Electronic Supplementary Material (ESI) for Physical Chemistry Chemical Physics.

Supporting Information: Computational prediction of the supramolecular self-assembling properties of organic molecules: the role of conformational flexibility of amide moieties

Laura Le Bras,[*,†] Yves Dory,[‡] and Benoît Champagne[*,†]

†*Unité de Chimie Physique Théorique et Structurale, Chemistry Department, Namur Institute of Structured Matter, University of Namur, Belgium*

‡*Laboratoire de Synthèse Supramoléculaire. Département de Chimie, Institut de Pharmacologie, Université de Sherbrooke, Sherbrooke, Québec, Canada*

E-mail: laura.lebras@unamur.be; benoit.champagne@unamur.be

Contents

Validation of GAFF parameters

To validate the use of GAFF force field for the description of the **B4s** and **B9s** families we have performed two types of calculation. The first one consisted in an optimization at the ωB97X-D/6-311+G(d,p) level in vacuum, denoted after as the QM level. For the second one, the structure is minimized by GROMACS with GAFF in vacuum also (GAFF level).

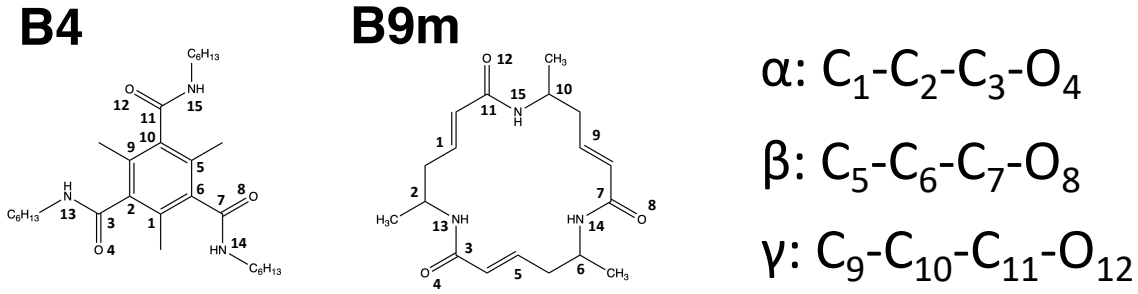

Figure S1: Representation of the structure of **B4** and **B9m**. The corresponding numbering of the atoms and the definition of dihedral angles are also provided.

We focused our study on the amide moieties as this is the most important part of the molecule for the type of application we target, i.e supramolecular assemblies ensured by amide-based hydrogen bond network. One can observe in Table S1 that all the structural parameters for both molecules are similar independently from the calculation scheme that is used (QM or GAFF). The distances within the amide moieties, 1.23 (1.23) Å and 1.34 (1.35) Å for the C=O and C-N bonds for **B4** (**B9**), are equivalent within the QM level. Within GAFF approaches, there are slight variations (± 0.01 Å) for all the distances in comparison to the QM approach. The most important feature of the molecules is the orientation of the amide groups with respect to the plane of the molecule. We thus reported the values relative to this orientation with for example α, β, and γ dihedral angles. Whether for **B4** or **B9** molecules, amide moieties are globally perpendicular to the plane of the molecules with values around 90° for the three previously mentioned dihedral angles. Once again,

Table S1: Selected geometrical parameters (see Figure S1 for the definition) obtained in vacuum at the QM and GAFF level for **B4** and **B9m**. Distances are in Å and dihedral angles are in degrees.

	B4		**B9m**	
	QM	GAFF	QM	GAFF
Distances				
C_3-O_4	1.23	1.22	1.23	1.22
C_7-O_8	1.23	1.23	1.23	1.24
C_{11}-O_{12}	1.23	1.22	1.23	1.22
C_3-N_{13}	1.34	1.34	1.35	1.36
C_7-N_{14}	1.34	1.34	1.35	1.34
C_{11}-N_{15}	1.34	1.34	1.35	1.34
Dihedral				
α	-89	-87	93	89
β	89	92	93	89
γ	88	91	92	91
O_4-C_3-N_{13}-H_{13}	179	178	178	178
O_8-C_7-N_{14}-H_{14}	180	179	178	176
O_{12}-C_{11}-N_{15}-H_{15}	179	178	178	179

there is a slighter larger flexibility within the GAFF approach but amide moieties are still perpendicular to the plane of the molecule. We also looked at the plane of the amide moiety itself. For both approaches and both molecules, the amide is planar with an almost perfect alignment of O, C, N, and H atoms, indicated by the values of the last three dihedral angles of Table S1 that are close to 180°.

