


 Cite this: *Mol. Syst. Des. Eng.*, 2021, **6**, 805

 Received 6th June 2021,  
 Accepted 13th July 2021

DOI: 10.1039/d1me00068c

[rsc.li/molecular-engineering](http://rsc.li/molecular-engineering)

## Variation of interaction zone size for the target design of 2D supramolecular networks†

 Łukasz Baran, \* Wojciech Rżysko  and Dariusz Tarasewicz

In this study, we have performed extensive coarse-grained molecular dynamics simulations of the self-assembly of tetra-substituted molecules. We have found that such molecules are able to form a variety of structures, depending on the parameters of the employed model. In particular, it has been demonstrated that even slight changes of the interaction zone and the shape of molecules can drastically alter the behavior of investigated systems. We have established the rules governing the formation of the Sierpinski triangles, Archimedean tessellation, Kagomé, and ladder networks. The appearance of Sierpinski triangles is rather surprising, since a majority of papers report the formation of such structures in completely different systems. The only general rule that has been established and proved experimentally is that the so-called “V-shape” molecules are able to order into Sierpinski triangles.

### Design, System, Application

The manuscript presents the results of molecular dynamics simulations of the self-assembly of tetrasubstituted molecules. We have demonstrated that even slight changes of the interaction zone and the shape of molecules can drastically alter the behavior of investigated systems. This parameter can be reflected by substitution of diverse chemically active groups. Hence, our results indicate how one can control experimental conditions of the self-assembly process for the target design of supramolecular networks presented herein. Therefore, we believe that the findings presented in this paper can be helpful for scientists working on such systems, as they can help select a building block that is able to create a 2D network with a predefined topology.

## 1 Introduction

Spontaneous self-assembly processes are ubiquitous in nature. Many biological systems such as phospholipids,<sup>1</sup> RNA<sup>2</sup> and DNA<sup>3</sup> complexes, viruses,<sup>4</sup> *etc.* undergo self-assembly. Furthermore, Janus (or generally “patchy”) particles with anisotropic interactions also show self-assembly of great importance for materials science,<sup>5,6</sup> medicine,<sup>7,8</sup> and other applications. Several organic molecules have been given attention due to their ability to undergo on-surface reactions, which often lead to the formation of products different than those appearing in bulk systems, or they would even not be possible to take place in the bulk. The main objective of this approach is to obtain single nanolayers, which have been proven to be used as membranes for the separation of liquid and gas phases,<sup>9</sup> in batteries,<sup>10</sup> and in molecular sieves.<sup>11</sup>

Another interesting behavior of some organic molecules is their assembly into mathematical fractals. There exist several

approaches attempting to fabricate such supramolecular networks. While it is impressive to see that the obtained structures are, for instance, high order Sierpinski triangles, they are often due to the expanded molecular structure of the building blocks.<sup>12,13</sup> On the other hand, one of the most popular simple precursors being able to form true mathematical fractals, being Sierpinski triangles, are the so-called “V-shaped” molecules. One of the most popular theoretical approaches that have shown the formation of this kind of self-similar network was the lattice Monte Carlo method (l-MC).<sup>14</sup> The predictions stemming from simulations were also proved experimentally.<sup>15–18</sup> The same authors proposed another molecular architecture that could also possibly form Sierpinski fractals using an identical approach.<sup>19</sup> However, it is well-known that lattice models should be treated as a preliminary method for the investigation of any problem of interest since the lattice type can enforce the formation of structures that correspond to its symmetry and are not necessarily reproducible by other off-lattice methods (or more importantly by experiments).

Due to the multiple possible applications and the variety of different paths of molecular assembly, it is indispensable to find general rules that could lead to the easier design of these kinds of systems. While the synthesis can be tedious

Department of Theoretical Chemistry, Maria Curie Skłodowska University, Lublin, Poland. E-mail: lukasz.baran@poczta.umcs.lublin.pl

† Electronic supplementary information (ESI) available. See DOI: 10.1039/d1me00068c



and costly, computer modelling can be a very convenient substitute for the exploration of problems of interest. Computer simulations can give valuable insight for experimentalists due to the possibility of examination of the influence of multiple factors, such as particle geometry, under different conditions on the self-assembly process in a reasonable time. The most popular methods involve quantum density functional theory,<sup>20</sup> classical Monte Carlo<sup>21–26</sup> and molecular dynamics simulations (MD).<sup>27,28</sup> In our laboratory, we have proven that the last method can be useful for examination of a variety of molecular geometries, both in one-component systems<sup>29,30</sup> and in binary mixtures.<sup>31</sup>

In this paper, we report various paths of self-assembly of tetra-topic building blocks. We have compared the prediction stemming from I-MC<sup>19</sup> with molecular dynamics and extended the model to different molecular geometries. It was shown in the literature that the precise control over selectivity of on-surface reactions is crucial for the design of desired nanostructures.<sup>32,33</sup> Here we demonstrate that even slight variations of the size of the interaction zone drastically change the behavior of the systems considered.<sup>34,35</sup> We have found the rules governing the formation of diverse networks including Sierpinski triangles, Archimedean tessellation, Kagomé, and ladder networks. The last case is of particular importance since it resembles the behavior of liquid crystals which tend to glue due to entropic effects. All of these structures have been characterized by appropriate order parameters.

## 2 Results and discussion

### 2.1 Model

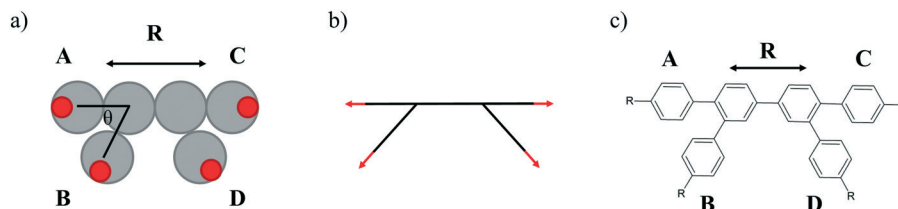
The model proposed in this paper has been designed to reflect the geometry of flat and rigid tetra-substituted chemical compounds as shown in Fig. 1. In the structure of molecules, we can distinguish two parts, the core and four arms attached to it, marked as **R** and **A–D**, respectively. We will refer to those as an entire backbone. The angle between neighboring arms has been set to  $\theta = 60^\circ$ . The influence of the core length, as well as the lengths of arms, has been investigated. Moreover, the terminal segment of each arm has been decorated with the so-called “sticky patch” (red segments in Fig. 1)), which will be referred to as active sites. The size of

every segment forming the entire backbone has been set to  $\sigma_b$ , whereas active sites were five times smaller  $\sigma_a = 0.2\sigma_b$ . The bonding distance between the neighboring backbone entities has also been set to  $\sigma_b$ , whereas the bonding distance between terminal arms and active sites has been varied between  $l = 0.4 \pm 0.04\sigma_b$ . Thus, the active sites are, to some extent, embedded into terminal segments to reflect directional interparticle interactions. A detailed description of molecular dynamics simulations can be found in the ESI.†

Fig. 2(a) gives a schematic representation of the interaction zone arising from the above model. Parts b–d of Fig. 2 demonstrate the influence of the bonding distance  $l$  on the angular behavior of the employed Lennard-Jones potential. We have estimated that the maximum size of the interaction zone changes with  $l$  and is equal to  $\alpha = 34.6^\circ$  for  $l = 0.36$ ,  $\alpha = 41.4^\circ$  for  $l = 0.40\sigma$ , and  $\alpha = 45.2^\circ$  for  $l = 0.44\sigma$ .

