Computational prediction of the supramolecular self-assembling properties of organic molecules

Le Bras, Laura; Dory, Yves L.; Champagne, Benoît

Published in:
Physical Chemistry Chemical Physics

DOI:
10.1039/d1cp02675e

Publication date:
2021

Document Version
Publisher's PDF, also known as Version of record

PCCP

ROYAL SOCIETY OF CHEMISTRY

PAPER

Cite this: *Phys. Chem. Chem. Phys.*, 2021, **23**, 20453

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

Laura Le Bras, *[a] Yves L. Dory[b] and Benoît Champagne *[a]

Received 14th June 2021,
Accepted 27th August 2021

DOI: 10.1039/d1cp02675e

rsc.li/pccp

Two families of organic molecules with different backbones have been considered. The first family is based on a macrolactam-like unit that is constrained in a particular conformation. The second family is composed by a substituted central phenyl that allows a larger mobility for its substituents. They have however a common feature, three amide moieties (within the cycle for the macrolactam-like molecule and as substituents for the phenyl) that permit hydrogen bonding when molecules are stacked. In this study we propose a computational protocol to unravel the ability of the different families to self-assemble into organic nanotubes. Starting from the monomer and going towards larger assemblies like dimers, trimers, and pentamers we applied the different protocols to rationalize the behavior of the different assemblies. Both structures and thermodynamics were investigated to give a complete picture of the process. Thanks to the combination of a quantum mechanics approach and molecular dynamics simulations along with the use of tailored tools (non covalent interaction visualization) and techniques (umbrella sampling), we have been able to differentiate the two families and highlight the best candidate for self-assembling purposes.

1 Introduction

Supramolecular chemistry is a key concept for many edifices or mechanisms that are essential for life.[1] So it is for some applications going from material sciences to medicine through information storage.[2–6] It has been defined by Lehn, one of its founders, as the chemistry of intermolecular bonds and interactions.[7] Supramolecular chemistry involves a wide variety of weak interactions of different strengths. Metal–ligand interactions are the strongest ones with an interaction energy of about 80 kcal mol^{-1}. Ionic, ion–dipole and dipole–dipole interactions are slightly less stabilizing with an interaction energy ranging from 15 to 50 kcal mol^{-1}. Aromatic interactions, encompassing π–π, π–cation and π-hydrogen bond, lead to a bounding of 3 kcal mol^{-1}. Finally, hydrogen bonds as well as van der Waals forces are the less energetic ones with interactions energies of 1–10 kcal mol^{-1} and lower than 1 kcal mol^{-1}, respectively. As the previously mentioned interactions are non-covalent, it means that the interactions can be reversible and thus confer a kind of flexibility to the system. Nevertheless, increasing the number or combining weak interactions can lead to really stable assemblies and one can cite for example the structure of ice, DNA[8] or the synthetic supramolecular polymer, namely nylon.[9] In both cases, hydrogen bonds (H-bonds) are at the origin of the large stability of the supramolecular assemblies (SMA). The association and dissociation processes within large SMA is ensured by the inter- and intramolecular H-bonds, the best example being the secondary and tertiary structures of proteins that are dictated by weak interactions involving peptide bonds as acceptors (C═O) and donors (N–H).[10–12] The low to medium strength of this type of H-bond, up to 5 kcal mol^{-1}, along with its flexibility, allow SMA to assemble and disassemble easily. One of the specificities of the amide-based H-bonds is their directionality. It has been used to build SMA that are characterized by a stacking of the molecules leading to a one-dimensional columnar SMA.[13,14] Many studies have reported SMA structures where the building block was a substituted benzene with various numbers of amide moieties for the formation of organogels or liquid crystals.[15–18] The number and the orientation of the amide moieties have been shown to

[a] *Unité de Chimie Physique Théorique et Structurale, Chemistry Department, Namur Institute of Structured Matter, University of Namur, Belgium. E-mail: laura.lebras@unamur.be, benoit.champagne@unamur.be*
[b] *Laboratoire de Synthèse Supramoléculaire, Département de Chimie, Institut de Pharmacologie, Université de Sherbrooke, Sherbrooke, Québec, Canada*

† Electronic supplementary information (ESI) available: Validation of GAFF parameters; general parameters for molecular dynamics simulations; particular structure of **B4s** and **B9s** families; hybrid hydrogen bond network; pentamers for **B4**, **B4c**, **B4pI**, **B9**, and **B9m**; interaction maps for **B4**, **B4c**, **B4pI** and **B9**. See DOI: 10.1039/d1cp02675e

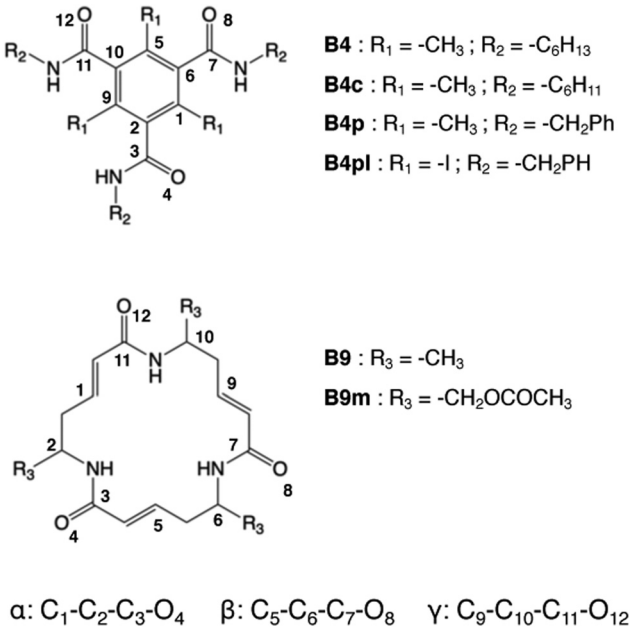

Fig. 1 Representation of the structures of the four molecules of the **B4s** family, **B4**, **B4c**, **B4p** and **B4pI** and of the two molecules of the **B9s** family, **B9** and **B9m**. The three important dihedral angles are also defined.

be crucial for the effectiveness of the formation of the stacking.[13] Another building block, with a backbone based on a macrolactam-like unit in which the amide moieties are incorporated, has been reported for the formation of columnar SMA.[19–21]

In this study we propose to consider two families of organic molecules that have different backbones. On the one hand, a 18-membered lactam ring with a C_3 symmetry and three amides conjugated to an alkene, has been considered (**B9s**, Fig. 1). Due to the conjugation of the enamides, the structure is globally flat. The *E* conformation of the alkene induces a constrained conformation for the entire molecule. One has to notice that the three amides are perpendicular to the mean plane of the molecule. This type of molecule is likely to stack as infinite supramolecular assemblies through backbone–backbone H-bond interactions and van der Waals contacts. Due to the involved moieties, the periodicity of the stacking, that is to say the distance between two identical atoms of two consecutive molecules, is about 4.8 Å. This is for information the same gap as the one observed between the two strands of β-sheets within proteins. Other scaffolds allow this kind of stacking, such as substituted benzenetricarboxamide (BTA) molecules (**B4s**, Fig. 1). The presence of the three amide moieties will lead to the same kind of stacking and the possibility to add substituents (position R_2, Fig. 1) with various lateral chains can be considered for targeted applications and properties. For a simple BTA molecule, with R_1 = H and R_2 being a benzyl group, the stacking distance is shorter than the one observed for macrolactam-like stacking (4.8 Å) with a value of about 3.6 Å.[22] This can be explained by the fact that the amide moieties are slightly twisted by 40 on average with respect to the plane of the molecule.[23] Our strategy to make **B4s** more similar to **B9s** is to add bulky groups on position R_1 (–CH_3 for **B4**, **B4c** and **B4p**) and –I for (**B4pI**) in order to make the amide moieties orthogonal to the plane of the molecule to have a periodicity close to 4.8 Å. This will make the two families, **B4s** and **B9s**, similar in terms of stacking periodicity and orientation of the side chains. This type of infinite stack, either based on a macrolactam-like unit or a substituted central phenyl ring, can be considered as a platform on which one can attach side chains (positions R_2 and R_3, Fig. 1), the nature of which will determine the global properties.[24] Lipophilic alkyl chains may lead to liquid crystals,[25,26] lateral chains based on pseudopeptides can be used for the development of metal complexation agents[27] and fluorinated substituents for the development of MRI contrast agent.[28]

Finally, six molecules will be studied (Fig. 1). Four of them will be part of the **B4s** family and the two others will form the **B9s** family. They will differ in the nature of the lateral chains with for example for R_2 position: –C_6H_{13} alkane chain for **B4**, –C_6H_{11} cyclic group for **B4c** or –CH_2-Ph moiety for both **B4p** and **B4pI**. For **B9s**, R_3 position will be substituted by a –CH_3 for **B9m** and a –CH_2OCOCH_3 for **B9**. Our idea is thus to compare those two apparently C_3-symmetric backbones for an efficient formation of columnar stacking involving amide-based H-bonds with an ideal stacking distance of 4.8 Å.