We have represented on Figure S2 a superpositon of the structures obtained with both approaches for **B4** and **B9**. The same conclusion hold as what has been said after the analysis of Table S1. There is no major distorsion of the core of the structure when GAFF is used when compared to the one obtained at the QM level. Nevertheless, there is a slight shift of the lateral groups for **B4**. The position of the alkyl chain of GAFF structure is shifted with respect to the position of the same alkyl chain for QM structure. This shift is illustrated through the value of δ angle (Figure S2) that is equal to 10° for the three lateral groups. This distorsion of the **B4** molecule does not affect the targeted properties, that is to say perpendicular amide moieties to ensure stacking of the molecules. We thus concluded

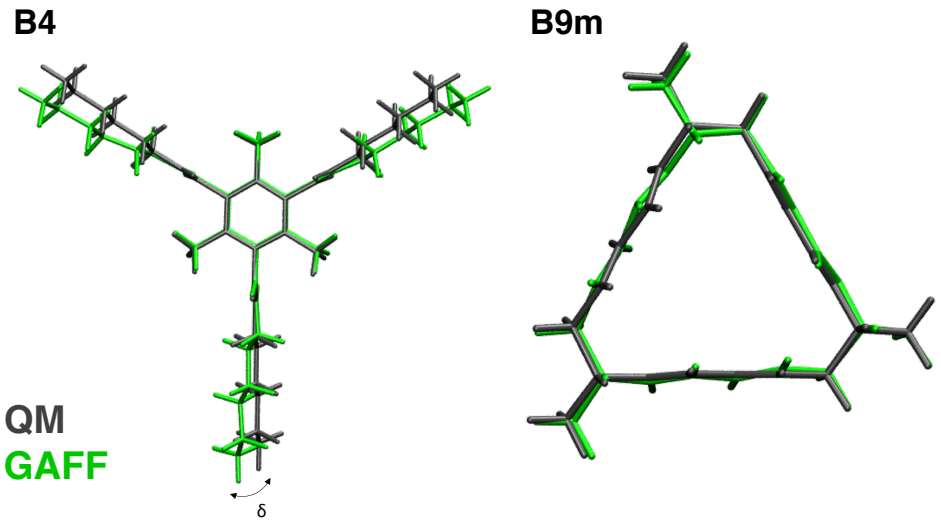

Figure S2: Structure superpositions for **B4** (left) and **B9m** (right) after a QM optimization (grey) and a minimization with GAFF (green).

that GAFF force field is correctly parametrized to describe our molecules and will therefore be considered for all our MD simulations.

General parameters for molecular dynamics simulations

In Table S2 are gathered the size and the corresponding number of water molecules present in the simulation box. Box were built by adding a pure 3 Å layer of water molecules from the edges of the organic molecule. In all cases, boxes were cubic.

Table S2: Size (in Å) of the boxes (length between the edges) and composition in term of number of water molecules for the different simulation boxes for **B4s** and **B9s** families for monomers and pentamers.

	B4	**B4c**	**B4p**	**B4pI**	**B9**	**B9m**
			Monomer			
Size	72	74	74	74	74	71
Water	11664	13632	13567	13548	13571	11908
			Pentamer			
Size	83	85	88	87	86	84
Water	20792	22345	25851	25426	22802	21476

Particular structure of B4s and B9s families

On Figure S3 we have represented the particular structures of both **B4s** and **B9s** families. Indeed, for each family, the three amide moieties are perpendicular to the plane of the core of the molecule.

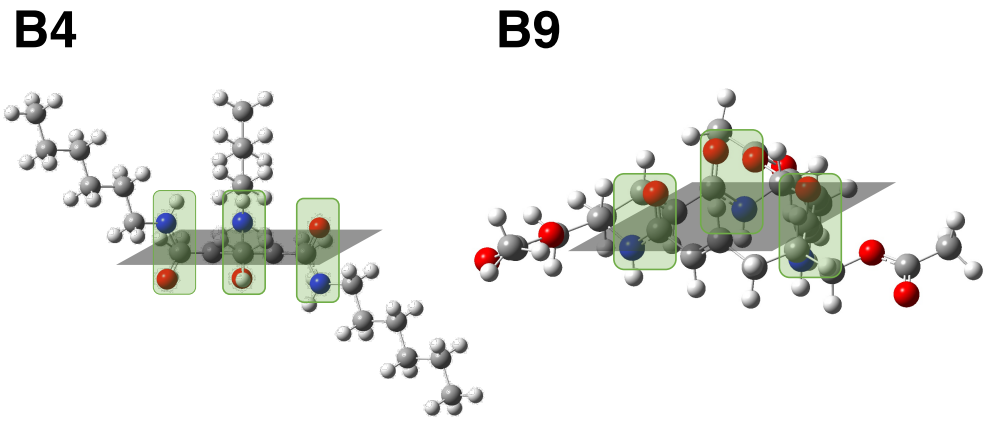

Figure S3: Representation of the structures of **B4** (left) and **B9** (right) molecules.

For **B4s**, the three amide moieties do not point towards the same side of the plane of the molecule.