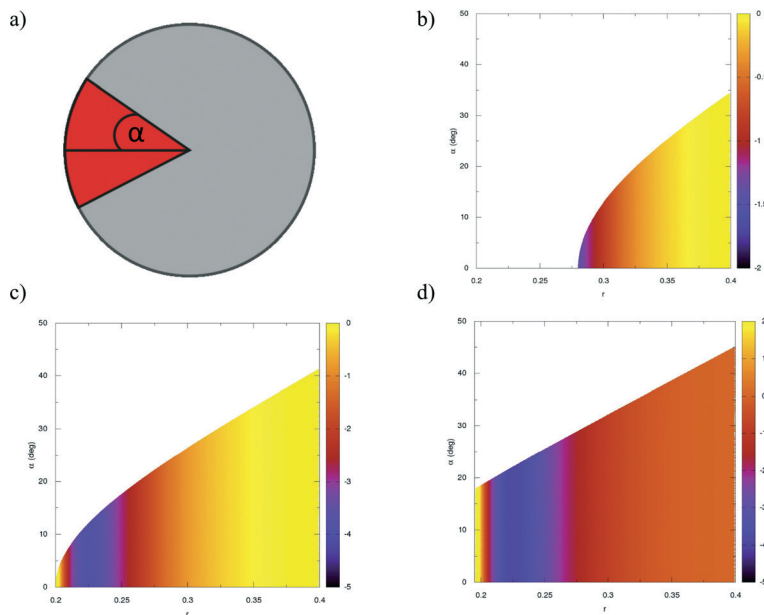
Kalyuzhnyi and Cummings<sup>35</sup> have shown that, due to geometrical constraints, the ranges of the interaction zone  $\alpha$  allowing for the connections between two ( $30^\circ \leq \alpha < 35.3^\circ$ ), three ( $35.3^\circ \leq \alpha < 45^\circ$ ), etc. molecules are well defined. However, these authors used a relatively short-ranged square-well potential and the above ranges of  $\alpha$  are expected to vary depending on the form of interparticle potential.

In the considered case of the Lennard-Jones potential, the ranges of  $\alpha$  are approximately the same, though the cut-off distance between active sites is different. We have found that for distances  $l = 0.36\sigma$  and  $l = 0.44\sigma$ , connections between two and three active sites are possible, respectively. In the case of  $l = 0.40\sigma$ , the association may involve the formation of bonds between two and three active centers. In our opinion, such a kind of coarse-grained modelling can be mapped on real experimental studies. One should substitute a chemical compound of similar geometry as shown in Fig. 1(c) with groups such as  $-\text{OH}$  or  $-\text{COOH}$  in single-component systems or alternatively halide groups in a mixture with two-coordinated metal atoms which in both cases allow interactions of two molecules with one another. This case would correspond to the results of  $l = 0.36\sigma$  presented in this paper. On the other hand, if one would substitute the same chemical compound with  $-\text{B}(\text{OH})_2$  in single-component systems or halide groups in a mixture with three-coordinated metal atoms, one should expect to obtain similar results to those for  $l = 0.44\sigma$ . The intermediate scenario where two phases may



**Fig. 1** Model of a tetra-substituted molecule (part a) with **R** and **A–D** being the lengths of the core and the arms, respectively. Silver circles correspond to the segments of the entire backbone, whereas red circles mark the active sites. Part b shows a schematic representation of the molecule from (part a). The example of a chemical compound that can correspond to our model (part c). Notation of **R** and **A–D** is the same as in part (a), and  $-\text{R}$  are active sites resulting from the presence of attached functional groups, such as  $-\text{CN}$ ,  $\text{B}(\text{OH})_2$ , etc.





**Fig. 2** Schematic representation of the interaction zone (a). Parts (b–d) present the angular dependence of the Lennard-Jones potential with respect to the separation distance between active sites for  $l = 0.36\sigma$  (b),  $l = 0.40\sigma$  (c), and  $l = 0.44\sigma$  (d).

occur, which is most likely the case in experimental studies due to the oscillations of chemical bonds of active groups, can be seen for  $l = 0.40\sigma$ . In our paper, we have shown the consequences of active zone manipulation on the formation of a variety of self-assembled networks.

All considered systems, summarized in Table 1, can be grouped into five models, abbreviated as **M1–M5**, with different numbers of segments in the arms **A–D**. For each of these models, we have also studied the influence of the core length, **X**, on self-assembly processes.

## 2.2 The influence of bonding distance on self-assembly

**2.2.1 Results for  $l = 0.40\sigma$ .** Let us start with the presentation of results for model **M1**. A typical snapshot for system **R2\_M1** is given in Fig. 3(a). In this case, we can see several patterns, schematically drawn in the right-hand side of this panel, showing how the molecules assemble. One can see that one network prevails, which is an isogonal tiling, in which molecules periodically form smaller and bigger triangles. The result for model **R3\_M1**, shown in Fig. 3(b) shows that elonga-

tion of the core length **X** leads to the stabilization of this structure and no other patterns can be distinguished. It is also noteworthy that a very similar network was recently obtained by Y. Zhang, *et al.* in the mixture of 1,3-bis(4-pyridyl) benzene molecules and Fe atoms.<sup>18</sup>

To better quantify this structure, we have examined the way how the active sites assemble, and the corresponding snapshot can be found in Fig. 3(c). It is clear that connecting the centers of mass of two associated active sites, the 3.4.6.4 Archimedean tessellation is formed. The results for model **R4\_M1** are qualitatively the same as for model **R3\_M1** and can be found in Fig. S1 in the ESI.† The only difference is the bigger pore size in model **R3\_M1**.

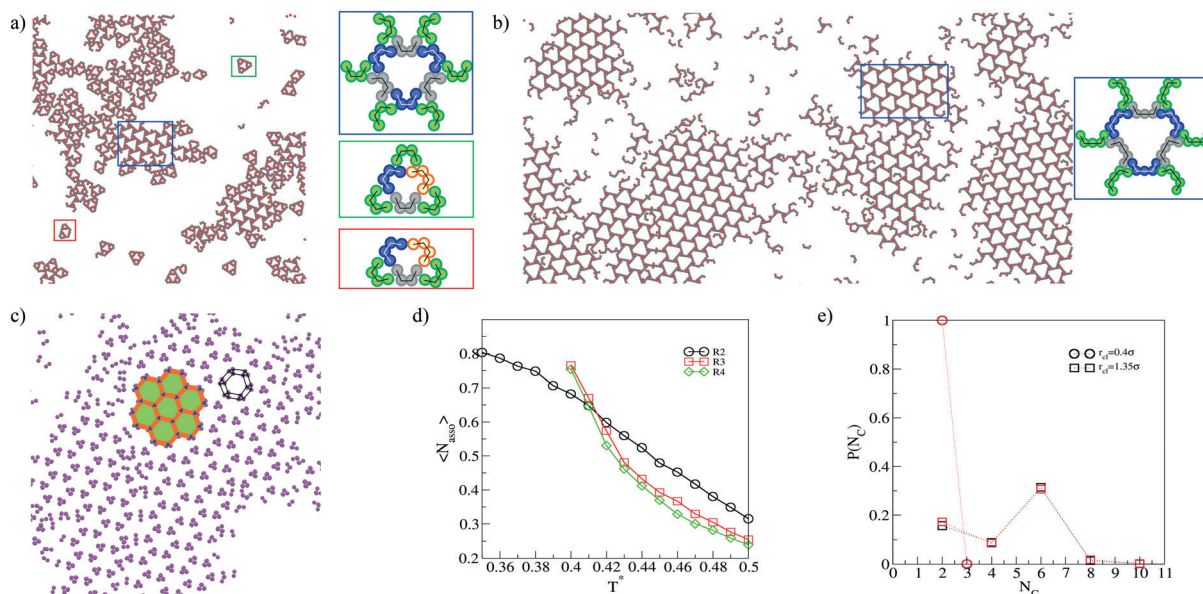
All of the observations concerning those networks can be proved by the calculation of several parameters. The first one is the average association number,  $\langle N_{\text{asso}} \rangle$ , which measures the average number of neighboring molecules within the interaction zone ( $r = 0.4\sigma$ ) of the active sites of a selected molecule. If it takes a value of one it means that two active sites of different molecules associate, *etc.* The changes of this parameter with temperature can be found in Fig. 3(d) and the value at the lowest temperatures is around  $\langle N_{\text{asso}} \rangle \approx 0.8$  for all the systems. This result corroborates with the snapshots showing that only two molecules associate. Since this is an average value and there also are non-associated particles present, the value is lowered in comparison with that predicted for the perfect ordering.

