"BetaSerpentine": a bioinformatics tool for reconstruction of amyloid structures

Stanislav A. Bondarev[1], **Galina A. Zhouravleva**[1] **and Andrey V. Kajava**[2,3,4]

[1]Department of Genetics and Biotechnology and the Laboratory of Amyloid Biology, Saint Petersburg State University, Saint Petersburg, Russia;
[2]Centre de Recherches de Biochimie Macromolculaire, CNRS, Universit Montpellier, Montpellier, Cédex 5, France;
[3]Institut de Biologie Computationnelle, Montpellier, Cédex, France; [4] University ITMO, Saint Petersburg, Russia;

e-mail: stanislavspbgu@gmail.com

Introduction

The core of amyloid fibrils (including prion fibrils) consists of cross-β structures that are extremely resistant against hydrogen/deuterium exchange, proteases, and chemical denaturation. The cross-β structure of most disease-related and functional amyloids is built of parallel and in-register stacks of structural elements called β-arches, consisting of two β-strands connected by a turn. The size of one β-arch can vary between 15-30 residues. At the same time it is known that a number of amyloid-forming regions can reach size of up to 80 residues. In the past few years, numerous experimental results have suggested that these regions form tandems of adjacent β-arches, called β-serpentines which, in turn, stack to form the superpleated β-structure. Recently, a computer program ArchCandy has been developed that analyzes protein sequences and predicts the location and structural arrangements of individual β-arches. In this work, we present an algorithm and computer program BetaSerpentine that predicts structural arrangements of adjacent β-arches.

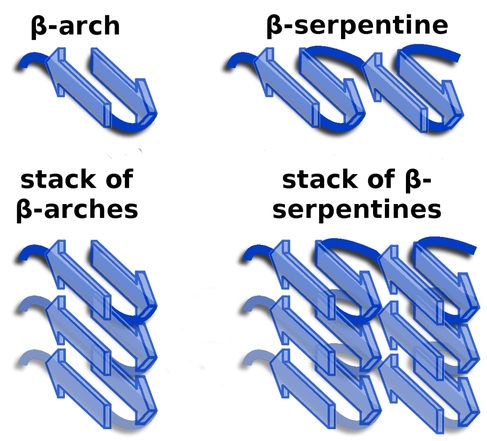

Figure 1: The structure of most amyloids is built of β-arches stacks, which in turn may form complicated structures called β-serpentines.

Rules of β-arch compatibility in β-serpentines

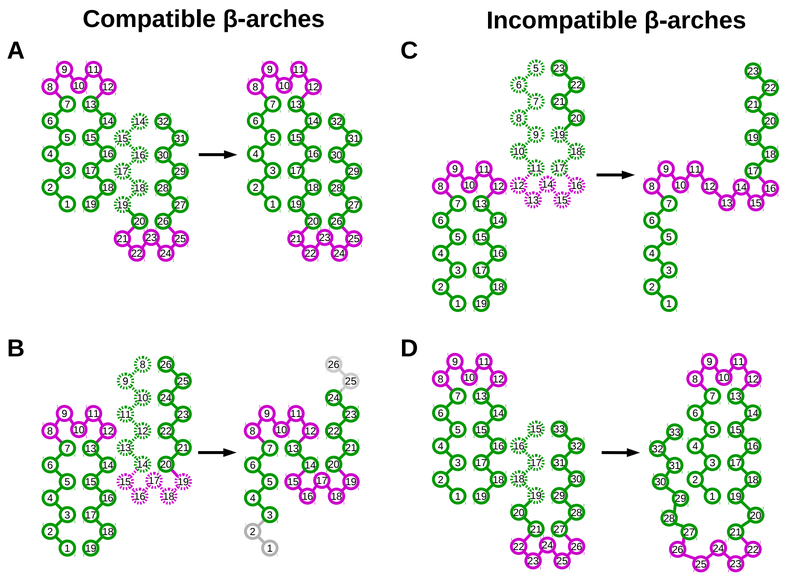

Figure 2: Examples of β-arches that can form β-serpentine (A,B) or not (C,D). Different variants of β-arches overlap are presented on the picture. Only half of them (A,B) leads to the β-serpentine formation. Positions of amino acids are marked within circles. Dashed lines correspond to the overlapping region within one of the compared arches. Regions of β-strands are colored in green, arcs — in magenta, amino acids located outside of the formed β-serpentine — in gray.

Serpentines scores

The final score of serpentine ($FScore$) is a multiplication of several scores listed below. All of them may vary between 0 and 1.

$$FScore = MAScore \times CScore_1 \times CScore_2 \times TScore$$

List of β-serpentines scores:

1. Mean arch scores ($MAScore$) is a mean of separate β-arches scores imported from ArchCandy program and corrected on the length scores of arches.

$$MAScore = \frac{\sum_{i=1}^{N} AScore_i / ALScore_i}{N}$$

$AScore$ — arch score counted by ArchCandy, $ALScore$ — lenght score of arch $(1 - (0.0003462 \times (2 \times L - 7 - 45)^2)$, where L — length of one of the strands), N — number of arches within serpentine.