Hybrid hydrogen bond network

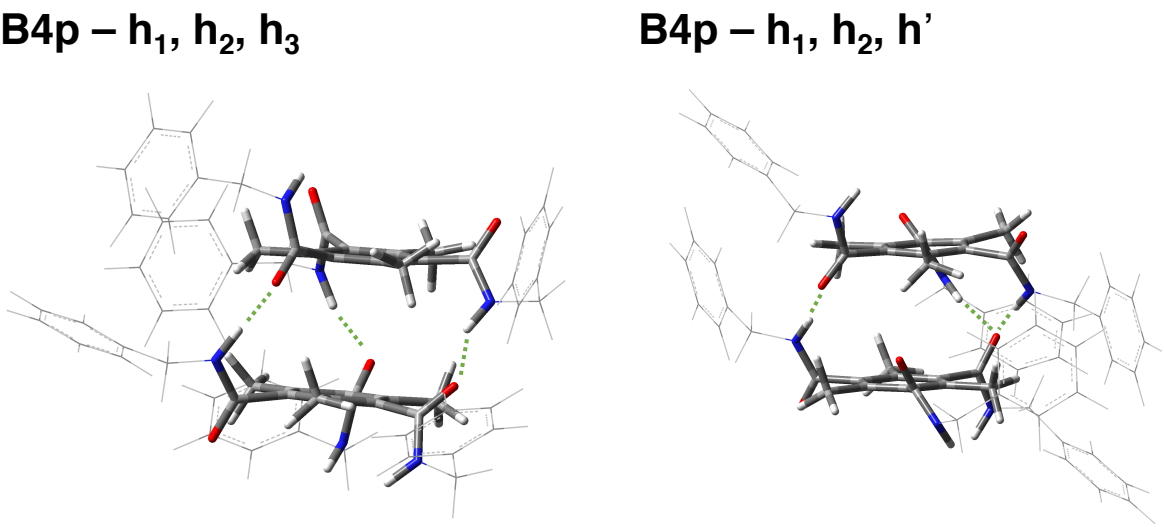

Figure S4: Representation of the structures involving "expected" H-Bonds, h_1, h_2, and h_3 (left) and the hybrid network, h_1, h_2, and h' (right).

Pentamers

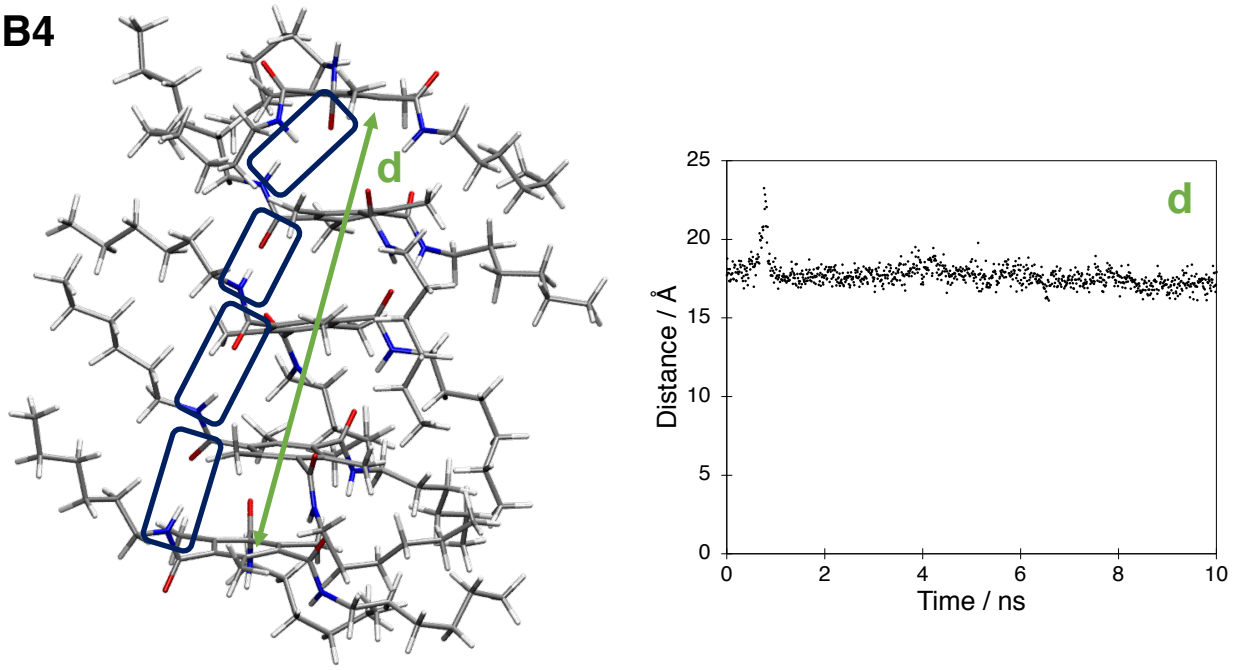

Figure S5: On the left, representative structure of a pentamer involving **B4** monomers. h_1 bonds are highlighted within blue frames and the total distance, d, is also indicated. On the right, time evolution of the d distance along the 10 ns simulation time.