Both those two families or close derivatives have already been studied either from an experimental or a computational point of view even if no direct comparison is available. Experimental studies provide information on the crystalline structures,[29] the behavior in solvated environments[30,31] and on the association process.[32,33] Some of the studies are also dedicated to the evaluation of chiral properties of such SMA.[23,34] From the computational point of view, studies involving calculations based on quantum mechanics (QM) or molecular dynamics (MD) simulations are available.[35] They provide insights into the atomistic description of the bonding between the monomers and quantitative analysis of the thermodynamics of the SMA with reported binding free energies.[36,37] In a related field, there are studies that are combining experimental and computational approaches (QM and/or MD) to assess the flexibility of organic molecules but when they are bind to carbon-based nanotubes.[38,39] In this study we propose to combine approaches that have already been used separately, namely QM calculations and MD simulations, in order to provide a complete picture of the self-assembling properties of those molecules based on the formation of a hydrogen bond network. QM calculations will provide the atomistic description of the weak interactions while MD simulations will highlight the dynamic of the association process. Recent studies have already shown that a fully computational approach, based on MD simulations and QM calculations, was able to unravel hydrogen bond interactions between organic molecules within supramolecular assemblies.[40] In parallel to QM and MD approaches, tailored techniques and tools are used to complete the understanding of the underlying interactions. Firstly, umbrella sampling, that is particularly relevant for the analysis

of supramolecular interactions, will be used in order to compute the binding free energy of the different SMA.[41–44] Secondly, non-covalent interactions (NCI) visualization will also be considered as it has been shown to be a valuable tool that can illustrate weak interactions such as hydrogen bonds or van der Waals interactions within SMA.[45,46] Finally, by combining those approaches, techniques and tools and by adopting a progressive and systematic approach for each of the molecules (Fig. 1), that is to say going from monomers to larger assemblies (decamers), it will be possible to unravel the parameters that are essential for the efficient formation of a columnar stacking.

2 Computational details

2.1 Quantum mechanics

2.1.1 Single point calculations. All the calculations have been performed using Gaussian16 package[47] within the density functional theory (DFT) framework. We used the ωB97X-D range-separated hybrid exchange correlation functional (XCF) combined with the 6-311+G(d,p) atomic basis set.[48] This XCF is known as one of the most efficient to consider structures and energies of assemblies involving hydrogen bond interactions.[49] The solvent (water) was modeled using an implicit solvation model, namely the polarizable continuum model (PCM).[50] In order to describe large assemblies and to reduce the computational time, we relied on a QM/QM′ ONIOM approach. Within this model it is necessary to define 2 subsystems. The model system, composed of the central phenyl and the amide groups for **B4s** and the macrolactam-like unit for **B9s**, will be described at both the high level of theory (ωB97X-D/6-311+G(d,p)) and the low level (HF/3-21G(d)) of theory. The real system, encompassing the model system and the lateral groups of both families, will only be treated at the low level of theory. A charge embedding framework was also added to the hybrid QM/QM′ ONIOM scheme.[51]

2.1.2 Non-covalent interactions. It has been possible to visualize non-covalent interactions (NCI) through the use of NCIPLOT code.[52] NCI analysis gives rise to an index that is based on the calculated electronic density and its reduced gradient, represented as a two-dimensional plot. For a given system, there will be a drastic change in the reduced density gradient (RDG) between the atoms that are interacting, leading to density critical points. The latter can be represented on the molecular structure as an isosurface to indicate the region where a weak interaction is occuring. Nevertheless, both attractive (H-bonds, van der Waals) and repulsive (steric repulsion) interactions can be spotted thanks to this index. By looking at the second derivative of the density and to the sign of its eigenvalue it is possible to distinguish attractive and repulsive interactions. Hence the density and the sign of the eigenvalue of the density second derivative give information about the strength and the type of interaction respectively and one can visualize them *via* isosurfaces. The electronic density used to compute the NCI index is the one calculated at the same level of theory as the one presented in Section 2.1.1.

2.2 Molecular dynamics

2.2.1 Classical molecular dynamics. Molecular dynamics (MD) simulations were run with the generalized AMBER force field (GAFF)[53] within GROMACS 2018.3 package[54] for both organic molecules and water molecules (TIP3P). As GAFF is not directly implemented in GROMACS, we used acpypi script[55–57] to convert the files from AMBER to GROMACS formalism. For each molecule, the atomic charges were derived following the parametrization procedure in GAFF, that is to say using HF/6-31G(d) RESP charges. The validity of the force field was checked by comparing the structures obtained after an optimization process in vacuum with GAFF and with DFT at the ωB97X-D/6-311+G(d,p) level. Results are provided in the ESI† section. During the simulations, the system is composed by n organic molecules (n being equal to 1, 2, 3, 5, and 10 and the organic molecules being either from **B4s** or **B9s** families). The general philosophy of our molecular dynamics simulations is represented schematically on Fig. 2. The size of the simulation box and the number of water molecules depend on the system under investigation. All those information for each system under investigation are gathered in ESI.†

To describe the electrostatic interactions, periodic boundary conditions were imposed along with a cut-off of 10 Å and the use of the Particle Mesh Ewald (PME) method.[58,59] Following a steepest descents minimization, each system was equilibrated in two steps. For the first step, a simulation in the canonical ensemble (*NVT*) during 100 ps was carried out. The temperature was set to 310 K using the Berendsen weak coupling method.[60] Organic molecules and solvent were coupled to separate temperature coupling baths. For the following second step, simulation under constant pressure (*NPT*) was performed. To maintain an isotropic pressure of 1 bar, we relied on the previously mentioned weak-coupling Berendsen method. The production phase was then carried out in the same *NPT* ensemble with a time step of 2 fs. Temperature was controlled thanks to the Nosé–Hoover thermostat (τ = 1 ps) while the isotropic character of the pressure was maintained *via* the

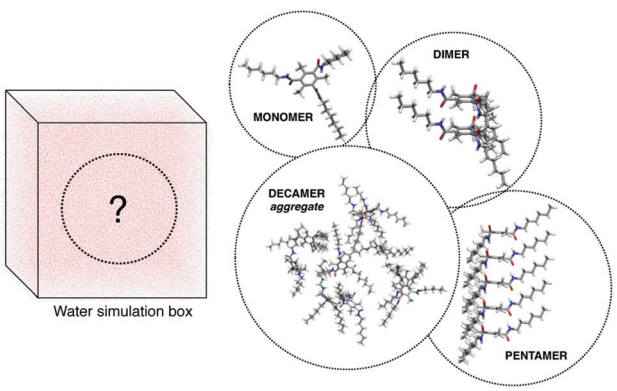

Fig. 2 Schematic representation of our strategy for molecular dynamics simulations. Within a simulation box full of water molecules we will successively consider either monomers, dimers, pentamers or aggregated decamers.

Parinello–Rahman barostat (τ = 1 ps). Combining this thermostat and barostat ensures the presence of a true *NPT* ensemble.[41,61–64] Simulation time, if not explicitly precised, was set to 10 ns.

2.2.2 Umbrella sampling. Within the umbrella sampling (US) approach we have considered dimers, trimers and pentamers. The systems were placed in a rectangular box that can allow a pulling simulation (*e.g.* a simulation box that is too small will lead to an interaction with the periodic images). In particular, the z length had to be large enough in order to satisfy the minimum image convention. The solvent molecules (water), were described through the TIP3P model. The first step consisted in an *NPT* equilibration of 100 ps, as it was described above. For the proper pulling simulation, restraints were applied to one of the monomers for dimers, to a dimer for trimers and to a tetramer in the case of the pentamer. The molecules that were restrained were thus considered as immobile references. The molecule that was not restrained was then pulled away from the immobile one, along the z-axis over 500 ps at a rate of 0.1 nm ps^{-1} with a spring constant of 250 kJ mol^{-1} nm^{-2}. The final COM (center of mass) distance between the two considered assemblies that was obtained was 4 nm. Snapshots were extracted from this pulling simulation in order to be as many starting points for the different umbrella sampling windows. For COM distances under 1 nm, a separation of 0.05 nm was considered between each window and then, for COM distances above 1 nm and up to 2.5 nm, the spacing between the windows was 0.1 nm. An example of the corresponding histogram is provided in ESI.† It allows a smoother and more accurate description of the interaction at small COM distances. This approach leads to around 25 windows. For each window, a 10 ns simulation was performed, resulting in a total simulation time of 250 ns for the US approach, for each assembly of each molecule. The analysis of the results was done using the Weight Histogram Analyzis Method (WHAM), implemented in the GROMACS 2018.3 package.[65] A schematic representation of the umbrella sampling approach is provided in Fig. 3.

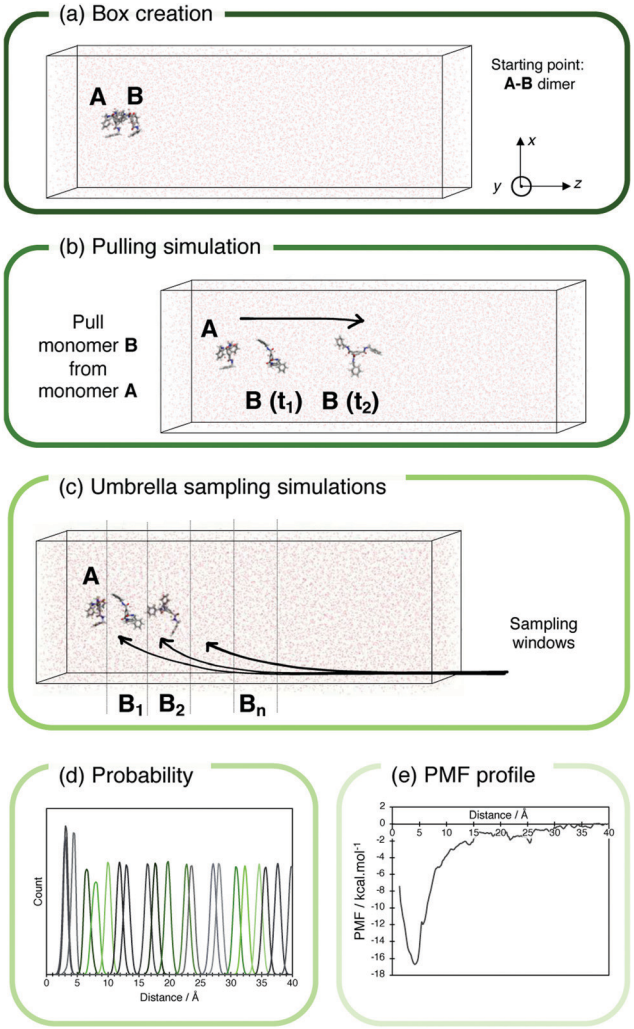

Fig. 3 Schematic representation of the umbrella sampling approach. With (a) the starting configuration for a dimer. The dimer is positioned voluntarily on one side of the simulation box and the z-axis is voluntarily set to a large value to allow the next pulling step. (b) Pulling simulation. B is pulled away from A at a constant rate. (c) Along the pulling simulations, different configurations with different A–B distances (A–B_1, A–B_2) are extracted in order to be as many starting configurations for different sampling window simulations that will be centered on the A–B distance. (d) For each sampling window, we count the number of times that the A–B distance is equal to each value. (e) A potential of mean force can be reconstructed and can be associated to the free binding energy, ΔG_{bind}.

