Supplementary data for article:

Ninković, D. B.; Malenov, D. P.; Petrović, P. V.; Brothers, E. N.; Niu, S.; Hall, M. B.; Belić, M. R.; Zarić, S. D. Unexpected Importance of Aromatic–Aliphatic and Aliphatic Side Chain–Backbone Interactions in the Stability of Amyloids. *Chemistry - A European Journal* **2017**, *23* (46), 11046–11053. https://doi.org/10.1002/chem.201701351

CHEMISTRY
A European Journal

Supporting Information

Unexpected Importance of Aromatic–Aliphatic and Aliphatic Side Chain–Backbone Interactions in the Stability of Amyloids

Dragan B. Ninković,[a, c] Dušan P. Malenov,[b] Predrag V. Petrović,[a, c] Edward N. Brothers,[a] Shuqiang Niu,[d] Michael B. Hall,[d] Milivoj R. Belić,[a] and Snežana D. Zarić*[a, b]

chem_201701351_sm_miscellaneous_information.pdf

Supplementary information

1. CCSD(T)/CBS limit calculations

To be able to computationally study the aggregation between two amyloid β-sheets, appropriate computational method needs to be selected. This is not an easy task since the size of the interacting system and the different type of interactions present in the amyloids should be taken into consideration. The B3LYP-D3 method was selected since it should be able to cope with these issues. Typical interactions between the side-chains present in the amyloid aggregates: aromatic-aromatic, aromatic-nonaromatic, and nonaromatic-nonaromatic interactions, are represented with four model systems shown in Figure S1.
As the "golden standard", CCSD(T)/CBS limit calculations using Mackie's method were performed on model systems obtained by isolating interacting side-chains of amino acids from 3OW9 sequence belonging to KLVFFA segment of Aβ peptide. Due to the high computational cost of the CCSD(T) calculations that considerably increases with the size of the system, Phe is represented by toluene (Phe*), Val by n-propane (Val*) and Leu by 2-methylpropane (Leu*) (Figure S1).

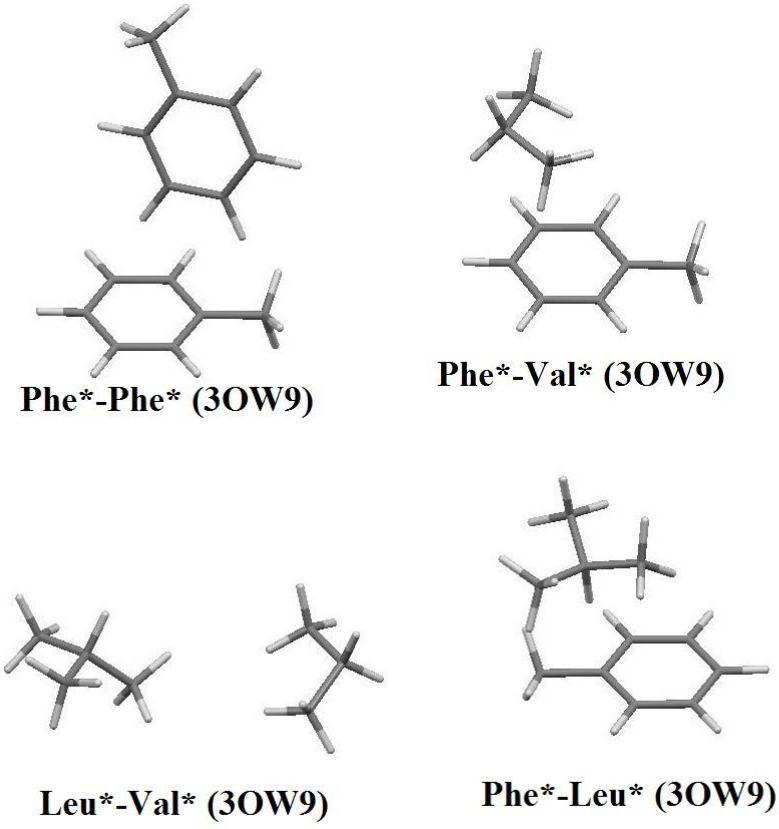

Figure S1. Model systems and interactions used to estimate CCSD(T)/CBS limit

The results calculated with B3LYP-D3 method using different basis sets was compared with the CCSD(T)/CBS calculations (Table S1). As the interaction energies for all three basis sets were similar, we have decided to use 6-31G* basis set since the calculations on B3LYP-D3/6-31G* level of theory are the least computationally demanding and suitable for large amyloid systems studied in this work.