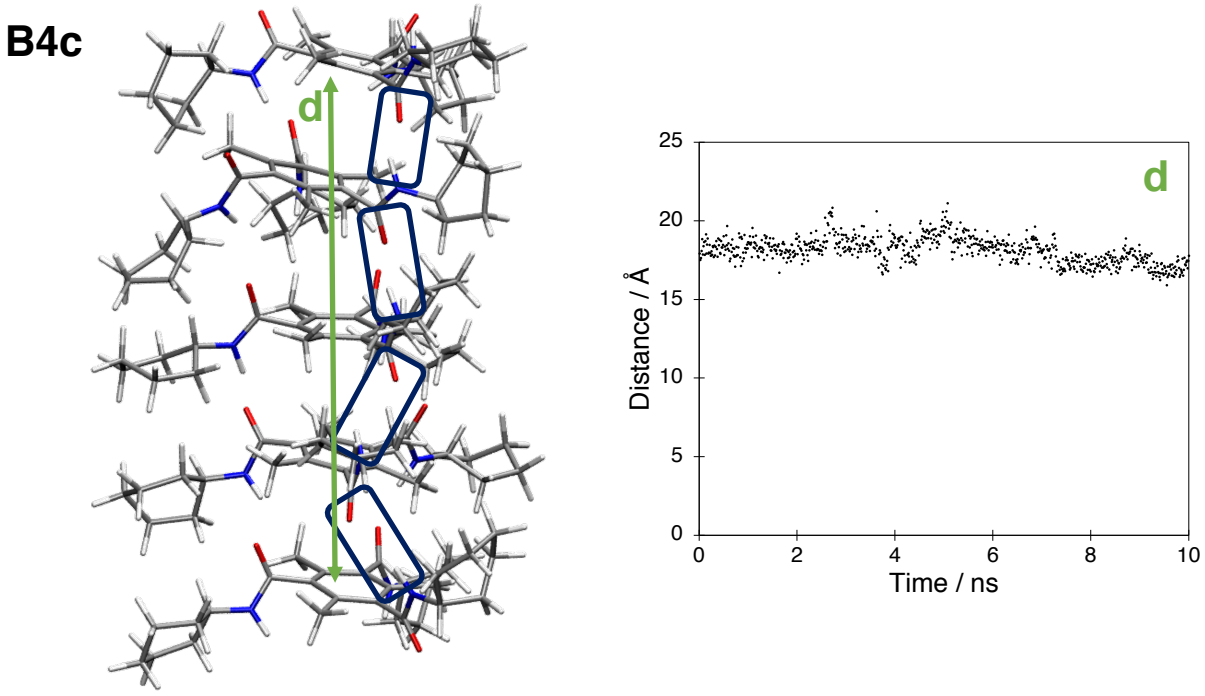

Figure S6: On the left, representative structure of a pentamer involving **B4c** monomers. h_1 bonds are highlighted within blue frames and the total distance, d, is also indicated. On the right, time evolution of the d distance along the 10 ns simulation time.

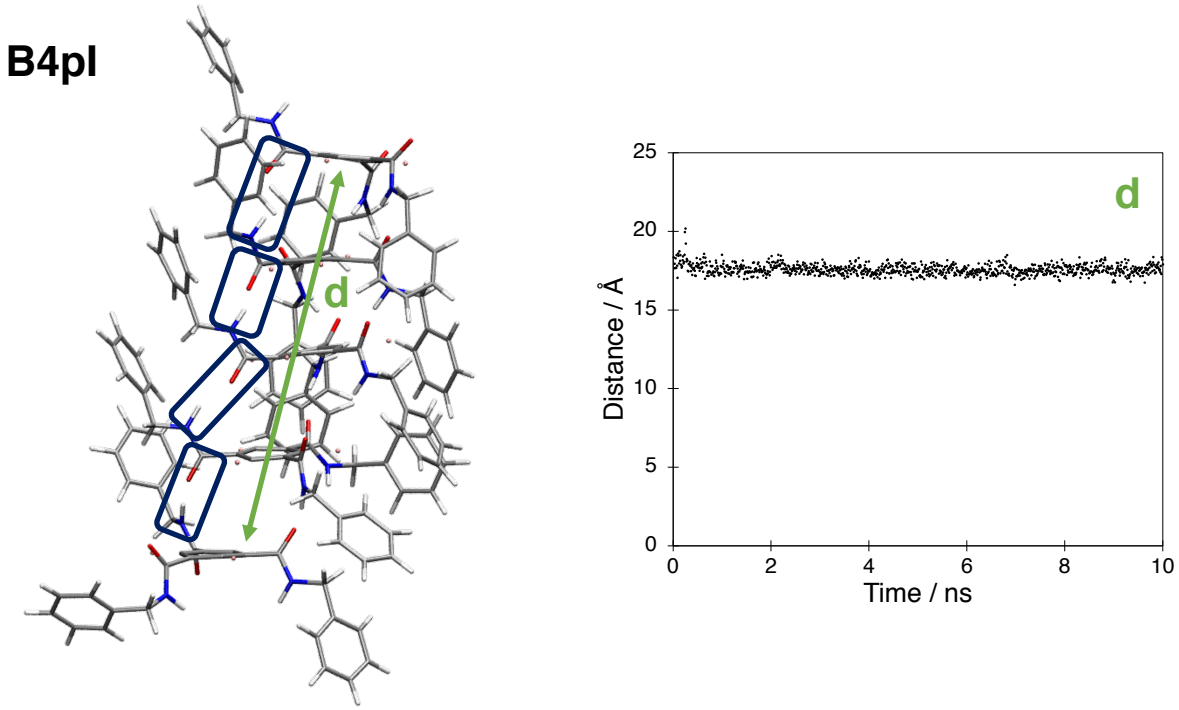

Figure S7: On the left, representative structure of a pentamer involving **B4pI** monomers. h_1 bonds are highlighted within blue frames and the total distance, d, is also indicated. On the right, time evolution of the d distance along the 10 ns simulation time.