3 Results and discussion

3.1 Monomers

Key structural parameters (see Fig. 1 for their definitions) have been selected to study the different monomers. For both **B4s** and **B9s** families, three dihedral angles were defined (α, β, and γ), illustrating the relative position of the amide groups with respect to the plane of the molecule. We followed their evolution along the 10 ns MD simulations. The average values are reported in Table 1.

For the **B4s** family, as previously mentioned, the three amide moieties, that are orthogonal to the plane of the molecule, do not point towards the same side of this plane. The particular structure of **B4s** family is provided on Fig. S3 in ESI.† As illustrated by α that is negative, there is one amide moiety that is antiparallel to the two other ones for **B4s**. This disymmetry in the orientation of the three amide moieties has been shown to be favorable for the formation of a columnar stacking.[37] For the **B9s** family, all the amide groups are oriented in the same direction. During the MD simulations, all the dihedral angle values are close to ideal value of 90 and no complete rotation of the amide moieties was observed. It indicates that all the amide moieties are and stay perpendicular to the plane of the molecule. Adding bulky groups on position R_1, instead of having only a hydrogen atom,[22] is thus a good strategy for inducing and imposing orthogonality to the three amide moieties of **B4s**. Nevertheless, considering the standard deviation associated to those values, **B9** and **B9m** show a smaller conformational

Table 1 Average values of the selected dihedral angles (in degrees) defined on Fig. 1 along with the RMSD value (in Å) for each of the molecules. Standard deviations are provided in parentheses

	α	β	γ	RMSD[a]
B4	−90 (13)	91 (14)	88 (13)	2.33 (0.48)
B4c	−90 (13)	89 (14)	89 (13)	2.08 (0.48)
B4p	−89 (13)	90 (13)	90 (13)	3.07 (0.56)
B4pI	−88 (14)	90 (12)	89 (13)	3.12 (0.51)
B9	95 (7)	95 (6)	96 (8)	0.44 (0.12)
B9m	83 (7)	83 (7)	84 (8)	0.80 (0.13)

[a] RMSD calculations were performed without considering lateral chains and hydrogen atoms. It leads to an amount of 18 atoms for **B4s** family and 21 for **B9s** family.

flexibility with deviations around 7 while this value increases up to 14 for the **B4s** family. To confirm this hypothesis, we computed the root mean square deviation (RMSD) of the different structures along the simulation, the reference structure being the ideal one with perfectly perpendicular amide groups. One has to notice that for the RMSD calculation, the hydrogen atoms were not considered and nor were the lateral chains of all the molecules. The results are gathered in Fig. 4 and Table 1.

The two families can be clearly distinguished. RMSD values are lower than 1 Å for **B9** and **B9m** while they are systematically larger than 1.5 Å for **B4s** family. It indicates that during the simulations, **B4s** molecules are more flexible than the **B9s**. Due to the inclusion of the amide groups within the macrolactam-like backbone of **B9s**, there is almost no degree of freedom for the molecule, as illustrated by the low average RMSD value for this family (0.44 and 0.80 Å for **B9** and **B9m** respectively). Fluctuations are also small with deviations equal to 0.12 and 0.13 Å for **B9** and **B9m** respectively. On the other hand, for **B4s** molecules, the average RMSD is ranging from 2.08 Å to 3.12 Å for **B4c** and **B4pI** respectively, indicating quite large structural modifications during the simulation. Standard deviations confirm this trend with values around 0.50 Å for the four molecules of **B4s** family. By studying the six different monomers in a polar solvent it has been possible to differentiate the two families on the basis of a conformational flexibility related to the orientation of the three amide moieties. For **B9s**, due to the cyclic structure that is encompassing the amides, there is a global conformational constraint for the entire molecule, explaining low RMSD values. For **B4s**, conclusions are different. The fact that the amide moieties are only substituents confers them a limited constraint-free movement. This freedom is only limitated by the presence of the $-CH_3$ and –I groups that prevent total rotation of the amides. This fundamental difference between the two families may have an impact on the self-assembling properties. The next sections are precisely dedicated to the comparison of the two families for the formation and the stabilities of these different SMAs.

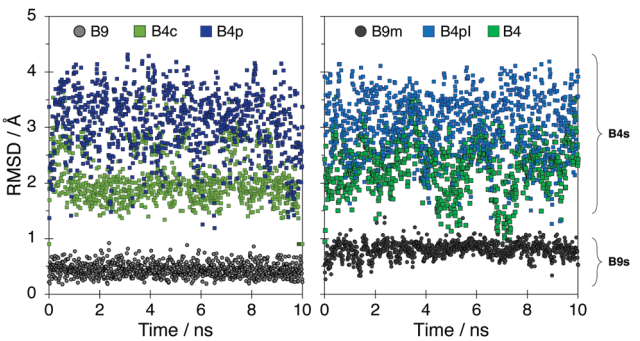

Fig. 4 Evolution of the RMSD values, in Å, for the considered molecules along the simulation time. Hydrogen atoms were not considered for the calculation of the RMSD.

3.2 Dimers

3.2.1 Qualitative approach. To study the possible self-assembling behavior of the different molecules, the smallest and thus simplest supramolecular entity was considered, namely a dimer. The interaction between the two stacking monomers is ensured by their amide moieties. Indeed it is possible to form a hydrogen bond involving the N–H part of the amide of one monomer and the C=O bond of the neighbouring monomer. As previously observed during the study of the monomers, the amide groups are perpendicular to the plane of the molecule and it is thus possible to stack monomers *via* a network of three hydrogen bonds. The dimers that have been built manually by stacking two monomers are represented on Fig. 5. Note that equivalent H-bonded structures (and dynamics) were obtained when starting from random orientations.

MD simulations were performed for each dimer. Three new key structural parameters were introduced, namely h_1, h_2, and h_3, which describe the three hydrogen bonds and d, that defines the distance between the centroids of the central phenyl for **B4s** and of the macrolactam-like unit for **B9s**. The average values for each dimer are gathered in Table 2.

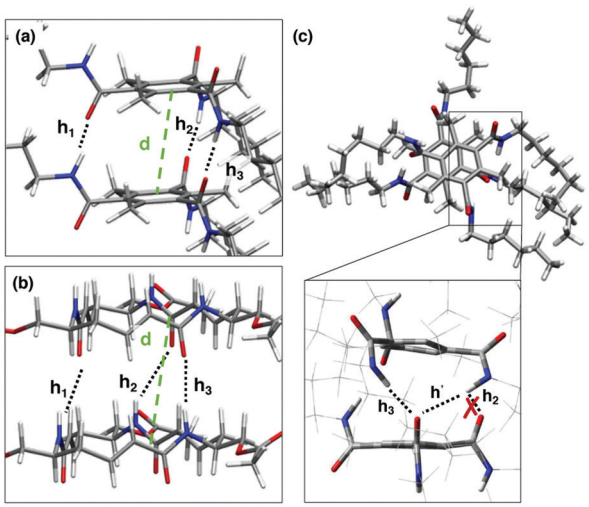

Fig. 5 Representation of the stacking of the dimers for the (a) **B4s** and (b) **B9s** families. Key distances are also identified for both stackings. (c) Representation of a particular stacking observed for **B4s** and definition of a new hydrogen bond, h'.

Table 2 Average values, in Å, for key distances of the different dimers. The mean deviations are also reported. The corresponding simulation time, in ns, during which the averages have been calculated is also provided in parentheses

	B4	B4c	B4p	B4pI	B9	B9m
d	4.23 ± 0.30 (10)	4.12 ± 0.23 (10)	4.11 ± 0.30 (10)	3.77 ± 0.27 (10)	—	3.82 ± 0.23 (10)
h_1	2.05 ± 0.23[a] (10)	2.05 ± 0.34 (10)	1.99 ± 0.24 (10)	2.01 ± 0.20 (10)	—	2.03 ± 0.22 (10)
h_2	2.49 ± 0.66[a] (8.4)	2.81 ± 0.61 (7.2)	2.49 ± 0.58 (8.9)	2.59 ± 0.71 (3.0)	—	2.01 ± 0.18 (10)
h_3	2.54 ± 0.64[a] (3.1)	2.08 ± 0.35 (1.1)	2.62 ± 0.63 (3.2)	2.64 ± 0.67 (6.4)	—	2.02 ± 0.21 (10)
h'	2.44 ± 0.60[a] (1.2)	3.17 ± 0.66 (0.5)	2.24 ± 0.45 (9.3)	2.26 ± 0.49 (6.3)	—	—
h''	2.39 ± 0.53[a] (7.7)	2.20 ± 0.50 (8.8)	—	2.36 ± 0.61 (2.3)	—	—

[a] For the calculation of the average value and the associated standard deviation, only the areas highlighted on Fig. 6 are considered. The same procedure is used for the other molecules, averages are made when the bonds are effective during the corresponding time.