Another quantity that we have computed is the cluster distribution function defined so that  $\int_0^N P(N_C) dN_C = 1$ , where  $N$  is the number of active sites. It has been calculated for active sites up to the threshold distance of  $r_{\text{cl}} = 0.4\sigma$  and  $r_{\text{cl}} = 1.35\sigma$ , which are the cut-off distance of the interparticle potential, and the location of the second minimum of the radial

**Table 1** The systems investigated in the course of our simulations. M marks the models of different molecular architectures, whereas **X** and **A–D** are the numbers of segments in the core and in each of the arms, respectively. In Fig. 1a, model **R2\_M2** ( $X = 2$ , **A–D** = 1) has been shown

Model abbreviation	Full description
<b>RX_M1</b>	<b>X</b> = 2–4, <b>A</b> , <b>C</b> = 0, <b>B</b> , <b>D</b> = 1
<b>RX_M2</b>	<b>X</b> = 2–4, <b>A–D</b> = 1
<b>RX_M3</b>	<b>X</b> = 2–4, <b>A</b> , <b>C</b> = 1, <b>B</b> , <b>D</b> = 2
<b>RX_M4</b>	<b>X</b> = 2–4, <b>A</b> , <b>C</b> = 2, <b>B</b> , <b>D</b> = 1
<b>RX_M5</b>	<b>X</b> = 2–4, <b>A–D</b> = 2





**Fig. 3** Fragment of the configurations for models **R2\_M1** in  $\rho^* = 0.2$ , at  $T^* = 0.35$  (a) and **R3\_M1** in  $\rho^* = 0.2$ , at  $T^* = 0.4$  (b). Schematically drawn association paths are shown on the right-hand sides of these panels. Part c) shows how active sites assemble into Archimedean tessellation. Different polygons are marked by different colors to better show the ordering. The average association number for models **R2–R4\_M1** (d). Cluster distribution functions for models **R3\_M1** (black symbols) and **R4\_M1** (red symbols) for two different threshold distances (e).

distribution function, respectively. The results obtained for models **R3\_M1** and **R4\_M1** are presented in Fig. 3(e). In model **R3\_M1**, the cluster distribution function confirms “visual” observations that only two molecules associate, while in the case of model **R4\_M1**, the most probable cluster size is equal to  $N_C = 6$ . However, there are also clusters with  $N_C = 2$  and 4, which appear due to imperfections of the network. In the case of  $r_{cl} = 1.35\sigma$ , we have computed the center of mass for every cluster found in the system, and these particles form a honeycomb network (see Fig. S1 in the ESI†). The results for the **R4\_M1** model are the same and can also be found in the ESI†.

Next, we proceed to the description of model **M2**. A typical snapshot for system **R2\_M2** at density  $\rho^* = 0.2$  can be found in Fig. 4(a). It has been found that molecules can associate into two different networks, namely into Sierpinski triangles up to the second-order, and into a denser phase. In the former structure, two molecules associate, just the same as in the previously discussed model, while in the latter, two or three active sites associate. As previously mentioned, we have checked the influence of the core length,  $X$  and found that in the range of  $X = 2–4$  the same tendency occurs (cf. Fig. S2 in the ESI†). The configuration for model **R4\_M2** at density  $\rho^* = 0.2$  is shown in Fig. 4(b). In this case, however, the amount of dense phase is lower than for smaller core lengths. On the other hand, the elongation of the core does not lead to the preferred formation of Sierpinski triangles of a higher degree. We conclude that at density  $\rho^* = 0.2$ , the formation of two competing structures is the reason why the system is not able to self-assemble into well-defined bigger clusters.

In all of the systems, we have also checked the influence of the density on self-assembly and found that at lower densi-

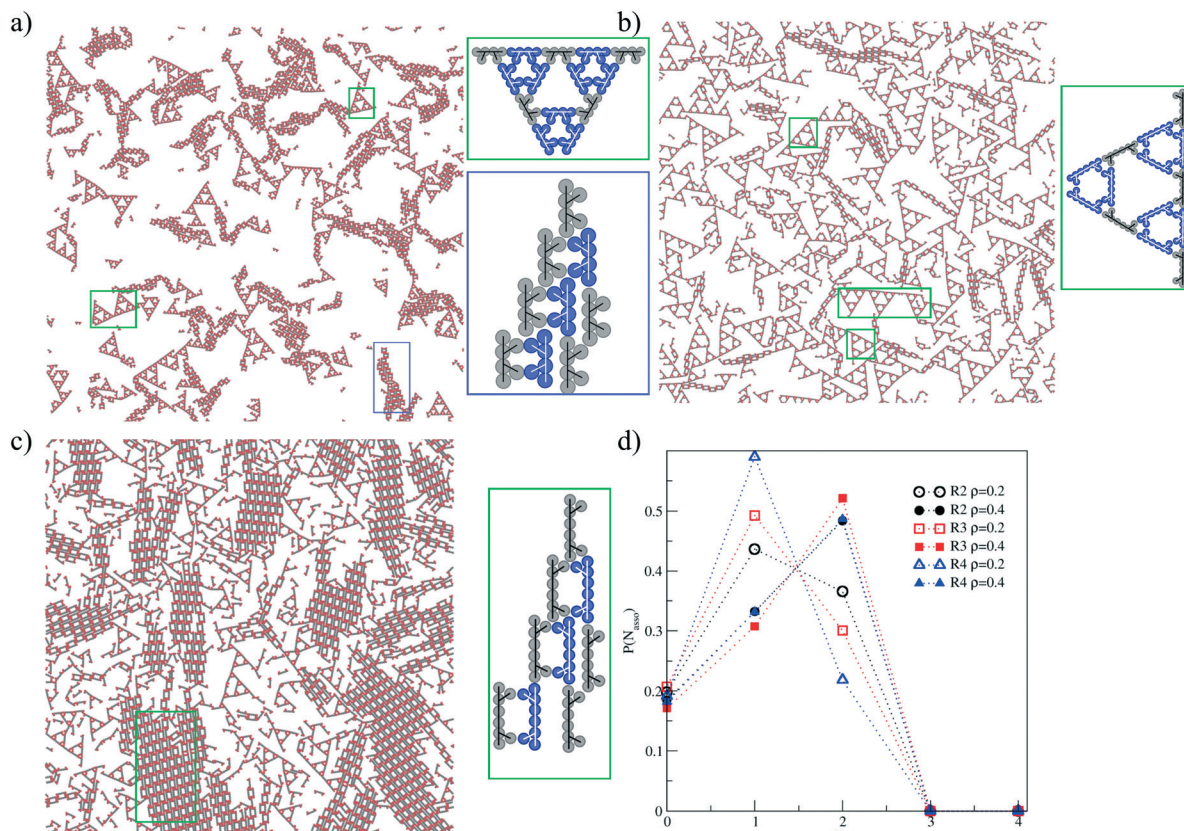
ties around  $\rho^* = 0.12$  there is no qualitative change within the range of  $X = 2–4$  (cf. Fig. S2 in the ESI†). On the other hand, at higher densities, around  $\rho^* = 0.4$ , the formation of only one well-developed ladder-like structure is observed for every **M2** model. A snapshot for the **R4\_M2** model, recorded at density  $\rho^* = 0.4$ , is given in Fig. 4(c), and other examples can be found in Fig. S2 in the ESI†.