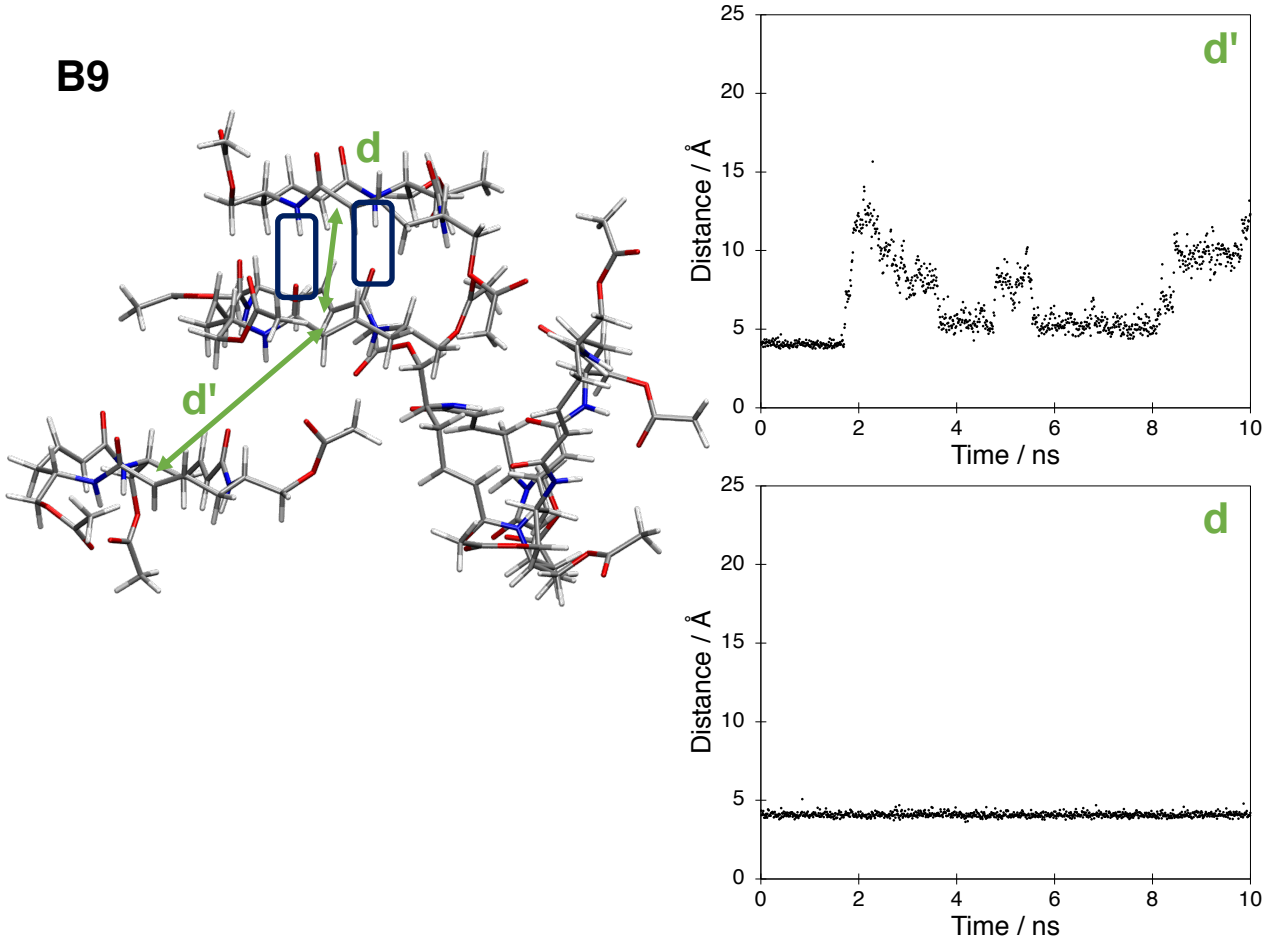

Figure S8: On the left, representative structure of a pentamer involving **B9** monomers. H-Bonds are highlighted within blue frames and the distances, d and d', are also indicated. On the right, time evolution of d and d' distances along the 10 ns simulation time.