By looking at the average value of d along the MD simulations it is already possible to have a qualitative idea of the stability of the different assemblies. If all the dimers involving monomers from the **B4s** family are stable, it is not the case for the **B9s** one. Indeed, the d values that are reported in Table 2 illustrate the fact that the two monomers involved within a dimer are remaining close to each other. If we go deeper in the analysis, we observe that the average value of d, ranging from 3.77 Å to 4.23 Å for **B4s** and **B9m**, is within the range of what is obtained experimentally in a crystalline environment (4.8 Å)[20,29,66] and what is obtained through other computational studies in the gas phase (3.7 Å,[36] 3.4 Å[37]) or in solvents (around 5 Å[35] and 3.4 Å[22]). However, one has to notice that the previously reported values are for close derivatives of **B4s** and **B9s**. Actually, both **B4s** and **B9s** have not been reported before in the literature.

The standard deviations presented in Table 2, around 0.2–0.3 Å, illustrate the fact that the dimers are not "broken" along the simulation. As the dimer involving **B9** monomers was not stable, it was not relevant to report d along the simulation.

One has to notice that if h_1, h_2, and h_3 are equivalent for **B9s** family, they are not for **B4s** dimers. Because of the asymmetry in the orientation of the amide moieties, only h_2 and h_3 are equivalent while h_1 is unique. Looking at those different values gives insights into the way the different monomers are interacting between each other and the interactions involved in the stability of the different dimers. We have represented on Fig. 6 the evolution of h_i distances for **B4**.

It is clear from Fig. 6(a) that the hydrogen bond h_1 is always effective during the 10 ns simulation. Moreover, with an average value of 1.99 ± 0.24 Å this hydrogen bond is within the range (1.5–3.5 Å) of a hydrogen bond with a medium strength (1 to 4 kcal mol^{-1}). The same conclusion can be drawn for all the dimers of the **B4s** family. The analysis is quite trickier for h_2 and h_3. As represented on Fig. 6b and c, those H-bonds are not always effective. The area that are colored represent the moment where h_2 and h_3 distances are within the 1.5–3.5 Å range. The time corresponding to the colored area are given in

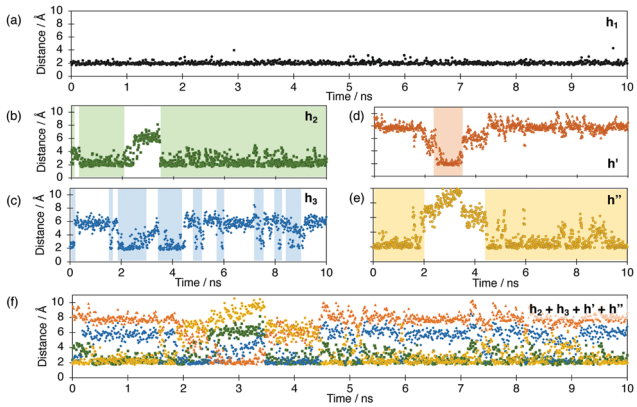

Fig. 6 Evolution for **B4** dimer of the distances (a) h_1, (b) h_2, (c) h_3, (d) h', (e) h''. On (f), h_2, h_3, h', h'' distances have been gathered. All the distances are in Å.

Table 2. If h_2 is effective during almost the 10 ns simulation (8.4 ns), it is not the case for h_3, that is only effective transiently for 3.1 ns. Nevertheless, the dimer of **B4** is stable and the stacking is maintained during all the simulation (d distance gives this information). By looking closer to the structure along the trajectory it has been possible to highlight a particular structure (Fig. 5(c)). During the simulation the two monomers do not stay aligned. There is a tilt of one monomer with respect to the other one, inducing a hybrid hydrogen bond. The keto part of one amide points between two amino moieties of two other amides of the other monomer. One expected H-bond is thus presents (h_2 or h_3), even if it is longer than a perfectly aligned H-bond, while the other one is new. We have thus introduced two new H-bonds, namely h' and h'' (Fig. 5(c)). Their evolutions are represented on Fig. 6d and e. One can observe that either h' and h'' is effective during the simulation. Their average values along with their effective time are gathered in Table 2. It is possible to say that the H-bonds complement each other. When h_2 is not effective, one can observe that h' is. When h_3 is not effective, h'' is taking over. We decided to combine

those four H-bonds on the same graphics (Fig. 6(e)). It appears that there is always at least one but most of the time two hydrogen bonds (h_2, h_3, h', or h'') that are present within the dimer, ensuring its stability along with h_1. The Boltzmann population ratio for the normal (with h_1, h_2 and h_3) and the hybrid (with h_1 and h_2/h'' or h_3/h') dimer (see Fig. S4 in ESI†) is always around 40/60 for all the dimers of the **B4s** family. Indicating that the hybrid dimer is more favorable than the normal and expected one. For the other dimers of the **B4s** family, the same behavior is observed with a compensation of h_2 and h_3 by h' and h'' to ensure the global stability of the dimers. For **B9m** dimer, h_1, h_2, and h_3 are effective during all the trajectory but no hybrid H-bond can be observed as the orientation of the amide moieties is conformationally constrained within the structure of the macrolactam-like molecule.

To summarize the findings about dimers, we have been able to highlight (i) the (non-)stability of **B4s** family and **B9m** (**B9**), (ii) the effectiveness of the hydrogen bond network to build supramolecular assemblies and (iii) a new hybrid H-bond pattern allowed by the conformational flexibility and the possibility for **B4s** dimers to orient their amide moieties. One has to notice that, contrary to what has been reported before, our dimers are showing a different stacking patterns with one pure and one hybrid H-bond instead of three pure H-bonds.[22] In particular, this hybrid interaction has been shown to be dynamic as it can be broken and reformed.

3.2.2 Quantitative approach: umbrella sampling. To go further in assessing the stability of the different dimers, we have undertaken US simulations. It allows us to retrieve the binding free energy, ΔG, along a reaction coordinate, x, that represents the preferential direction for the stacking pattern. Using around 25 sampling windows along this axis, one can construct a one-dimensional potential of mean force (PMF) profile for each system under study, leading to a binding energy, E_{bind}. US simulations for larger assemblies, namely a trimer and a pentamer, were also performed. For each PMF profile, the minimum energy is associated to a particular distance, d_{com}, that can be roughly compared to the d distance discussed in the previous section as it is the distance between the center of masses of the different assemblies. All those values, E_{bind} and d_{com}, for each system and for each molecule, are gathered in Table 3. For the following discussion, the comparison and evolution of E_{bind} will be made on the absolute value.

Table 3 Computed binding free energy, E_{bind}, in kcal mol^{-1} for different assemblies and the corresponding optimal interaction distance, d_{com}, in Å

	Dimer		Trimer		Pentamer	
	E_{bind}	d_{com}	E_{bind}	d_{com}	E_{bind}	d_{com}
B4	17	4.2	21	4.3	15	4.3
B4c	24	4.2	17	4.0	12	4.1
B4p	12	4.2	12	3.9	8	4.1
B4pI	21	3.6	23	3.9	22	3.8
B9	7	3.8	9	3.8	7	4.2
B9m	12	4.2	6	4.0	4	4.1

The binding energy that are obtained for the **B4s** dimers are globally higher than the one for the **B9s**. They are equal to 17, 24, 12, and 21 kcal mol^{-1} for **B4**, **B4c**, **B4p**, and **B4pI** respectively while they are equal to 7 and 12 kcal mol^{-1} for **B9** and **B9m**. It indicates that it is more stabilizing for **B4s** to form dimers than for **B9s**. Moreover, even if a new dimer structure has been highlighted and appears to be favorable, the binding energies that are found are comparable to what is already reported for similar molecules involving such H-bond pattern[37]. One can notice that the US d_{com} values are consistent with the MD average values, with a distance around 4.0 Å between the two monomers, corresponding to a classic stacking distance for this kind of molecules.

The question that motivated the new simulations involving trimers and pentamers is the following: how is the binding energy evolving when the supramolecular assembly is getting larger? More specifically, is it getting harder to add a monomer to a dimer, a tetramer? In other words, is there a cooperative effect occurring within those SMA? The answer to this question will help us to understand the self-supramolecular assembling behavior of **B4s** and **B9s** into larger assemblies. The same umbrella sampling approach presented before was used to study first the interaction within a trimer. Two subsystems were considered, a dimer and a monomer and we were looking for the binding energy of the monomer with the dimer. For the pentamer, the binding energy is computed for the interaction of a tetramer and a monomer. The PMF profile for trimers and pentamers, for each molecule, is provided on Fig. 7. For **B4s** family, the three possible different behavior are observed, a slight increase of E_{bind} for **B4** (+4) and **B4pI** (+2), a decrease for **B4c** (−7) and no evolution for **B4p**. One can nevertheless notice that the formation of trimers still remains favorable in all the cases. For **B9** and **B9m**, E_{bind} is slightly increasing (+2) and decreasing (−6) respectively. In conclusion, adding a monomer to an already formed dimer is more favorable for **B4s** than for **B9s**. Trimers of **B4s** appears even more stable than dimers of **B9s**. Going further and considering pentamers for both families leads to a unique conclusion. The binding energy is always decreasing when compared to the energies obtained for dimers and trimers. Nevertheless, the values obtained for **B4s** (15, 12, 8, and 22 kcal mol^{-1} for **B4**, **B4c**, **B4p**, and **B4pI** respectively) are still higher than the ones obtained for **B9s** (7 and 4 kcal mol^{-1} for **B9** and **B9m** respectively), indicating that adding a monomer to a dimer of B9s family will be less favorable than for the B4s family and the same appears when adding a monomer to a tetramer.