To quantify the behavior of the considered systems better, we have computed the distribution of association number for every **M2** model at densities  $\rho^* = 0.2$  and  $\rho^* = 0.4$ . The results, shown in Fig. 4(d), display that at the lower density the most probable is the association involving two molecules ( $N_{\text{asso}} = 1$ ), while at the higher density three active sites associate ( $N_{\text{asso}} = 2$ ). It is also worth mentioning that at density  $\rho^* = 0.2$ , the maximum of the association distribution increases with the core length. This is consistent with the picture emerging from the snapshots showing that in model **R4\_M2** the Sierpinski triangles are the dominating entities. At density  $\rho^* = 0.4$ , we have observed a higher contribution of well-ordered ladder-like structure of higher density, independent of the core length. The schematic representation of the structure, shown on the right-hand side of Fig. 4(c), shows that two and three molecules can associate to form the network, in agreement with the behavior of the distribution function.

Then, we have examined the behavior of model **M3**. Results for particles **R2\_M3** are shown in Fig. 5(a). In this case, the molecules assemble into one and well-defined dense network. Similar to the denser phase observed in model **M2**, two and three active sites can associate and the pattern of the ordering can be found in Fig. 5(b). Clusters of sizes equal to  $N_C = 2$  and  $N_C = 3$  form lines, arranged vertically in the 2–3–3–2 sequence. We have checked the influence of the core length







**Fig. 4** Configurations for models **R2\_M2** in  $\rho^* = 0.2$ , at  $T^* = 0.4$  (a) and **R4\_M2** in  $\rho^* = 0.2$ , at  $T^* = 0.4$  (b). Snapshot for model **R4\_M2** in  $\rho^* = 0.4$ , at  $T^* = 0.45$  (c). Schematically drawn association paths are shown on the right-hand side of these panels. The distribution of the association number for models **R2–R4\_M2** in density  $\rho^* = 0.2$ , at  $T^* = 0.4$  (open symbols) and  $\rho^* = 0.4$ , at  $T^* = 0.45$  (filled symbols) (d).

on the behavior of this model and found qualitatively the same behavior for  $X = 2-4$  (cf. Fig. S3 in the ESI†).

Fig. 5(c) gives the results for model **R2\_M4**. The structure formed by those particles is very similar to the one formed in the previous case (and to the dense phase in model **M2**). Likewise, one can distinguish two pores of different sizes, and two, as well as three active sites, can assemble. The main difference, however, is on how the latter “atoms” are ordered. Namely the pairs and the triplets are now ordered alternately in lines and form a “zigzag-like” pattern (cf. Fig. 5(d)).

Unlike in model **M3**, elongation of the core length  $X$  leads to completely different results. The snapshot for particles **R3\_M4** at  $\rho^* = 0.2$  and  $T^* = 0.4$  is presented in Fig. 5(e). The molecules assemble into an aligned in one direction ladder network. It is rather surprising that these “wires” remain separated and do not glue into a denser structure. From the cluster analysis computed for the segments of the entire backbone, up to the threshold distance taken from the first minimum of the radial distribution function ( $r_{cl} = 1.05\sigma$ ), we have found that the average length of such a wire consists of  $\langle N_C \rangle \approx 433 \pm 10$  segments, with the maximum value equal to  $N_C = 3357$ . It should be noted that the linear system size was equal to  $L_x = L_y = 336\sigma$ . Thus, the wire length is much larger than the system size. Besides that, the formation of small patches of a dense

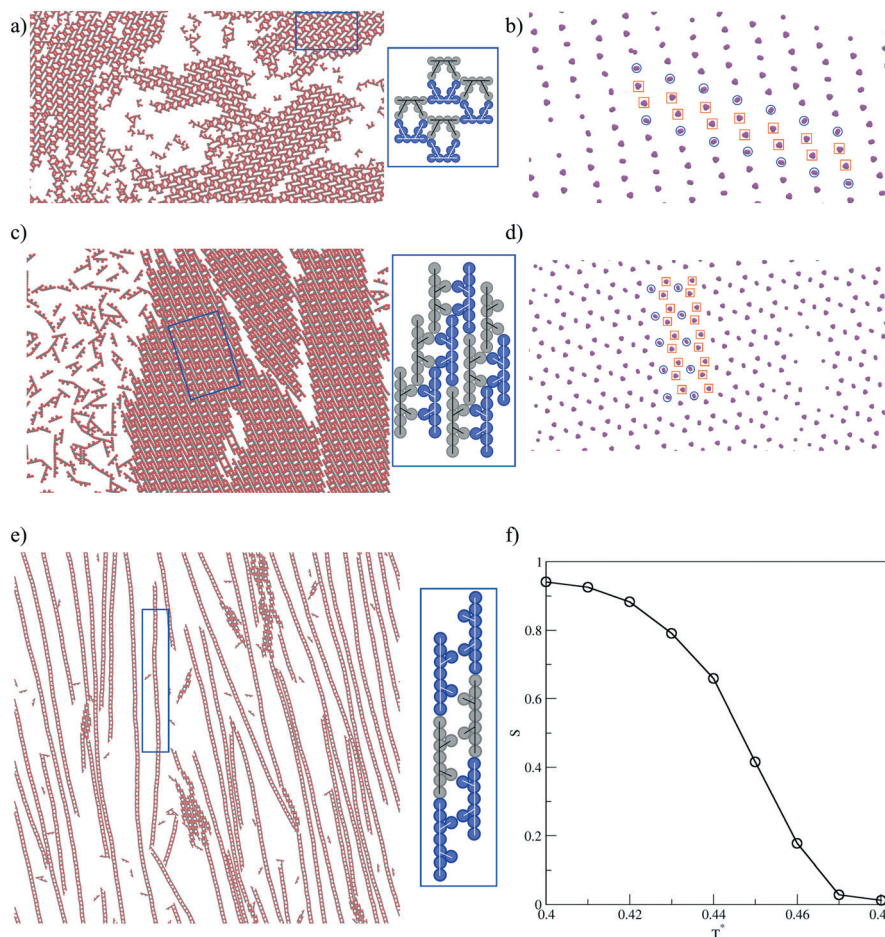
phase, with a structure similar to system **R2\_M4** is also observed.

In order to check whether the length of these wires is finite or goes to infinity, we have enlarged the system size. It has been found that the average cluster size decreases and is around  $\langle N_C \rangle \approx 315 \pm 20$  segments in every case, with the maximum value of  $N_C = 4000$ , for the linear system size,  $L_x = L_y = 469\sigma$  ( $N_C = 5800$  for the linear system size,  $L_x = L_y = 610\sigma$ ), and the number of lines of different lengths increased when the system size becomes larger. This means that the wires reach their certain threshold length and they are not growing further upon the increase of the system size.

Another interesting feature is that this structure vanishes when the density becomes high enough. Instead of gluing with another as one could suspect. We have performed simulations for the system in  $\rho^* = 0.5$  and in this case, the dense phase prevails at  $T^* = 0.5$  (see Fig. S3 in the ESI†). However, it is important to note that at higher temperatures, between  $T^* = 0.59 - 0.56$ , lines that are glued to the growing dense phase occur, which is not the case for temperatures below  $T^* = 0.56$  (cf. Fig. S3 in the ESI†).