Table S1. Interaction energies calculated at CCSD(T) level at complete basis set (CBS) and with D3 dispersion corrected B3LYP functional using three different basis sets, with removal of basis set superposition error (BSSE); the energies are given in kcal/mol

	CCSD(T)/CBS	**B3LYP-D3**		
		*6-31G**	*def2-SVP*	*def2-TZVP*
Phe*-Phe*	**-3.08**	-3.12	-3.07	-3.10
Phe*-Val*	**-2.78**	-2.69	-2.87	-2.90
Leu*-Val*	**-0.80**	-0.86	-0.87	-0.83
Phe*-Leu*	**-2.47**	-2.71	-2.73	-2.74

2. Presentations of various model systems

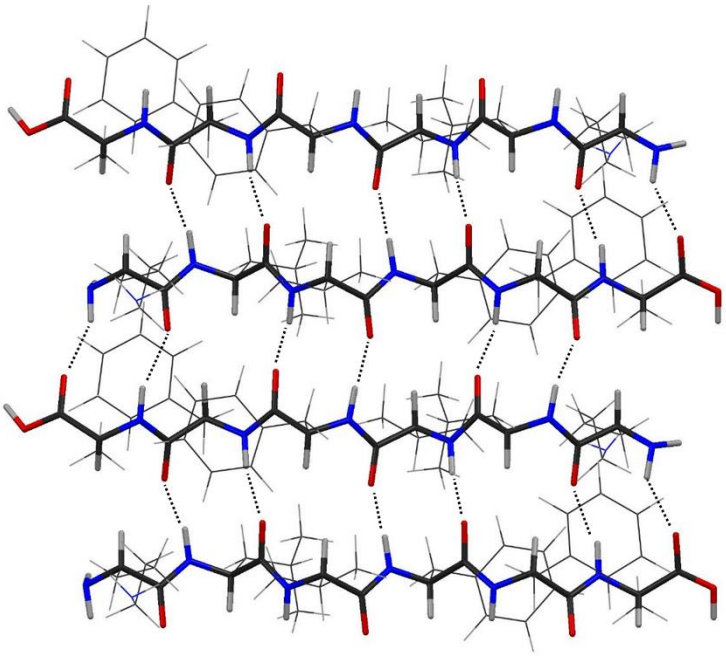

Figure S2. Example of an amyloid β-sheet presenting 2Y2A structure. H-bonds between the monomers are presented.