The conclusion that can be drawn after the US simulations is that, contrary to other studies,[15,67] no clear cooperative effect can be observed for either **B4s** or **B9s** families. In most of the cases for **B4s** assemblies, adding new monomers to small oligomers is not strongly favorable nor unfavorable. It indicates that the formation of small oligomers to larger assemblies may be possible, regardless of the already existing building block. For **B9** and **B9m**, the different binding energies indicate that either the different assemblies are poorly favored (**B9**) with low E_{bind} or are getting unfavorable with a constant decrease of E_{bind} (**B9m**).

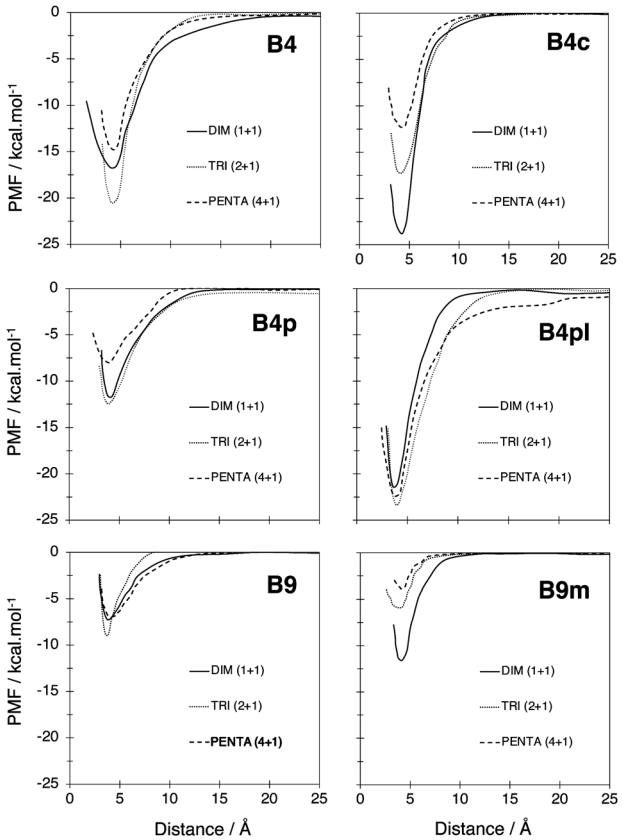

Fig. 7 Potential of mean force (PMF) curves for each of the molecules in different configurations: dimer (bold line), trimer (dotted line) and pentamer (dashed line).

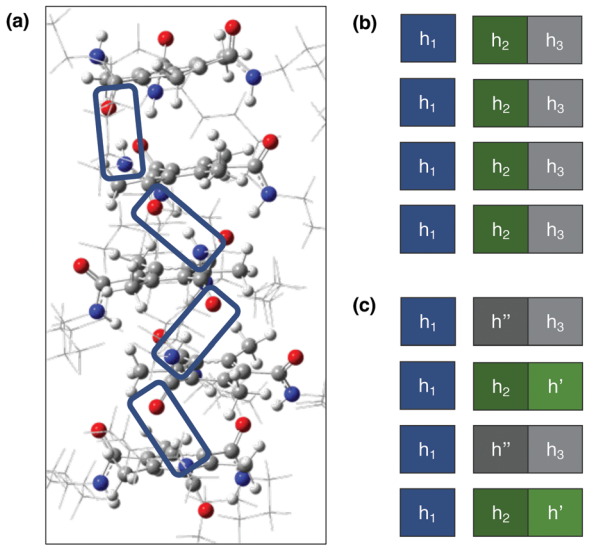

Fig. 8 (a) Representation of the pentamer structure of **B4p**. h_1 bonds are also highlighted by the blue frames to illustrate the tilt of the amide moieties. Schematic representation of the H-bond framework between the monomers with (b) the expected framework with perfectly stacked monomers and (c) the framework that is observed during the simulations with the formation and alternation of hybrid bonds.

3.3 Pentamers

Pentamers were built manually in the same way as dimers and classical molecular dynamics simulations have been performed for each of the molecules. Two behaviors were observed for the two families. For **B9s**, almost no H-bond interactions were maintained throughout the trajectory leading to non-stability for both of the assemblies. Though, dimers were observed ponctually (see Fig. S8 and S9 in ESI†). The conclusions that were drawn in the previous sections are thus confirmed with (1) a low but still possible stability for dimers and (2) an unfavorable binding energy for systems involving a large number of monomers.

For **B4s** family the conclusions are different with respect to **B9s**. Indeed, for the entire **B4s** family, the pentamers appeared as stable along the trajectory (Fig. 8 for **B4p** and Fig. S5, S6 and S7 in ESI† for **B4**, **B4c**, **B4pI** respectively). The H-bond network is at the origin of this stability. As for the dimers, h_1 is always effective and can thus be considered as the backbone of the entire supramolecular assembly. Within the pentamer, amide moieties are no longer perpendicular to the plane of the molecules, leading on average to a tilt of 15 (Fig. 8(a)). The fact that the amide moieties are quite flexible also allows the formation of the previously mentioned hybrid bonds. During the simulation, there is an alternation of h_2/h' and h_3/h'' bonds (Fig. 8(c)) which implies that the monomers are stacked in staggered rows. Similar stacking have been reported but in this case, the non-alignment of the monomers was due to the twist of the three amide moieties of about 40 and the H-bond pattern was thus different.[68] The interactions between each pair of monomers and the stability of this interaction is provided by the formation of one expected H-bond (h_1) and an hybrid scheme composed of two H-bonds (h_2/h' or h_3/h''). This finding confirms the importance of the slight conformational flexibility of the amide moieties for the global stability of SMA. It also highlights the fact that the stability of the SMA is dynamic as there is a constant compensation of the H-bonds.

3.4 Aggregation mode

Once the strength of the different supramolecular assemblies (dimers to pentamers) have been considered from both a qualitative and quantitative point of view, we decided to have a look at the formation of those assemblies. In the previous sections, already formed supramolecular assemblies have been considered as starting point of the dynamics, in this section, the aggregation and association process will be studied, aiming at answering the question: how do we go from monomers to larger assemblies? Different simulations with similar starting configurations were set up as follows to address different problems:

– Non-interacting monomers (no H-bond between them) placed in the center of the simulation box for a long (100 ns) simulation time.

– Non-interacting monomers (no H-bond between them) placed in the center of the simulation box for 100 short (1 ns) simulation times.

The first type of simulation will allow us to know if SMA can self-assemble spontaneously and if this SMA will be stable along a long simulation time. The second type of simulation will provide information on the frequency of formation of such SMA and more precisely the frequency of formation of H-bonds. For simplicity, results and detailed analysis on **B4p** are presented here and in ESI† for the others.

3.4.1 One long simulation. We used PACKMOL[69] to generate a starting configuration encompassing 10 monomers that are loosely compacted, meaning they are relatively close to each other but with no hydrogen bonds or particular interaction between them. They were then placed in the center of a simulation box for a 100 ns long simulation time. To ensure that the monomers are not interacting at the beginning of the simulation and more particularly that no hydrogen bond is effective for the starting configuration, we represented the radial distribution function (RDF) for the oxygen and hydrogen atoms involved in those interactions (see Fig. 9(a)). The atoms O_1, O_2 and O_3 and H_1, H_2 and H_3 are the ones associated to h_1, h_2 and h_3 hydrogen bond identified on Fig. 5. One can observe on Fig. 9(a) that for each oxygen atom (O_1, O_2 and O_3) there is no peak indicating a possible hydrogen bond with either H_1, H_2 or H_3 for the starting configuration. The first and most intense peak corresponds to an intramolecular peak at around 3.0 nm corresponding to oxygen and hydrogen atoms of the same amide moiety. The following peaks, appearing at distances larger than 3.0 nm, are indicating the relative spacing of the different O and H atoms and thus of the different monomers. We have then represented the same RDF but for a 100 ns long simulation, differentiating the three possible H-bond that can be formed between two monomers (Fig. 9(b)). One can observe that the O_1–H_1 is predominantly formed when compared to possible O_1–H_2 and O_1–H_3 H-bonds. Looking at the two other graphics, one can say that O_2 is almost not implied in H-bond while some O_3–H_2 and O_3–H_3 are effective. If O_3–H_3 is an expected H-bond in the case of a perfectly aligned dimer, O_3–H_2 H-bond corresponds to the hybrid H-bond already mentioned in the previous section, proving once again the importance of such interaction.