In order to quantify this quite unexpected structure, we wanted to see that these lines do not rotate and are arranged in one direction. The parameter that can be used here is the two-dimensional nematic order parameter,<sup>36</sup> defined as





**Fig. 5** Fragments of the configurations for models **R2\_M3** at  $\rho^* = 0.42$  and  $T^* = 0.45$  (a) and **R2\_M4** at  $\rho^* = 0.5$  and  $T^* = 0.5$  (c). Parts b) and d) show the locations of active sites for the snapshots shown in parts a) and c), respectively. The configuration for model **R3\_M4** at  $\rho^* = 0.2$  and  $T^* = 0.4$  (e). Part f) displays the temperature changes of the nematic order parameter for model **R3\_M4**. Schematically drawn association paths are shown on the right-hand side of panels a), c), and e).

$$Q_{\alpha\beta} = N^{-1} \sum_i^N [2b_{\alpha}(i)b_{\beta}(i) - \delta_{\alpha\beta}] \quad (1)$$

$$\kappa = 2 \frac{\lambda_x^4 + \lambda_y^4}{(\lambda_x^2 + \lambda_y^2)^2} - 1 \quad (2)$$

where  $b_{\alpha}(i)$  is the  $\alpha$ -th coordinate of the unit vector  $b$ , specifying the orientation of the molecule  $i$ , and  $\delta_{\alpha\beta}$  is the Kronecker delta function. Since the eigenvalues of  $Q$  are equal to  $\pm S$ , the order parameter may be positive or negative. However, the absolute value of  $S$  tends to 1 in a perfectly ordered phase, and is expected to vanish in a disordered phase when  $N \rightarrow \infty$ . In real systems, it is very difficult to reach the value of  $|S|$  equal to 1, owing to possible imperfections in the ordered structure, or rotations of differently oriented domains. Fig. 5f shows how the nematic order parameter changes with temperature. At the lowest considered temperature ( $T^* = 0.4$ ) the order parameter reaches the value of about 0.95, allowing us to conclude that the wires are aligned along one direction.

Moreover, to prove that the ladder structure is indeed ordered into straight lines, we have computed the relative shape anisotropy parameter,<sup>37</sup> defined as

where  $\lambda_x$  and  $\lambda_y$  are the diagonal eigenvalues of the gyration tensor. This quantity, like the nematic order parameter, takes values between 0 and 1, when the system is spherically symmetric or the molecules lie along a straight line, respectively.

This parameter has been calculated separately for every wire extracted from the cluster analysis, and its value was found to vary between  $\kappa = 0.85$ – $0.95$ . These values are good enough to prove that the structure consists of nearly straight lines. It is important to note that we have also observed lower values of  $\kappa \approx 0.2$ , which corresponded to the clusters in which the elements of the ladder network were connected to a more dense phase.

Next, we proceed to the description of the last model **M5**. It has appeared that similar to model **M2**, the molecules associate into Sierpinski triangles and into a more dense phase. In general, the behavior of model **M5** is qualitatively very similar to that found for model **M2**. In particular, the changes of the



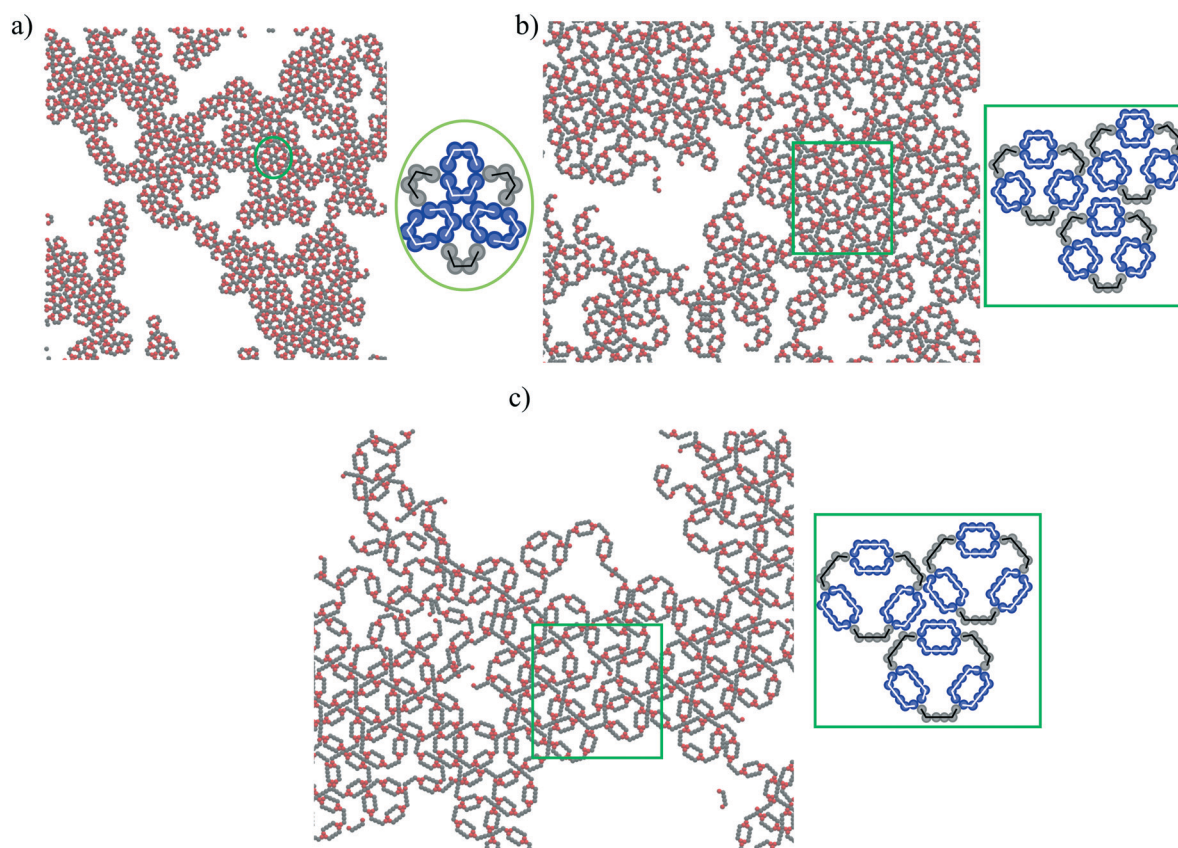


structure when the core length,  $\mathbf{R}$ , and the density are varied remain the same. For the sake of brevity, we have omitted a detailed description of the results in the main text, and the reader is advised to look at Fig. S4 and S5 in the ESI†

**2.2.2 Results for  $l = 0.36\sigma$  and  $l = 0.44\sigma$ .** Let us start with model **M1** of the increased bonding distance  $l = 0.44\sigma$ . In this case, three molecules connect and form completely different networks than in the systems with  $l = 0.4\sigma$  (cf. Fig. 3). The results for different core lengths  $\mathbf{R}$  are shown in Fig. 6(a–c). Similar to the systems with the bonding distance  $l = 0.4\sigma$ , model **R2\_M1** does not form any well-defined ordered network. There are several patterns formed by the associating molecules. Likewise in the case of shorter  $l$ , the core elongation  $\mathbf{R}$  leads to the formation of a single dominant network, and a “flower-like” motif appears. At higher temperatures, those small aggregates are randomly connected but do not form a single well-defined network. A decrease in the temperature leads to the condensation of those patterns, as depicted in Fig. 6(b and c). Unfortunately, it is difficult to determine the structure of this ordered network. We have tried to compute the two-dimensional structure factors, which should show the type of ordering,<sup>30</sup> but could not obtain statistically relevant results. Therefore, in our opinion, the increase of the system size could lead to the formation of better-ordered larger networks.