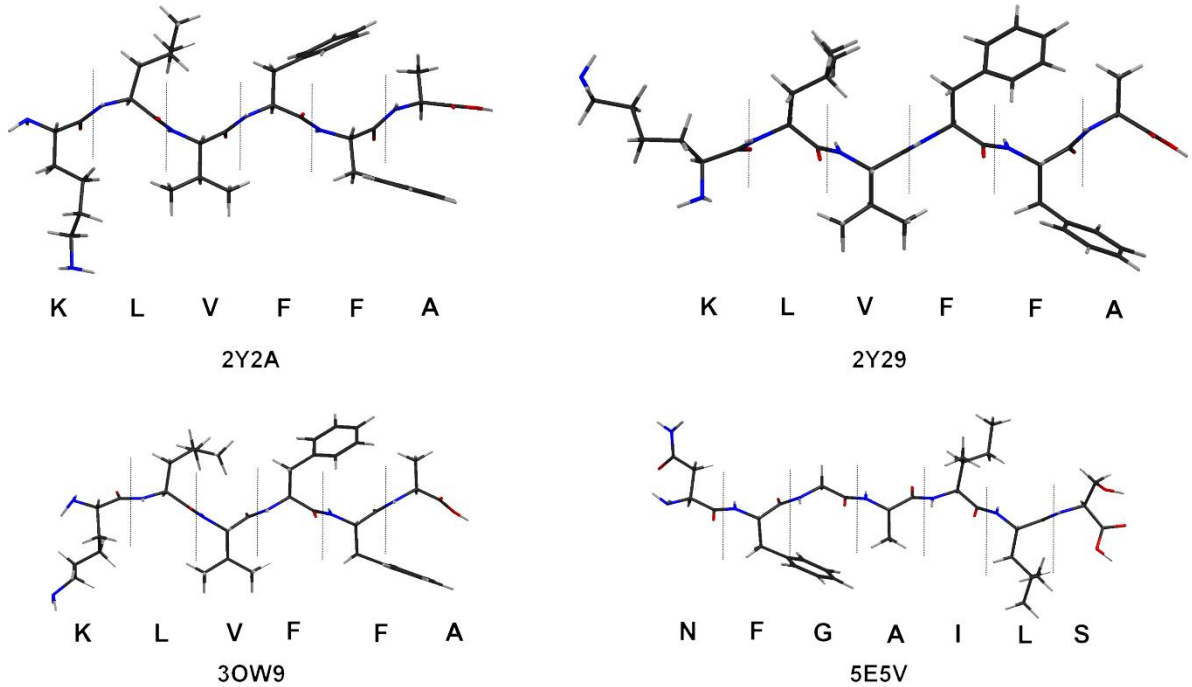

Figure S3. Monomers of the amyloids with aromatic residues: 2Y2A, 2Y29, 3OW9 (KLVFFA sequence), and 5E5V (NFGAILS sequence)

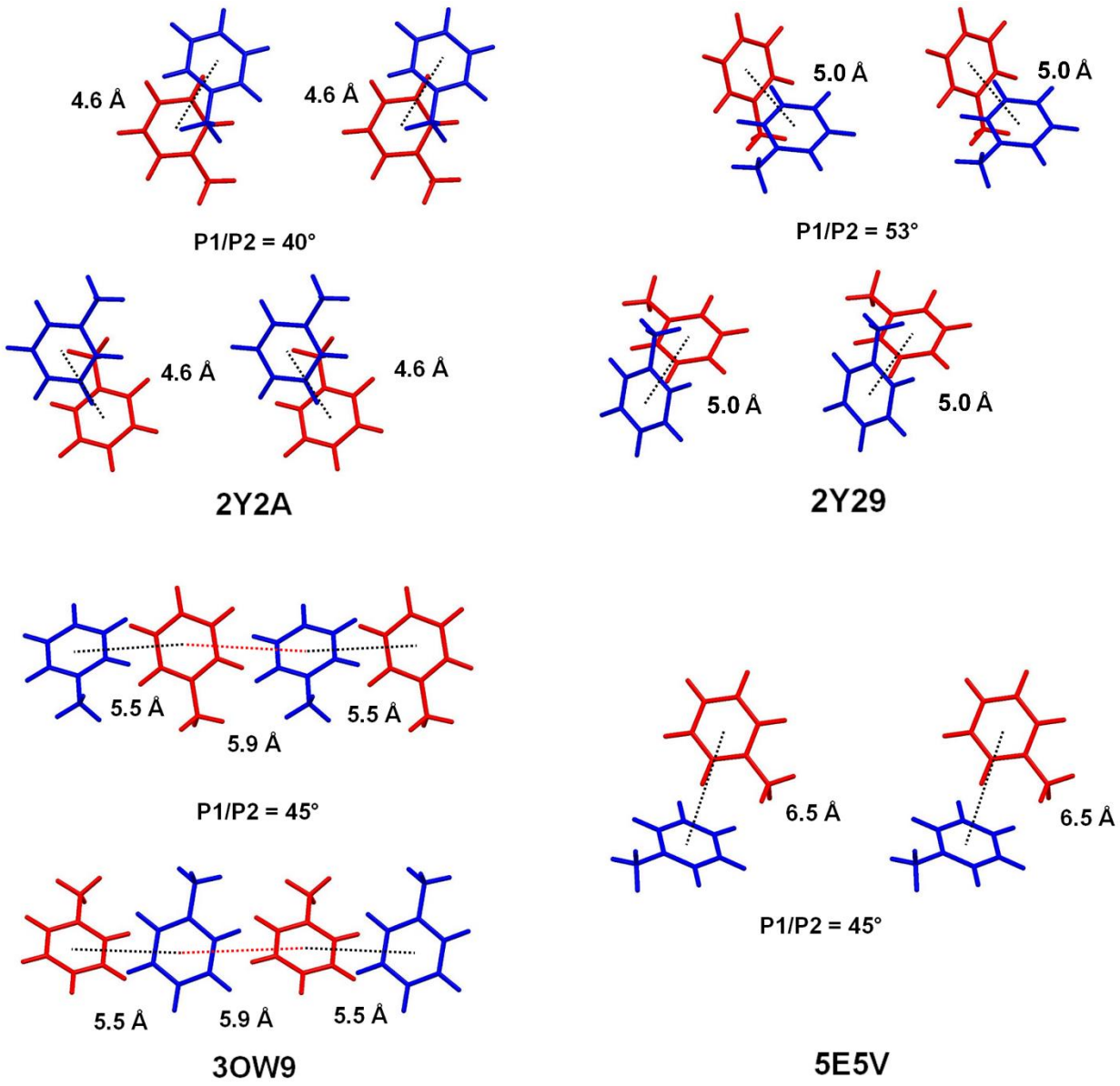

Figure S4. Model systems used to calculate aromatic-aromatic interaction energies between two β-sheets. Top view was used to show the aromatic side chain interactions. Different color coding is used to show the difference between two interacting tetramers. P1/P2 is the angle between the planes of two interacting aromatic rings; these angles are the same for each pair of interacting aromatic rings in one amyloid structure.

