University Libraries, 2011.

[13] P. Koehl and M. Delarue, "Aquasol: An efficient solver for the dipolar poisson–boltzmann–langevin equation," *The Journal of chemical physics*, vol. 132, no. 6, pp. 064101–064101, 2010.

Algorithm 1 Mutation Generator

```
1: Input 1: Structural description of the protein shape
2: Input 2: Any aminoacid chain (for example, Aβ)
3: for all aminoacids in the chain do
4:    hydophobicTendency = highestHydrophobe - currentHydrophobicity
5:    polarTendency = highestPolar - currentPolarity
6:    tendency = polarTendency + hydrophobicTendency
7: end for
8: sort according to tendency
9: return  first n aminoacids in the rank
   {to suggest a change:}
10: for aminoacid 0 to n in the rank do
11:    if hydrophobicTendency > polarTendency then
12:       return  highestHydrophobe and median{currentAminoacid, high-
          estHydrophobe}
13:    else
14:       return  highestPolar and median{currentAminoacid, highestPolar}
15:    end if
16: end for
```

-	Single	2-Stack	3-Stack	2-Ring	3-Ring	4-Ring	3-Polygon	4-Polygon	5-Polygon
Single	-	18-25 PH	18-25, 31-41 PH	26-31 P, 33-34 H	26-31 P, 33-34 H, 18-25 PH	18-25 P, 26,31,33,34 H, 18-25 PH	31-42 H, 26,42 P	31-42 H, 26,42 P	31-42 H, 26,42 P
2-Stack	-	-	18-25, 31-41 PH	26-31 P, 33-34 H	26-31 P, 33-34 H, 18-25 PH	18-25 P, 26,31,33,34 H, 18-25 PH	31-42 H, 26,42 P	31-42 H, 26,42 P	31-42 H, 26,42 P
3-Stack	-	-	-	26-31 P, 33-34 H	26-31 P, 33-34 H, 18-25 PH	18-25 P, 26,31,33,34 H, 18-25 PH	31-42 H, 26,42 P	31-42 H, 26,42 P	31-42 H, 26,42 P
2-Ring	-	-	-	-	26,31 P, 26-31 H	26,31 P, 26-31 H	31-42 H, 26,42 P	31-42 H, 26,42 P	31-42 H, 26,42 P
3-Ring	-	-	-	-	-	26-31 H	31-42 H, 26,42 P	31-42 H, 26,42 P	31-42 H, 26,42 P
4-Ring	-	-	-	-	-	-	31-42 H, 26,42 P	31-42 H, 26,42 P	31-42 H, 26,42 P
3-Polygon	-	-	-	-	-	-	-	31-42 Y, 26,42 P	31-42 Y, 26,42 P
4-Polygon	-	-	-	-	-	-	-	-	31-42 Y, 26,42 P

Table 1: Suggested foreseen input for Aβ. Abbreviations: P Polar; H Hydrophobic; Y Hydrophilic. Below the diagonal of the table is the reverse of the upper diagonal (Polar will turn into non-polar, hydrophobic to hydrophilic and vice versa).