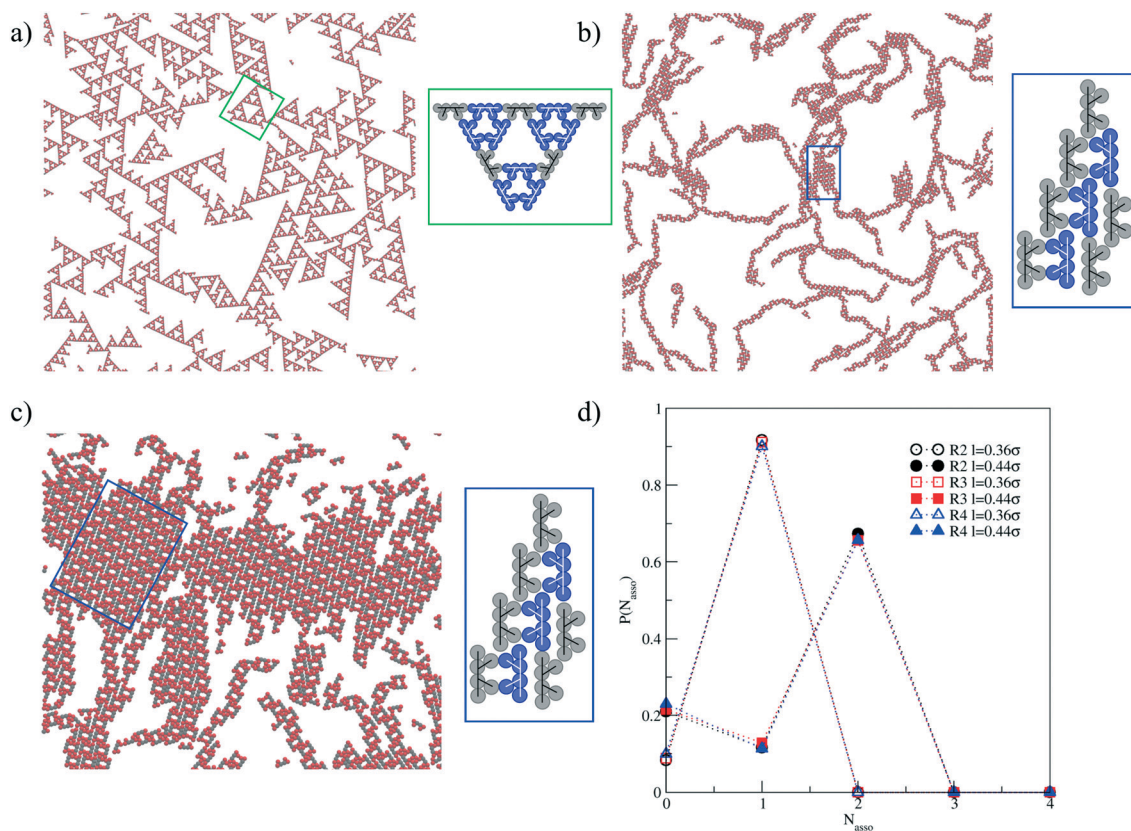
Next, we proceed to the description of model **M2**. Since in the case of the bonding length  $l = 0.40\sigma$ , two different structures were found, we have considered the systems with lower and higher  $l$ , to see whether the bonding length affects self-assembly. The results for  $l = 0.36\sigma$  are presented in Fig. 7(a). In this case, the molecules form only Sierpinski triangles and no other network can be distinguished. It is also important to note that in comparison with the systems of  $l = 0.4\sigma$ , the formation of Sierpinski triangles of the third degree with some defects can be observed. Contrary to this, the longer  $l = 0.44\sigma$  (Fig. 7b and c) stabilizes a more dense phase and no self-similar structures are observed, both in lower ( $\rho^* = 0.2$ ) and higher densities ( $\rho^* = 0.4$ ). The same tendency remains for different core elongations  $\mathbf{R}$  (see Fig. S6 in the ESI†).

As previously, we have calculated the distributions of the association numbers for models **R2–R4\_M2** with different bond lengths,  $l$  and the results are given in Fig. 7(d). We can see that for bond length  $l = 0.36\sigma$  the highest value of  $N_{\text{asso}}$  is equal to 1, confirming that only two molecules connected one to another, and form Sierpinski triangles. On the other hand, for the molecules with bond length  $l = 0.44\sigma$ , the most probable value of the association number is equal to  $N_{\text{asso}} = 2$ . One can also observe a peak at  $N_{\text{asso}} = 1$ , and these results corroborate with the formation of a more dense phase. It is also worth mentioning that the course of the distribution



**Fig. 6** Parts of the configurations for models **R2\_M1** (a), **R3\_M1** (b), and **R4\_M1** (c) at  $\rho^* = 0.4$  and  $T^* = 0.42$ , and for  $l = 0.44\sigma$ . Schematically drawn association paths are shown on the right-hand sides of the panels.





**Fig. 7** The configurations for model **R2\_M2** at  $\rho^* = 0.2$  and  $T^* = 0.25$  for  $l = 0.36\sigma$  (a) and at  $\rho^* = 0.2$  and  $T^* = 0.4$  for  $l = 0.44\sigma$  (b). The snapshot for model **R2\_M2** at  $\rho^* = 0.4$  and  $T^* = 0.65$  for  $l = 0.44\sigma$  (c). Schematically drawn association paths are shown on the right-hand sides of the panels. Distributions of the association number for model **R2-R4\_M2** with the density equal to  $\rho^* = 0.2$ , recorded at  $T^* = 0.25$  and using  $l = 0.36\sigma$  (open symbols) and recorded at  $T^* = 0.50$  with  $l = 0.44\sigma$  (filled symbols) (d).

function does not change within the range of the core length values **R2-R4**.

The results for model **M3** with the bond length  $l = 0.36\sigma$  can be found in Fig. 8. In the case of  $l = 0.40\sigma$ , we have observed only one phase, in which the connections of two and three molecules were possible. Now, the results are significantly different (*cf.* Fig. 5(a)). For model **R2\_M2** (Fig. 8(a)) one can see a lot of small aggregates that are similar to Sierpinski triangles of the first order. However, we have not observed the formation of a single ordered network. Those small clusters glue randomly and form complex structures.

Contrary to the bonding distance  $l = 0.4\sigma$ , where the core length **R** did not change qualitatively the structure, for  $l = 0.36\sigma$  we observe significant changes. The structure for both **R3\_M3** (Fig. 8(b)) and **R4\_M3** (Fig. 8(c)) is the same isogonal tiling, which has been previously observed for model **M1** (*cf.* Fig. 3). Likewise, we have examined the arrangement of active sites and this kind of representation reveals similar ordering to the case of model **M1** with  $l = 0.4\sigma$ . The connection of the centers of mass of two associated active sites results in the formation of 3.4.6.4 Archimedean tessellation in models **R3\_M3** (Fig. 8(d)) and **R4\_M3**.