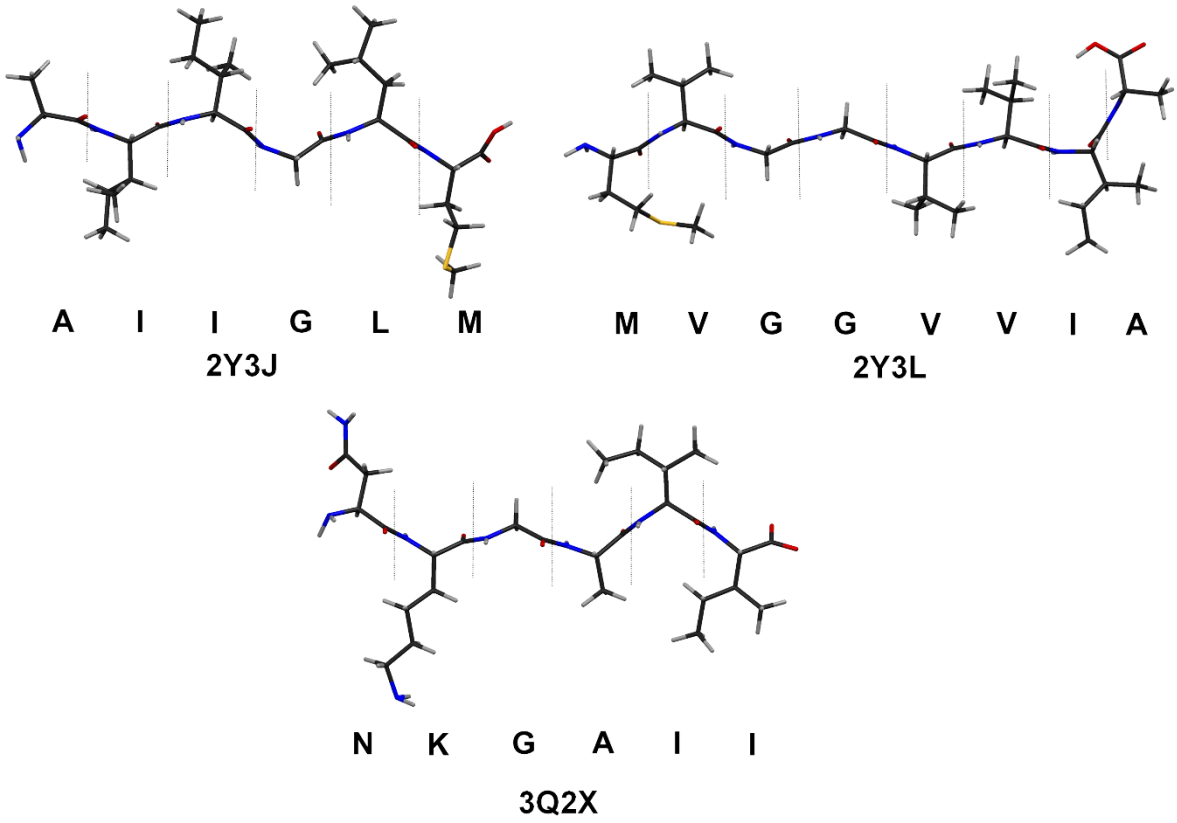

Figure S5. Monomers of the amyloids without aromatic amino acids: 2Y3J (AIIGLM sequence), 2Y3L (MVGGVVIA sequence), and 3Q2X (NKGAII sequence)