2. The indexes of compactness ($CScore_1$,$CScore_2$). To evaluate this scores all lenghts are counted according approximation shown on Figire 3.

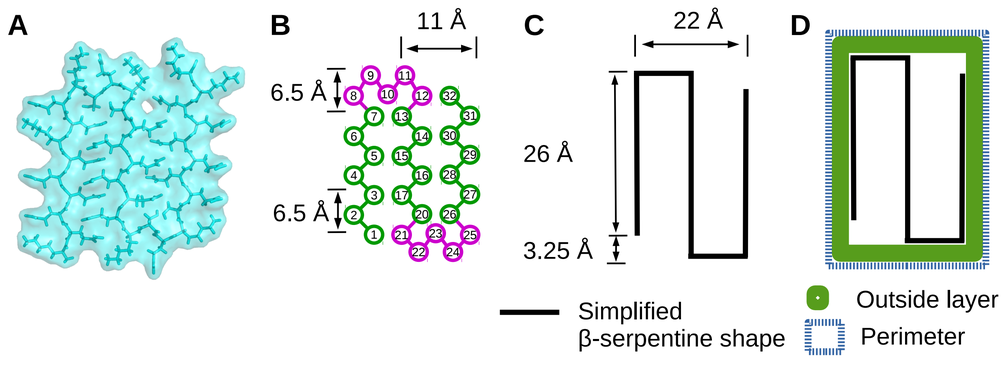

Figure 3: Approximation of β-serpentine shape. To calculate several scores we consider a serpentine as a miander. Linear charecteristics of such structures were calculated according assuptions described below. Length of two amino acids within β-strands is 6.5 Å, this distance is also equal to "height" of arc region, the interval between two β-strands — 11 Å. Width of outside layer was 5 Å. 3D structure of arbitrary β-serpentine is shown on panel A. On B the same structure is presented in schematic view. Panel C represents approximation of this serpentine as miander. Finally, D illustrates calculation of the β-serpentine perimeter.

- $CScore_1 = 1 - \delta l / L$
  δl — difference in length between two neighboring strands, L — total length of strands (Fig. 4).
- $CScore_2 = \frac{4 \times \sqrt{S}}{P}$
  S — square of serpentine, P — perimeter of serpentine (Fig. 4).

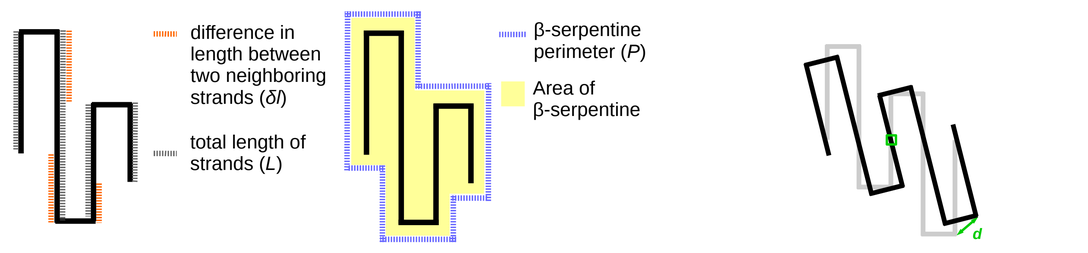

Figure 4: Parameters of β-serpentines that are estimated to calculate indexes of compatness. Perimeter and area of serpentine are evalauted with outside layer.

Figure 5: Calculation of maximal divergense of polypeptide chains caused by fibril twist.

3. Twist Score ($TScore$).

$$TScore = \frac{1}{1 + \exp^{\frac{1.05 - d}{-0.04}}}$$

d — maximal divergence of polypeptide chains caused by fibril twist on 1.8° (Fig. 5).

Table 1: Number of potential β-serpentines predicted by Beta-Serpentine for known amyloids. Summarized results are presented on this table. * — prion-forming domain of corresponding protein are used.