These changes of the behavior with the core length are due to the difference in saturation per molecule. For the **R2\_M3** model, small clusters are formed by six molecules in

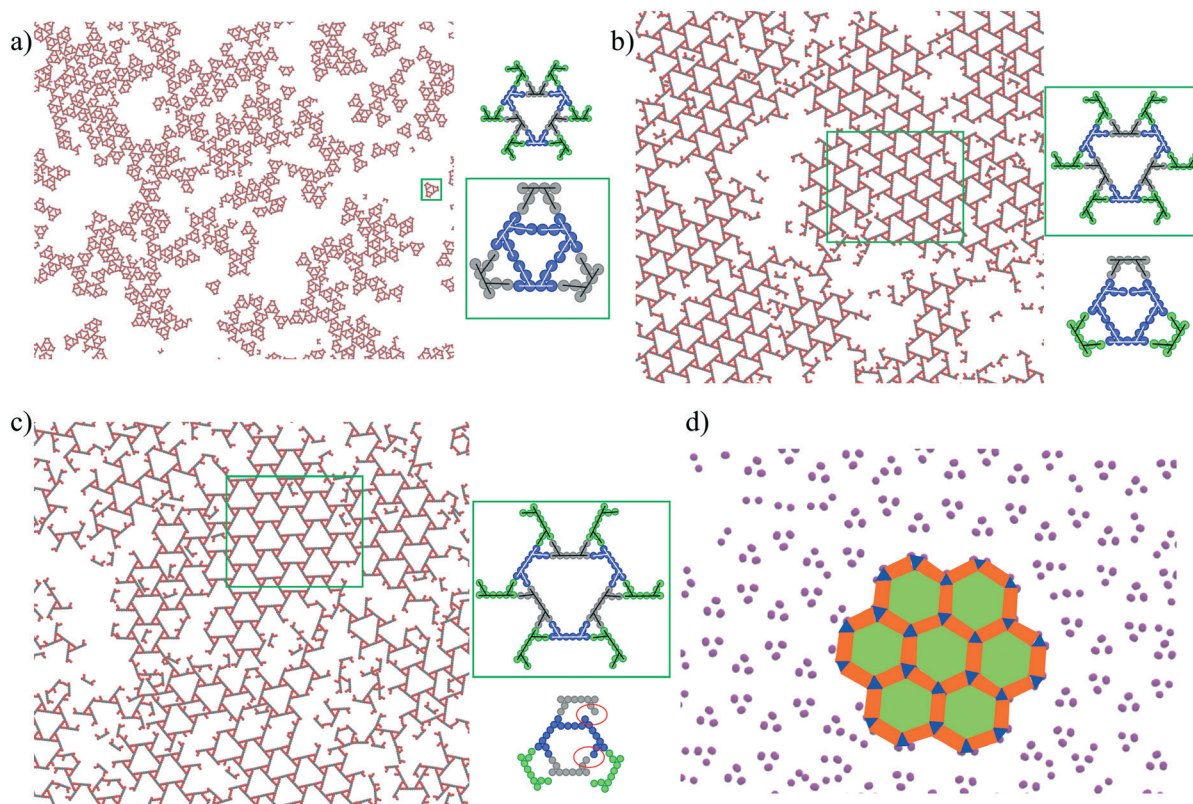
which three tectons have every active site associated. To form a seed, leading to an isogonal tiling, there have to be at least twelve molecules.

The probability of formation of such a group is rather low, and energetically less favorable in this case. On the contrary, in model **R3\_M3**, the core elongation leads to a slight mismatch in the direction of association between active sites; hence the formation of small aggregates is energetically less favorable compared to an isogonal tiling. This effect is even better pronounced for model **R4\_M3**.

Let us now proceed to model **M4**. For the bond length  $l = 0.4\sigma$  we have observed the formation of only one well-defined ordered network, for all systems with different core lengths  $R = 2, 4$ . Surprisingly, for greater embedment the formation of a similar isogonal tiling, as in models **M1** and **M3** with the same  $l = 0.4\sigma$ , has been observed (Fig. 9(a)). Nevertheless, it has to be emphasized that in the case of  $l = 0.36\sigma$ , this structure has not been formed for the core length  $R = 2$ . However, the arrangement of active sites is quite different than for those isogonal tilings observed previously. Despite the formation of qualitatively the same network if one would connect the center of mass of every two associated active centers, the formation of a structure resembling a Kagomé network can be observed (Fig. 9(b)). For model **R4\_M4** (Fig. 9c), similar to **R2\_M4**, the more dense







**Fig. 8** Parts a), b), and c) present the configurations for models **R2–R4\_M3** with  $l = 0.36\sigma$ , at  $\rho^* = 0.2$  and  $T^* = 0.28$ . Schematically drawn association paths are shown on the right-hand sides of the panels. Part d) shows how the active sites in model **R3\_M3** assemble into Archimedean tessellation. Different polygons are marked by different colors to better pronounce the ordering.

phase that has been observed for  $l = 0.4\sigma$  has not been formed and an isogonal tiling is the only ordered network. The active site layout is comparable to the **R2\_M4** model and has been omitted for the sake of brevity.

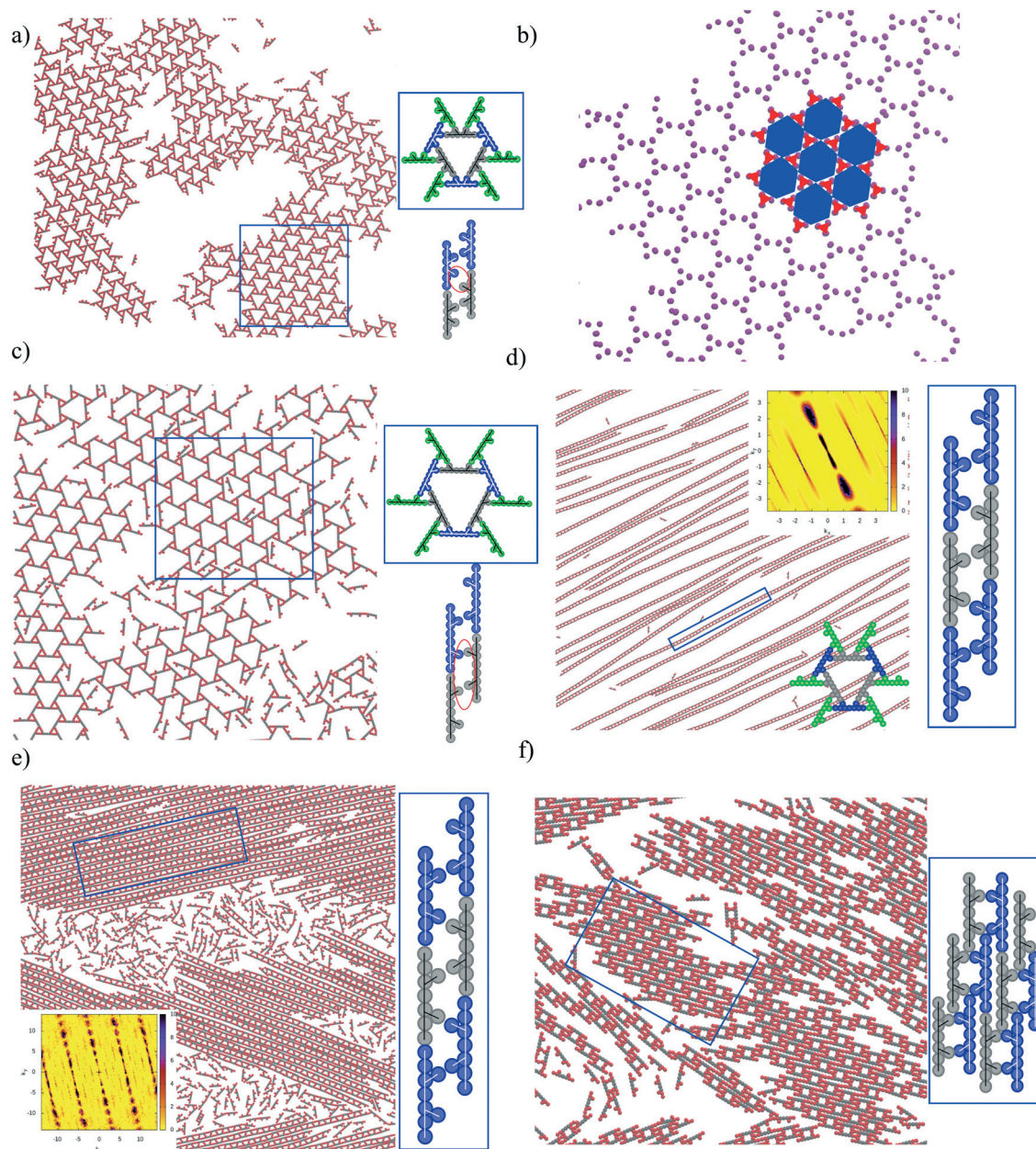
For  $l = 0.4\sigma$ , elongation of the core length leads to the formation of finite one-dimensional wires (ladder network) and a more dense phase. By changing the embedding distance for model **R3\_M4** to  $l = 0.36\sigma$ , two phases are no longer present and the formation of a ladder network is preferred (Fig. 9(d)). The two-dimensional structure factor calculated with respect to the backbone segments, and given in the inset to part (d) of Fig. 9 confirms the alignment of the ladder network along one direction.