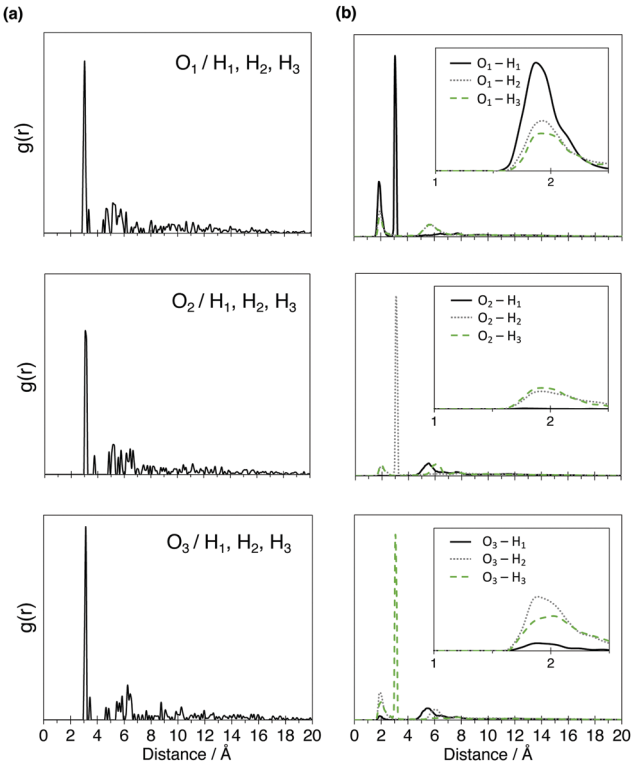

Fig. 9 Radial distribution function for a **B4p** aggregate containing 10 monomers. (a) For the starting configuration (t = 0 s) and (b) for a 100 ns long simulation. The attention has been focused on the oxygen and hydrogen atoms involved in possible H-bond (see Fig. 5).

On Fig. 10 we have provided a representative structure of a SMA of **B4p**, involving 8 monomers, that has been formed during the simulation. One has to notice that the two remaining monomers are involved in a dimer that is not interacting

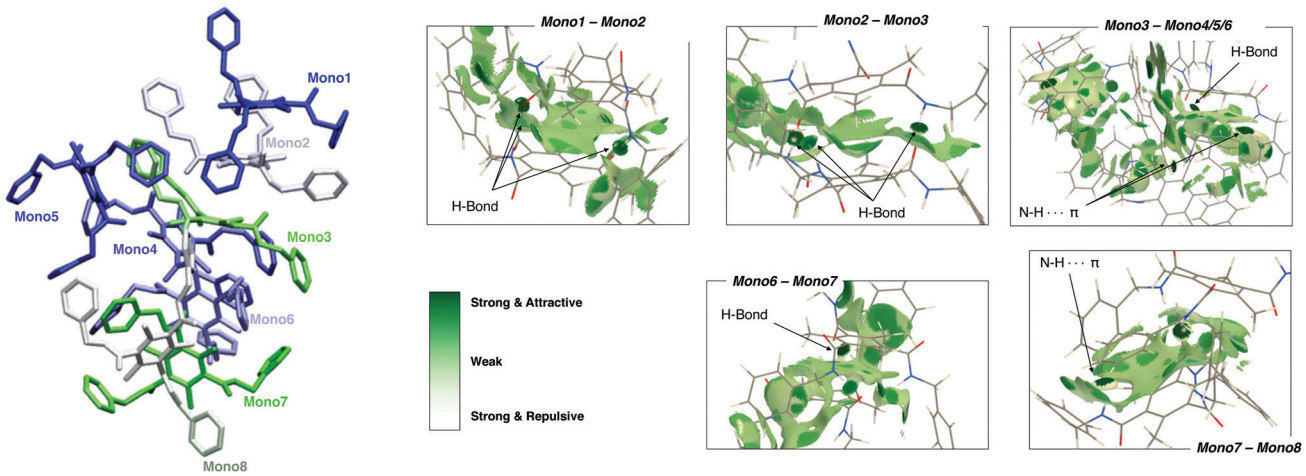

Fig. 10 Representative structure of **B4p** aggregate extracted from the 100 ns long simulation along with the NCI analysis. Isosurfaces (s = 0.5 a.u.) in dark green represent strong and attractive interactions (*e.g.* H-bonds), whereas weak interactions such as van der Waals interactions are depicted in light green. For the sake of clarity for the structure on the left, hydrogen atoms have not been represented, except those of the amide moieties.

with the octamer. To illustrate the interactions between the different monomers we also provided interaction surfaces extracted from the NCI analysis for each pair of monomers. Various interactions have been highlighted among which the predominant ones are:

– h_1 H-bond (Mono2–Mono3, Mono6–Mono7)

– h_2 (h_3) associated with h' (h'') hybrid H-bond (Mono3–Mono4/5/6)

– N–H$\cdots\pi$ bond (Mono3–Mono4/5/6, Mono7–Mono8)

– H-bond involving O_1 atoms and H_2/H_3 atoms (Mono1–Mono2, Mono7–Mono8)

One has to notice that other interactions can be found within the SMA, for example π–π interactions between the phenyls of different monomers and C–H$\cdots\pi$ interactions. Those interactions are nevertheless more labile.

If we go deeper in the analysis of the different interactions, it is possible to observe that all the interactions are not formed simultaneously but sequentially. Indeed the first interaction that is effective within the SMA is the h_1 bond, with a formation at the nanosecond timescale (see Fig. 11). We then have the formation of the previously mentioned interactions (h_2 and h') at a slightly higher timescale (few nanoseconds), followed by N–H$\cdots\pi$ interactions and O_1–H_2/H_3 after tens of nanosecond. If h_1 is the first interaction being observed, it is also the most stable one as it is effective during almost the entire trajectory (Fig. 11(b)). For other molecules of the **B4s** family, there is always the formation of a SMA involving at least five monomers (**B4** and **B4pI**). **B4p**, with 8 monomers involved is the most efficient one while **B4c** forms a seven-members SMA (see Fig. S10–S12 in ESI†). The same kind of interactions are present within all the SMA with a predominance for the h_1 bond, completed with the previously mentioned interactions.

On the opposite, when considering **B9** and **B9m** molecules, no large SMA were detected and only poorly stable dimers were observed. As a conclusion one can say that the self-assembling process is (1) efficient for **B4s** family but not for the **B9s** one, (2) quite fast, of the order of the nanosecond, (3) can be a long process as the interactions are added sequentially, and (4) dynamic in the sense that some interactions (N–H$\cdots\pi$, O_1–H_2, O_1–H_3) are labile but also compensated by other interactions (C–H$\cdots\pi$, π–π), ensuring the global stability of the SMA.

3.4.2 Many short simulations. We observed in the previous section that, when a H-bond is formed, it is then quite stable along the trajectory. The analysis we propose to perform in this section aims at retrieving the frequency of formation of the H-bond. To do so, we generated 100 different starting configuration encompassing 10 monomers close enough to each other but with no H-bond between them and performed a 1 ns simulation. We then extracted the final structure of each simulation and counted the H-bonds between all the monomers. One has to notice that no distinction was made between all the possible H-bonds (9 in total). On Fig. 12 we have represented, for the 100 simulations and for **B4p** and **B9m** only, the total number of H-bonds that have been observed between each pair of dimers. The results for **B4**, **B4c**, **B4pI** and **B9** are also provided in ESI† (Fig. S13 and S14).

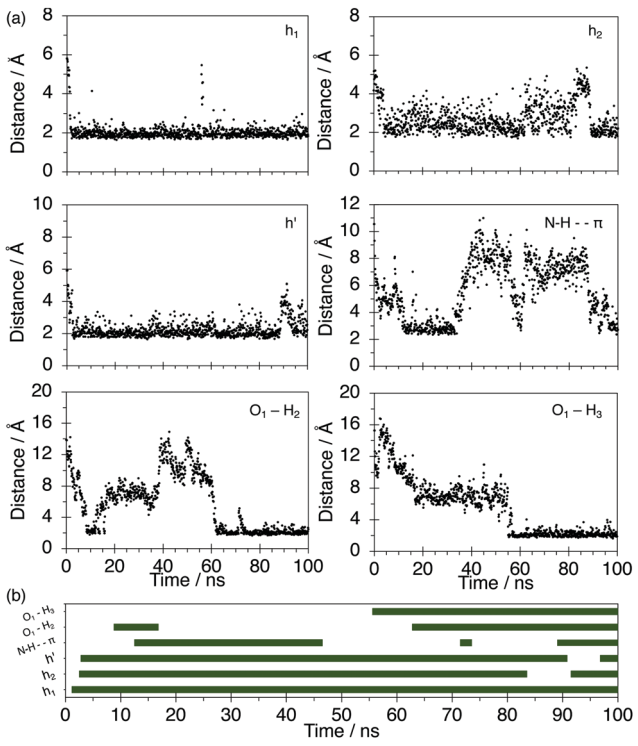

Fig. 11 (a) Evolution for **B4p** octamer of the distances h_1, h_2, h', N–H$\cdots\pi$, O_1–H_2, O_1–H_3 during a 100 ns long simulation. All the distances are in Å. (b) Schematic representation of the formation time and effectiveness of the previously mentioned interactions.

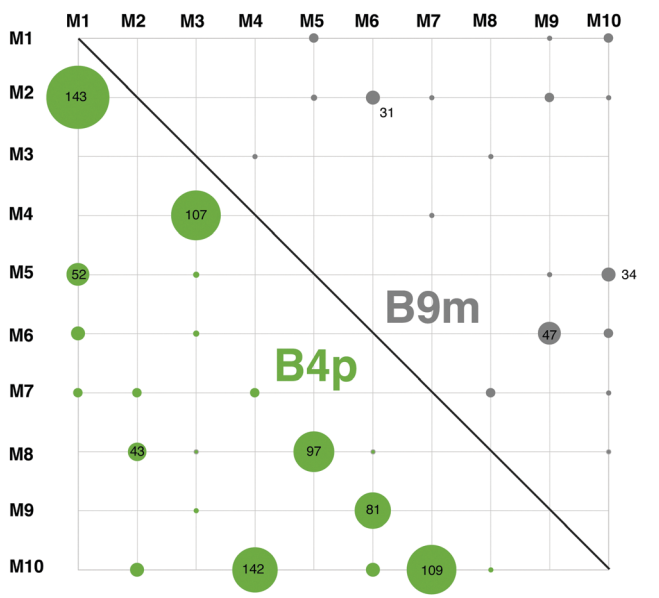

Fig. 12 Schematic representation of the occurrence of the formation of H-bonds between all the monomers (M*i* with *i* = 1,. . .,10) for **B4p** (bottom left, green) and **B9m** (up right, gray) for a total of 100 simulations. The structures that are considered for the count are the final ones obtained 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.