Protein	Uniprot ID	N of arches	N of serpentines	Time (s)
IAPP	P10997	108	2848	43
PrP	P04156	150	838	10
Rip1	Q13546	348	3380	53
Rip3	Q9Y572	224	536	9
α-synuclein	P37840	157	7020	187
Sup35 1-123	P05453 *	179	4372	106
Ure2 1-89	P23202 *	157	7020	186
Rnq1 153-405	P25367 *	758	154647	13714 (≈4)
CPEB	Q967R6	538	13093	620
Orb2	A4V1P2	475	5558	88

Predictions of Sup35p structure for known [PSI$^+$] prion variants

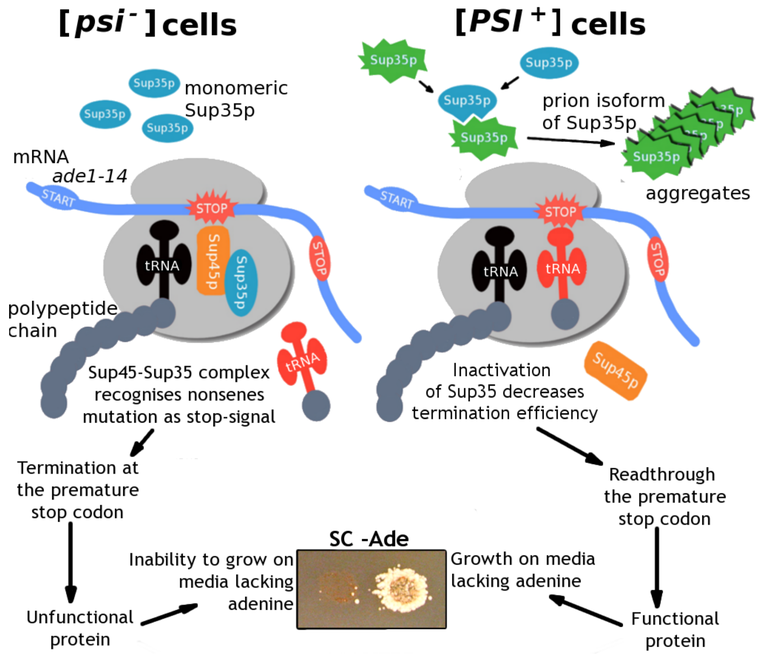

Figure 6: [PSI^+] cells show nonsense-suppressor phenotype. Sup35p (yeasts termination factor eRF3) is a structure protein of [PSI^+] prion. In prion isoform this factor is inactivated forming aggregates. This leads to decrease of termination efficiency and to readthrough of nonsense-mutations (their suppression). In most works mutation *ade1-14* are used to detect [PSI^+] status of cells.

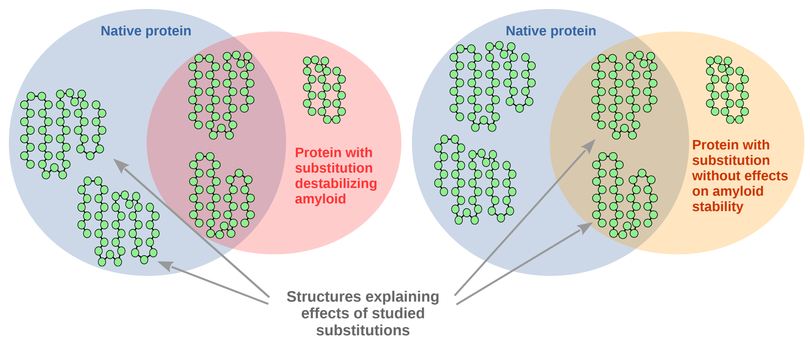

Figure 7: Comparison of β-serpentines sets allows to find structures explaining effects of substitutions.

Table 2: Summarized results of the Sup35 aggregates structures in different yeast strains.

Prion variant or yeast strain	**7A-D794**	**10-7A-D832**	**[VH]**	**[VK]**	**[VL]**
Substitutions destabilizing amyloid	13 (+*PNM2*)	5	13	4	5
Substitutions compatible with amyloid	0	3	0	6	5
Number of serpentines	**35 (2)**	**97**	**0**	**208**	**43**

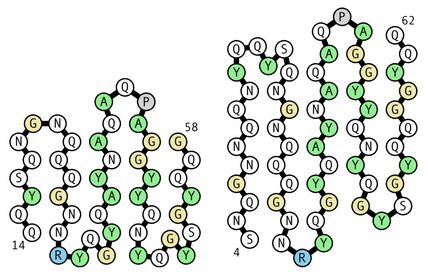

Figure 8: Two predicted structures of Sup35 amyloid aggregate in 74-D694 yeast strain.

Conclusions

Thus, the application of the BetaSerpentine tool to the set of known sup35 mutant alleles showed its usefulness, for the structural interpretation of the experimentally observed effects on Sup35p fibrillogenesis and prion properties. This opens avenues to the interpretation of the experimental results and design of new mutations for the other amyloid-forming proteins.

Acknowledgements

The authors acknowledge Saint-Petersburg State University for research grants 1.41.682.2016 and 1.37.291.2015, Russian Foundation of Basic Research for a research grant 16-34-00582 and ANR INVESTISSEMENTS DAVENIR BIO-INFORMATIQUE "Institut de Biologie Computationnelle".