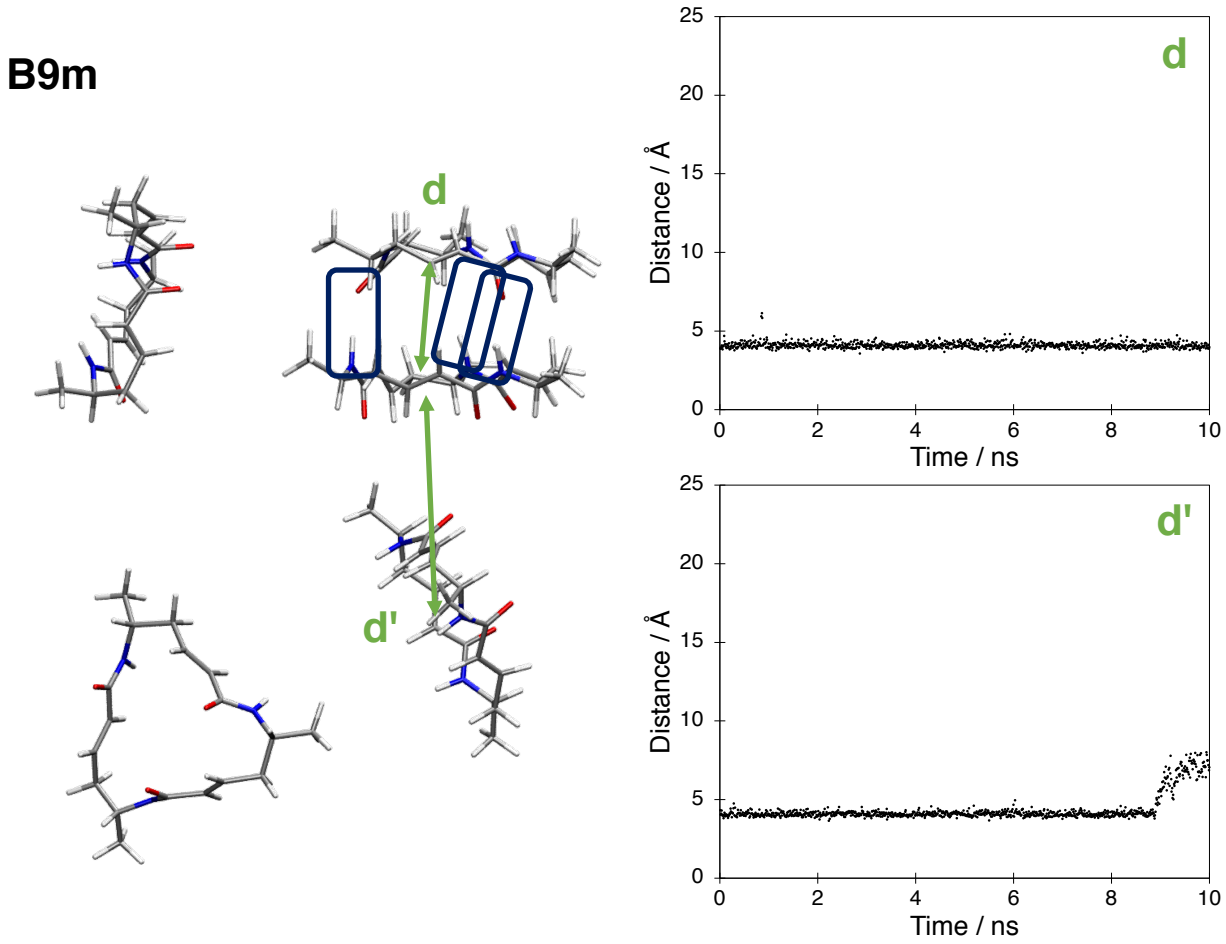

Figure S9: On the left, representative structure of a pentamer involving **B9m** monomers. H-Bonds are highlighted within blue frames and the distances, d and d', are also indicated. On the right, time evolution of d and d' distances along the 10 ns simulation time.

Aggregation process - Long Simulation

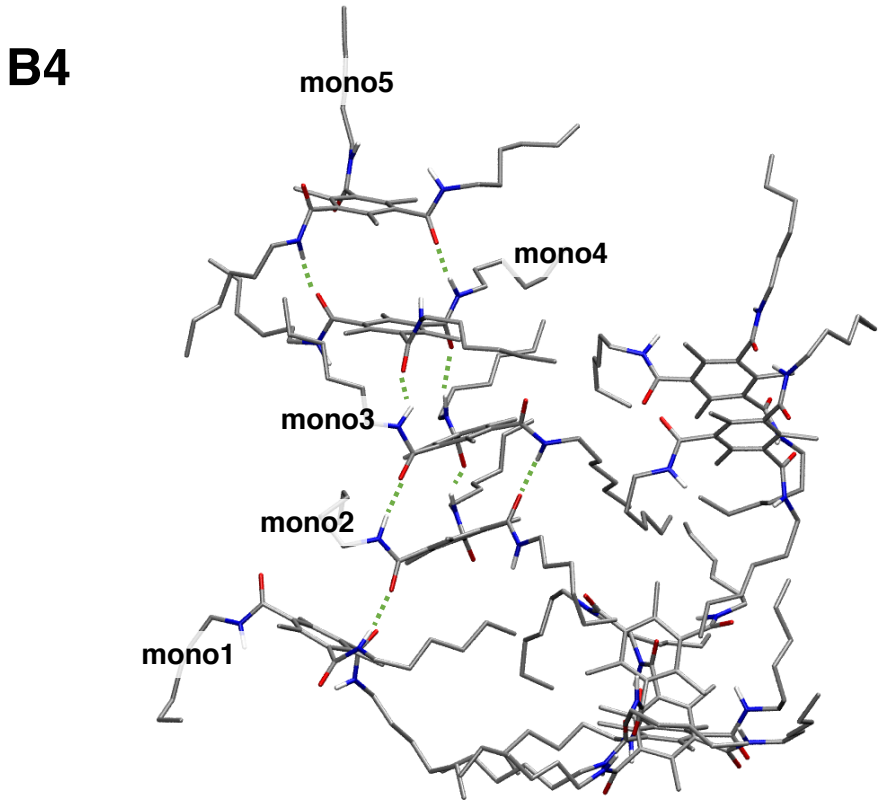

Figure S10: Representative structure of **B4** aggregate extracted from the 100 ns long simulation. For the sake of clarity hydrogen atoms have not been represented, except those of the amide moieties.