One can observe immediately on Fig. 12 that the number of H-bonds present within **B4p** aggregate is larger than the ones of **B9m**. Due to the particular arrangement of the molecules, some preferential interactions are observed. For example, for **B4p**, the formation of H-bond between monomer 1 and monomer 2 is almost systematic. Indeed, the count reveals 143 H-bonds between those 2 monomers for a maximum of 300. Nevertheless, one can also observe that monomer 1 is also involved in interactions with monomer 5 (52) and in a lesser extent with monomer 6 (31) and 7 (18). So does monomer 2 with monomer 7 (17), 8 (43) and 10 (29). For **B4p**, there are six main interactions involving M1–M2, M3–M4, M4–M10, M7–M10, M5–M8, and M6–M9. For information, all the monomers are staying close to each other, as a loose aggregate, during all the simulations. This can be explained by the fact that other interactions (C–H$\cdots\pi$ and N–H$\cdots\pi$) are formed and thus enforce the stability of the entire supramolecular assembly. For **B9m**, no systematic interaction was observed with a maximum count observed for a M6–M9 interaction. This result may also be due to the fact that **B9m** monomers are not staying ''packed'' during the simulation. The loose aggregate is thus not even stable for **B9m** molecules.

4. Conclusion

A complete computational protocol mixing molecular dynamics simulations and calculations based on quantum mechanics has been defined. It was aimed to study the ability of organic molecules to form supramolecular assemblies and their resulting stabilities. It has thus been possible to provide some hints for an effective supramolecular self-assembly process. A total of 6 molecules (**B4**, **B4c**, **B4p**, **B4pI**, **B9**, and **B9m**) divided into 2 families (**B4s** and **B9s**), based on different backbones but similarly bearing three amide groups as a common feature have been considered. The study of the monomers allowed us to validate our molecular dynamics approach and also to understand the properties of the molecules when they are isolated. These calculations proved that bulky methyl and iodine groups can indeed impose orthogonality to the amides to make them more similar to **B9s** family. When studying the dimers, new hydrogen bond networks were discovered for these assemblies. Expected hydrogen bonds (h_1, h_2, and h_3) were always observed for both families. Nevertheless, they were not always effective altogether and a novel H-bond pattern (h' and h'') has been highlighted. This bifurcated H-bond network, that we called hybrid and which appears as dynamic and adaptive, was observed only for one of the two families, namely **B4s**. The calculation of the binding energies clearly showed that dimers and even trimers or pentamers of the **B4s** family were more favorable than those of the **B9s**. Nevertheless, the evolution of the different binding energies going from dimers to pentamers for **B4s** indicates that there is no obvious cooperative effect within the SMA. There is rather a favorable interaction energy regardless of the size of the SMA that is considered. It may indicate that all the small oligomers act as building blocks for the formation of larger assemblies. For **B9s**, interaction energies are less favorable than the ones of **B4s**. Instead of a cooperative effect, there is rather a non-cooperative effect for **B9m**, with a systematic decrease of the binding energies when the SMA is getting larger. Looking at the formation of the assemblies starting from a loose aggregate has allowed to observe that not only H-bonds can ensure the stability of the aggregate but also N–H$\cdots\pi$ or C–H$\cdots\pi$ in a lesser extent. Finally, if SMAs of **B4s** family are more stable than those of **B9s** it is because they have a relative conformational flexibility regarding the orientation of the amide moieties. It is worth noting that one fundamental characteristic has been highlighted in this study, namely the dynamic behavior of the stability of the SMA. Indeed, we have shown that everytime a SMA was found to be stable, there was a constant compensation of the H-bonds, allowed by conformation flexibility while non stable SMA were characterized by really constrained structures. Finally, by providing a computational protocol allowing a complete description of the amide-based H-bond SMA, it has been possible to drawn up some general rules for the design of efficient self-assembling SMA.

Conflicts of interest

There are no conflicts to declare.

Acknowledgements

The authors thank Dr J. Quertinmont, C. Bouquiaux, P. Beaujean and C. D Matatu MBengo for fruitful scientific discussions. This work was carried out thanks to fundings from the Fonds de la Recherche du Québec (F.R.Q.) et du Fonds de la Recherche Scientifique (F.R.S.-FNRS) in the frame of bilateral research projects under convention R.P001.19. L. L. B. thanks the F.R.S.-FNRS for her post-doctoral grant (convention R.P001.19). The calculations were performed on the computers of the ''Consortium des équipements de Calcul Intensif (CÉCI)'' (http://www.ceci-hpc.be), including those of the ''UNamur Technological Platform of High-Performance Computing (PTCI)'' (http://www.ptci.unamur.be), for which we gratefully acknowledge the financial support from the FNRS-FRFC, the Walloon Region, and the University of Namur (Conventions No. 2.5020.11, GEQ U.G006.15, U.G018.19, 1610468, and RW/GEQ2016) as well as on Zenobe, the Tier-1 facility of the Walloon Region (Convention 1117545).

References

1 P. J. Cragg, *Supramolecular Chemistry and the Life Sciences*, Springer Netherlands, Dordrecht, 2010, pp. 49–89.
2 J. D. Hartgerink, E. Beniash and S. I. Stupp, *Science*, 2001, **294**, 1684–1688.
3 N. Kameta, M. Masuda, G. Mizuno, N. Morii and T. Shimizu, *Small*, 2008, **4**, 561–565.