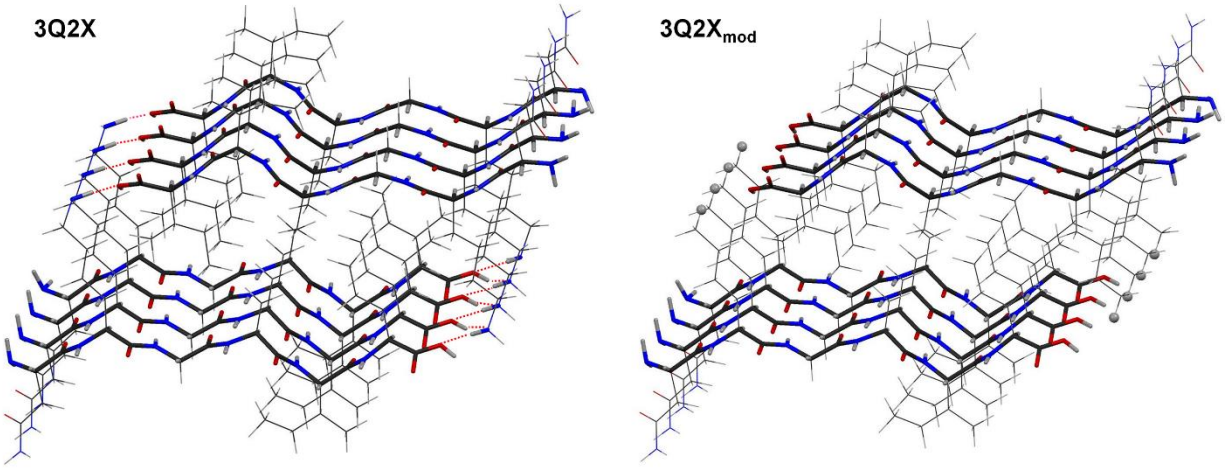

Figure S6. Model systems derived from 3Q2X crystal structure used for calculation of interaction energies between two tetramers. Left - 3Q2X model system with highlighted H-bonds between two β-sheets; right - 3Q2X$_{mod}$ model system with highlighted hydrogen atom that replaced amino group in order to remove the H-bonds between two β-sheets.

3. Scaling of interaction energies between small aromatic molecules

Valence electron method can be justified by looking at the example of the interaction of small aromatic molecules. For example, interacting energies for parallel stacked benzene-benzene (-2.7 kcal/mol) and benzene-naphthalene (-4.7 kcal/mol) differ by 2 kcal/mol. As the nature of the interactions in these systems is the same, the interaction energy divided by the product of valence electrons in two interacting fragments (representing the interaction of all valence electrons), should be similar. When this scaling method is used, the scaled interaction energies are -3.03×10^{-3} and -3.29×10^{-3} kcal/mol for benzene-benzene and benzene-naphthalene respectively. As the difference in scaled interaction energies is less than 10%, the method should be applicable on larger interacting systems.

4. Alternative scaling method

To evaluate the previous scaling method, an alternative way to scale the interaction energies was used. In the alternative approach, we account for the size of the interacting systems by dividing the interaction energies in Table 1 and Table 2 by the number of amino acids involved in the interaction between two β-sheets. This alternative method results in similar conclusions. The scaled interaction energies between aromatic side chains are with energies ranging from -0.69 to -1.46 kcal mol^{-1}. For the interactions of aromatic and nonaromatic side-chains, average scaled energies range from -0.54 to -1.12 kcal mol^{-1}, while the scaled interaction energies between the nonaromatic side-chains range from 0.33 kcal mol^{-1} to -0.80 kcal mol^{-1}.

Table S2. Scaled interaction energies[a] per interacting amino acid (in kcal mol^{-1}) for different types of interactions between two tetramer β-sheets in model systems of 2Y29, 2Y2A, 3OW9 (KLVFFA sequence), 5E5V (NFGAILS sequence), 2Y3J (AIIGLM sequence), 2Y3L (MVGGVVIA sequence), and 3Q2X (NKGAII sequence)/$3Q2X_{mod}$[b] amyloid structures

	Structure	Ar-Ar	Ar^1-nAr^2	Ar^2-nAr^1	Avg (Ar-nAr)[c]	nAr-nAr
With aromatic amino acids	2Y2A	-1.46	-0.66	-0.52	-0.59	-0.30
	2Y29	-1.28	-0.70	-0.57	-0.64	-0.34
	3OW9	-1.14	-1.12	-1.12	-1.12	-0.28
	5E5V	-0.69	-0.74	-0.35	-0.54	-0.48
Without aromatic amino acids	2Y3J	-	-	-	-	0.33
	2Y3L	-	-	-	-	0.19
	3Q2X	-	-	-	-	-0.80
	$3Q2X_{mod}$	-	-	-	-	-0.69

[a] Interactions: Ar-Ar – aromatic-aromatic, Ar-nAr – aromatic-nonaromatic, nAr-nAr – nonaromatic-nonaromatic, nAr-B – nonaromatic-backbone; suffixes in superscript designate association with one of the two interacting β-sheets; [b] a model system modified to remove the influence of the hydrogen bonds; [c] average value for Ar^1-nAr^2 and Ar^2-nAr^1