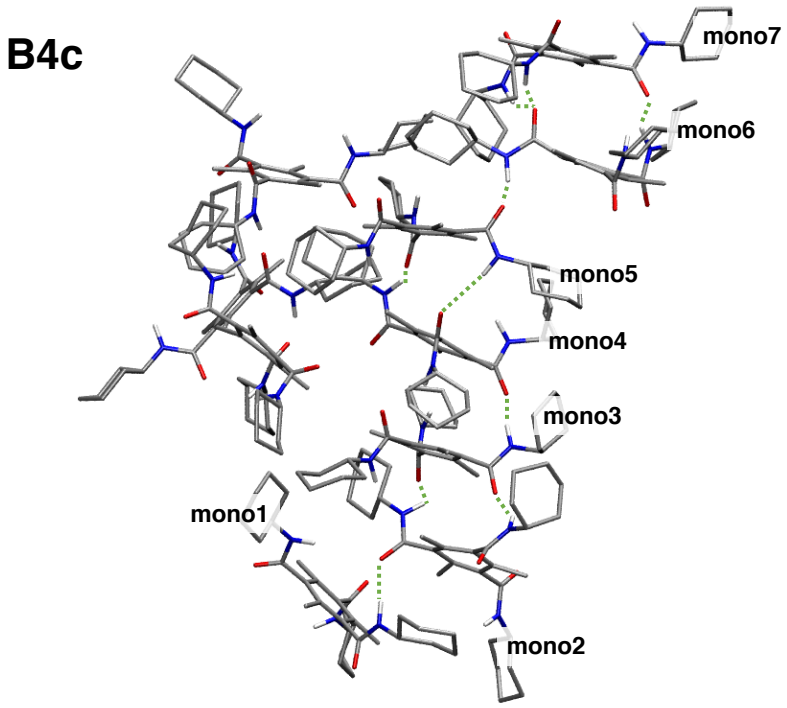

Figure S11: Representative structure of **B4c** aggregate extracted from the 100 ns long simulation. For the sake of clarity hydrogen atoms have not been represented, except those of the amide moieties.

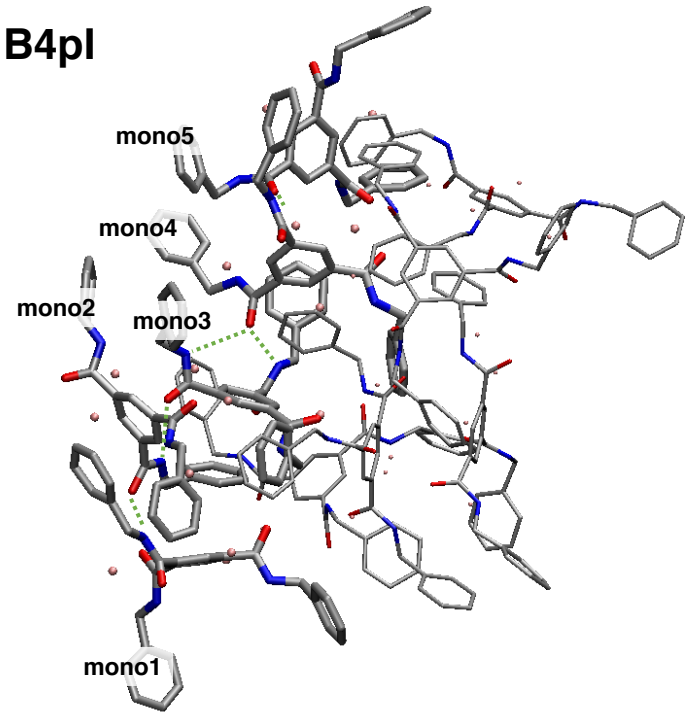

Figure S12: Representative structure of **B4pI** aggregate extracted from the 100 ns long simulation. For the sake of clarity hydrogen atoms have not been represented, except those of the amide moieties.