4 J. T. Ernst, O. Kutzki, A. K. Debnath, S. Jiang, H. Lu and A. D. Hamilton, *Angew. Chem., Int. Ed.*, 2002, **41**, 278–281.
5 D. H. Williams and B. Bardsley, *Angew. Chem., Int. Ed.*, 1999, **38**, 1172–1193.
6 Y. Wen, Y. Song, G. Jiang, D. Zhao, K. Ding, W. Yuan, X. Lin, H. Gao, L. Jiang and D. Zhu, *Adv. Mater.*, 2004, **16**, 2018–2021.
7 J.-M. Lehn, *Supramolecular Chemistry*, John Wiley & Sons, Ltd, 1995.
8 P. J. Cragg, *An Introduction to Supramolecular Chemistry*, Springer Netherlands, Dordrecht, 2010, pp. 1–48.
9 C. W. Bunn, E. V. Garner and W. L. Bragg, *Proc. R. Soc. London, Ser. A*, 1947, **189**, 39–68.
10 A. V. Morozov and T. Kortemme, *Peptide Solvation and HBonds*, Advances in Protein Chemistry, Academic Press, 2005, vol. 72, pp. 1–38.
11 E. T. Powers, S. Deechongkit and J. W. Kelly, *Peptide Solvation and HBonds*, Advances in Protein Chemistry, Academic Press, 2005, vol. 72, pp. 39–78.
12 A. Greenberg, C. Breneman and J. F. Liebman, *The amide linkage: structural significance in chemistry, biochemistry, and materials science*, 2003.
13 Y. Matsunaga, N. Miyajima, Y. Nakayasu, S. Sakai and M. Yonenaga, *Bull. Chem. Soc. Jpn.*, 1988, **61**, 207–210.
14 Y. Shishido, H. Anetai, T. Takeda, N. Hoshino, S.-I. Noro, T. Nakamura and T. Akutagawa, *J. Phys. Chem. C*, 2014, **118**, 21204–21214.
15 L. Brunsveld, A. Schenning, M. Broeren, H. Janssen, J. Vekemans and E. Meijer, *Chem. Lett.*, 2000, 292–293.
16 J. Roosma, T. Mes, P. Leclère, A. R. A. Palmans and E. W. Meijer, *J. Am. Chem. Soc.*, 2008, **130**, 1120–1121.
17 P. J. M. Stals, M. M. J. Smulders, R. Martín-Rapún, A. R. A. Palmans and E. W. Meijer, *Chem. – Eur. J.*, 2009, **15**, 2071–2080.
18 P. J. M. Stals, P. A. Korevaar, M. A. J. Gillissen, T. F. A. de Greef, C. F. C. Fitié, R. P. Sijbesma, A. R. A. Palmans and E. W. Meijer, *Angew. Chem., Int. Ed.*, 2012, **51**, 11297–11301.
19 D. Gauthier, P. Baillargeon, M. Drouin and Y. L. Dory, *Angew. Chem., Int. Ed.*, 2001, **40**, 4635–4638.
20 P. Baillargeon, S. Bernard, D. Gauthier, R. Skouta and Y. Dory, *Chem. – Eur. J.*, 2007, **13**, 9223–9235.
21 S. Leclair, P. Baillargeon, R. Skouta, D. Gauthier, Y. Zhao and Y. L. Dory, *Angew. Chem., Int. Ed.*, 2004, **43**, 349–353.
22 M. F. J. Mabesoone, S. Kardas, H. Soria-Carrera, J. Barberá, J. M. de la Fuente, A. R. A. Palmans, M. Fossépré, M. Surin and R. Martín-Rapún, *Mol. Syst. Des. Eng.*, 2020, **5**, 820–828.
23 A. J. Wilson, J. van Gestel, R. P. Sijbesma and E. W. Meijer, *Chem. Commun.*, 2006, 4404–4406.
24 S. Cantekin, T. F. A. de Greef and A. R. A. Palmans, *Chem. Soc. Rev.*, 2012, **41**, 6125–6137.
25 Y.-J. Choi, W.-J. Yoon, D.-Y. Kim, M. Park, Y. Lee, D. Jung, J.-S. Kim, Y.-T. Yu, C.-R. Lee and K.-U. Jeong, *Polym. Chem.*, 2017, **8**, 1888–1894.
26 S. Devi, I. Bala, S. P. Gupta, P. Kumar, S. K. Pal and S. Venkataramani, *Org. Biomol. Chem.*, 2019, **17**, 1947–1954.
27 M. Gelinsky, R. Vogler and H. Vahrenkamp, *Inorg. Chem.*, 2002, **41**, 2560–2564.
28 P. Besenius, J. L. M. Heynens, R. Straathof, M. M. L. Nieuwenhuizen, P. H. H. Bomans, E. Terreno, S. Aime, G. J. Strijkers, K. Nicolay and E. W. Meijer, *Contrast Media Mol. Imaging*, 2012, **7**, 356–361.
29 T. Marmin and Y. L. Dory, *Chem. – Eur. J.*, 2019, **25**, 6707–6711.
30 N. M. Matsumoto, R. P. M. Lafleur, X. Lou, K.-C. Shih, S. P. W. Wijnands, C. Guibert, J. W. A. M. van Rosendaal, I. K. Voets, A. R. A. Palmans, Y. Lin and E. W. Meijer, *J. Am. Chem. Soc.*, 2018, **140**, 13308–13316.
31 C. Oliveras-González, M. Linares, D. B. Amabilino and N. Avarvari, *ACS Omega*, 2019, **4**, 10108–10120.
32 Y. Li, A. Hammoud, L. Bouteiller and M. Raynal, *J. Am. Chem. Soc.*, 2020, **142**, 5676–5688.
33 B. Adelizzi, N. J. Van Zee, L. N. J. de Windt, A. R. A. Palmans and E. W. Meijer, *J. Am. Chem. Soc.*, 2019, **141**, 6110–6121.
34 M. P. Lightfoot, F. S. Mair, R. G. Pritchard and J. E. Warren, *Chem. Commun.*, 1999, 1945–1946.
35 M. Calvelo, C. I. Lynch, J. R. Granja, M. S. P. Sansom and R. Garcia-Fandiño, *ACS Nano*, 2021, **15**, 7053–7064.
36 I. A. W. Filot, A. R. A. Palmans, P. A. J. Hilbers, R. A. van Santen, E. A. Pidko and T. F. A. de Greef, *J. Phys. Chem. B*, 2010, **114**, 13667–13674.
37 K. K. Bejagam, G. Fiorin, M. L. Klein and S. Balasubramanian, *J. Phys. Chem. B*, 2014, **118**, 5218–5228.
38 C. Huang, R. K. Wang, B. M. Wong, D. J. McGee, F. Léonard, Y. J. Kim, K. F. Johnson, M. S. Arnold, M. A. Eriksson and P. Gopalan, *ACS Nano*, 2011, **5**, 7767–7774.
39 Y. Joo, G. J. Brady, M. J. Shea, M. B. Oviedo, C. Kanimozhi, S. K. Schmitt, B. M. Wong, M. S. Arnold and P. Gopalan, *ACS Nano*, 2015, **9**, 10203–10213.
40 L. Le Bras, R. Berthin, I. Hamdi, M. Louati, S. Aloïse, M. Takeshita, C. Adamo and A. Perrier, *Phys. Chem. Chem. Phys.*, 2020, **22**, 6942–6952.
41 J. A. Lemkul and D. R. Bevan, *J. Phys. Chem. B*, 2010, **114**, 1652–1660.
42 F. Mahmoudinobar, J. M. Urban, Z. Su, B. L. Nilsson and C. L. Dias, *J. Chem. Theory Comput.*, 2019, **15**, 3868–3874.
43 I. M. Ilie and A. Caflisch, *Chem. Rev.*, 2019, **119**, 6956–6993.
44 N. T. Tung, P. Derreumaux, V. V. Vu, P. C. Nam and S. T. Ngo, *ACS Omega*, 2019, **4**, 11066–11073.
45 E. R. Johnson, S. Keinan, P. Mori-Sánchez, J. Contreras-García, A. J. Cohen and W. Yang, *J. Am. Chem. Soc.*, 2010, **132**, 6498–6506.
46 A. Perrier, M. Eluard, M. Petitjean and A. Vanet, *J. Phys. Chem. B*, 2019, **123**, 582–592.
47 M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, G. A. Petersson, H. Nakatsuji, X. Li, M. Caricato, A. V. Marenich, J. Bloino, B. G. Janesko, R. Gomperts, B. Mennucci, H. P. Hratchian, J. V. Ortiz, A. F. Izmaylov, J. L. Sonnenberg, D. Williams-Young, F. Ding, F. Lipparini, F. Egidi, J. Goings, B. Peng, A. Petrone, T. Henderson, D. Ranasinghe, V. G. Zakrzewski, J. Gao, N. Rega, G. Zheng, W. Liang, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao,

H. Nakai, T. Vreven, K. Throssell, J. A. Montgomery Jr, J. E. Peralta, F. Ogliaro, M. J. Bearpark, J. J. Heyd, E. N. Brothers, K. N. Kudin, V. N. Staroverov, T. A. Keith, R. Kobayashi, J. Normand, K. Raghavachari, A. P. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, J. M. Millam, M. Klene, C. Adamo, R. Cammi, J. W. Ochterski, R. L. Martin, K. Morokuma, O. Farkas, J. B. Foresman and D. J. Fox, *Gaussian 16 Revision C.01*, Gaussian Inc., Wallingford CT, 2016.

48 J.-D. Chai and M. Head-Gordon, *Phys. Chem. Chem. Phys.*, 2008, **10**, 6615–6620.

49 K. S. Thanthiriwatte, E. G. Hohenstein, L. A. Burns and C. D. Sherrill, *J. Chem. Theory Comput.*, 2011, **7**, 88–96.

50 J. Tomasi, B. Mennucci and R. Cammi, *Chem. Rev.*, 2005, **105**, 2999–3094.

51 H. P. Hratchian, A. V. Krukau, P. V. Parandekar, M. J. Frisch and K. Raghavachari, *J. Chem. Phys.*, 2011, **135**, 014105.

52 J. Contreras-García, E. R. Johnson, S. Keinan, R. Chaudret, J.-P. Piquemal, D. N. Beratan and W. Yang, *J. Chem. Theory Comput.*, 2011, **7**, 625–632.

53 J. Wang, R. M. Wolf, J. W. Caldwell, P. A. Kollman and D. A. Case, *J. Comput. Chem.*, 2004, **25**, 1157–1174.

54 M. J. Abraham, T. Murtola, R. Schulz, S. Páll, J. C. Smith, B. Hess and E. Lindahl, *SoftwareX*, 2015, **1-2**, 19–25.

55 J. Wang, R. M. Wolf, J. W. Caldwell, P. A. Kollman and D. A. Case, *J. Comput. Chem.*, 2004, **25**, 1157–1174.

56 J. Wang, W. Wang, P. A. Kollman and D. A. Case, *J. Mol. Graphics*, 2006, **25**, 247–260.

57 P. R. Batista, A. Wilter, E. H. A. B. Durham and P. G. Pascutti, *Cell Biochem. Biophys.*, 2006, **44**, 395–404.

58 T. Darden, D. York and L. Pedersen, *J. Chem. Phys.*, 1993, **98**, 10089–10092.

59 U. Essmann, L. Perera, M. L. Berkowitz, T. Darden, H. Lee and L. G. Pedersen, *J. Chem. Phys.*, 1995, **103**, 8577–8593.

60 H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren, A. DiNola and J. R. Haak, *J. Chem. Phys.*, 1984, **81**, 3684–3690.

61 S. Nosé and M. Klein, *Mol. Phys.*, 1983, **50**, 1055–1076.

62 S. Nosé, *J. Chem. Phys.*, 1984, **81**, 511–519.

63 W. G. Hoover, *Phys. Rev. A: At., Mol., Opt. Phys.*, 1985, **31**, 1695–1697.

64 M. Parrinello and A. Rahman, *J. Appl. Phys.*, 1981, **52**, 7182–7190.

65 J. S. Hub, B. L. de Groot and D. van der Spoel, *J. Chem. Theory Comput.*, 2010, **6**, 3713–3720.

66 C. H. Görbitz, *Curr. Opin. Solid State Mater. Sci.*, 2002, **6**, 109–116.

67 A. Sarkar, T. Behera, R. Sasmal, R. Capelli, C. Empereurmot, J. Mahato, S. S. Agasti, G. M. Pavan, A. Chowdhury and S. J. George, *J. Am. Chem. Soc.*, 2020, **142**, 11528–11539.

68 J. van Gestel, A. R. A. Palmans, B. Titulaer, J. A. J. M. Vekemans and E. W. Meijer, *J. Am. Chem. Soc.*, 2005, **127**, 5490–5494.

69 L. Martínez, R. Andrade, E. G. Birgin and J. M. Martínez, *J. Comput. Chem.*, 2009, **30**, 2157–2164.