Aggregation process - Frequency

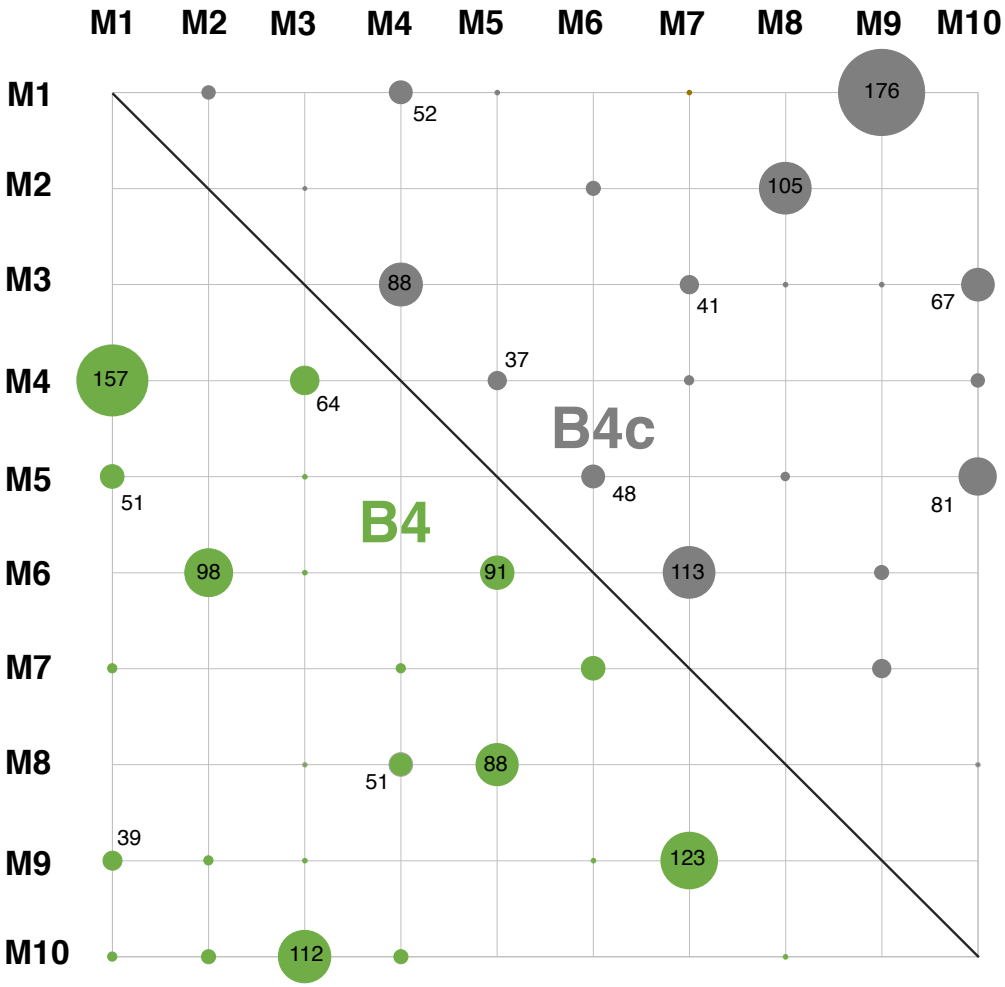

Figure S13: Schematic representation of the occurrence of the formation of H-Bonds between all the monomers (Mi with i= 1,...,10) for **B4** (bottom left, green) and **B4c** (top right, gray) for a total of 100 simulations. The structures that are considered for the count are the final ones obtained at the end of the 1 ns simulation. The size of the dots is proportionnal to the number of H-Bonds that have been detected. For the largest dots, we provided the corresponding exact number of interactions.

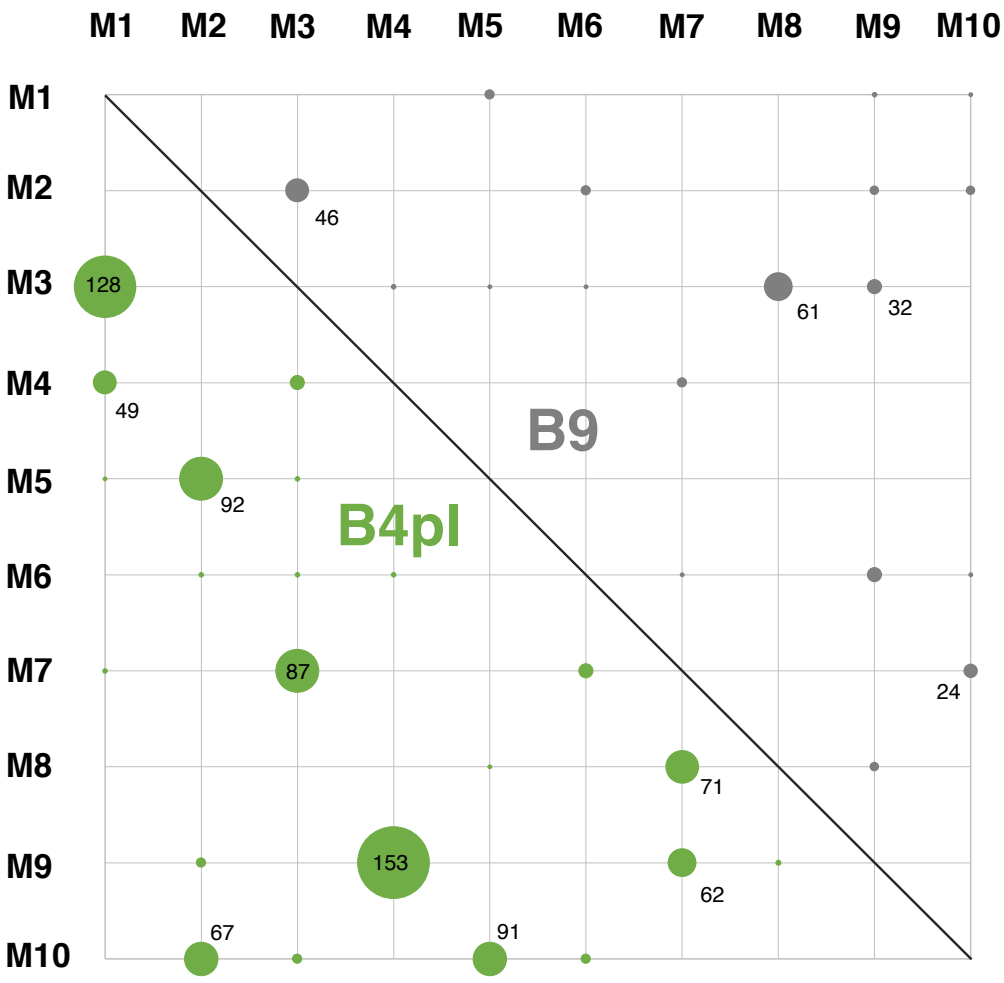

Figure S14: Schematic representation of the occurrence of the formation of H-Bonds between all the monomers (Mi with i= 1,...,10) for **B4pI** (bottom left, green) and **B9** (top right, gray) for a total of 100 simulations. The structures that are considered for the count are the final ones obtained at the end of the 1 ns simulation. The size of the dots is proportional to the number of H-Bonds that have been detected. For the largest dots, we provided the corresponding exact number of interactions.