On top of that is diffused and the distances in reciprocal space are small which corroborates with the observation that those wires are separated (large distance in real space). Contrary to  $l = 0.4\sigma$ , the increase of density does not lead to structural changes of the system, and the wires start to glue laterally one to another, instead (Fig. 9(e)). This resembles the behavior of liquid crystals which, even without the presence of attractive interactions, tend to form densely-packed, ordered networks.<sup>38</sup> The structure factor calculated for  $\rho = 0.5$  is shown in the inset of Fig. 9(e). The wires are aligned along one direction, but the reflexes are not diffused. The distances in the reciprocal space are larger than at lower density, which corroborates with the observation from snapshots.

For every core length **R2–R4** we have geometrically analyzed a possibility to form either isogonal tiling or a ladder network. We have found that due to the architecture of molecules **R2** and **R4** the formation of the former is the only possibility (*cf.* right-hand side of panels (a) and (c) of Fig. 9). We conclude that the reason is an incomplete saturation of possible interparticle interactions in the formation of one-dimensional wires. On the other hand, for **R3** both structures reach the same saturation, and they only differ by the packing fraction of the networks. Therefore, the formation of a denser phase should be preferred.

For the molecules **R3\_M4** and with  $l = 0.36\sigma$ , we have performed cluster analysis which has shown that the linear clusters are nearly two times longer relative to the case of  $l = 0.40\sigma$ . The average cluster size was estimated to consist of  $\langle N_C \rangle = 776 \pm 50$  segments with the maximum value of  $N_C = 6687$  (linear system size,  $L_x = L_y = 336\sigma$ ). Similarly, the relative shape anisotropy parameter varied between  $\kappa \approx 0.92 \pm 0.05$ . However, it has to be emphasized that in this system, no clusters with smaller values have been observed, which in the system with  $l = 0.40\sigma$  was caused by the occurrence of the second, more dense phase. On the other hand, an increase of the bonding distance to  $l = 0.44\sigma$  results in the formation of a more dense phase in systems at both low ( $\rho^* = 0.2$ ) and high ( $\rho^* = 0.5$ ) densities and the ladder networks have not been observed at all (Fig. 9f).





**Fig. 9** Fragment of the configuration for model **R2\_M4** recorded at  $\rho^* = 0.2$  and  $T^* = 0.28$  for  $l = 0.36\sigma$  (a). Part b) shows how the active sites of model **R2\_M4** assemble into a Kagomé-like pattern. Different polygons are marked by different colors to better pronounce the ordering. Part c) shows a part of the configuration for model **R4\_M4** recorded at  $\rho^* = 0.2$  and  $T^* = 0.28$  for  $l = 0.36\sigma$ . Configurations for model **R3\_M4** recorded at  $\rho^* = 0.2$  and  $T^* = 0.28$  for  $l = 0.36\sigma$  (d), at  $\rho^* = 0.5$  and  $T^* = 0.42$  for  $l = 0.36\sigma$  (e), and at  $\rho^* = 0.5$  and  $T^* = 0.67$  for  $l = 0.44\sigma$  (f). 2-D structure factors are shown in the insets to parts (d) and (e). Schematically drawn association paths are shown on the right-hand side of the panels.

Finally, we proceed to the description of the last architecture of the tetra-topic building blocks *i.e.*, model **M5** with bond lengths  $l = 0.36\sigma$  and  $l = 0.44\sigma$ . Similar to model **M2**, the molecules associate into Sierpinski triangles and into a more dense phase. The effects of the elongation of core length **R** as well as the change of density have been found to be the same as in model **M2** (*cf.* Fig. 7). For the sake of brevity, we have omitted the details of analysis in the main text and one should refer to Fig. S7 and S8 in the ESI† for more details.

### 3 Conclusions

In this study, we have performed extensive coarse-grained molecular dynamics simulations of the self-assembly of tetra-substituted molecules. We have found that they are able to form a variety of different structures depending on the parameters of the employed model. All of these ordered networks have been characterized by cluster analysis, the nematic order parameter, the relative shape anisotropy parameter, and the two-dimensional structure factors.





We have established the following rules depending on the interaction zone  $\alpha$  (cf. Fig. 2) and the architecture of tetra-topic molecules. In the case of building blocks with the same numbers of segments in each of the arms (**M2**, **M5**), for  $l = 0.4\sigma$ , we could distinguish two different networks. Slight changes of the interaction zone allowed us to selectively obtain either Sierpinski triangles ( $l = 0.36\sigma$ ) or a denser phase ( $l = 0.44\sigma$ ) for all core lengths **R**. On the other hand, the molecules with unequal numbers of segments in the arms have demonstrated another tendency. For bond lengths  $l = 0.36\sigma$  and  $l = 0.40\sigma$  they assemble into isogonal tiling and into a denser phase, respectively. In the former case, it is worth mentioning that the representation of active sites reveals that although those models form qualitatively the same network, geometrical centers of associations form a 3.4.6.4 Archimedean tessellation (**M3**) or Kagomé network (**M4**).

However, there are two exceptions. For model **R2\_M3** we have observed the formation of small clusters only, which are similar to the STs of the first-order, whereas the behavior of molecules **R3\_M4** is more complex. For the bond length  $l = 0.4\sigma$  we have seen the formation of the ladder network and a denser phase. However, the former disappears at high densities. Similar to models **M2**, **M5**, we have been able to selectively obtain each of those networks depending on the bonding distance. Surprisingly, for  $l = 0.36\sigma$  we have found that the wires are much longer as compared to the greater jointing distance. Moreover, they are gluing one to another at higher densities, and do not form a denser phase. On the other hand, for the system with  $l = 0.44\sigma$ , the denser phase is observed even at lower densities, and a ladder network has not been found at all.

Another interesting feature found is that denser phases are qualitatively the same for every model irrespective of their architecture. The connections between molecules are the same; however, the pore sizes are different due to the geometry of investigated models. This observation is better pronounced by the visualization of the active site ordering.

Moreover, we have found that model **M1**, irrespective of the bonding length  $l$ , assembled into unique structures compared to other examined systems. For  $l = 0.4\sigma$  isogonal tiling, whereas for  $l = 0.44\sigma$  a flower-like motif has been found.

Despite the formation of a variety of different structures it has to be emphasized that ordering of tetra-topic molecules into Sierpinski triangles is quite surprising since the majority of the papers report their formation in completely different systems. The only general rule that has been established to date is that the so-called “V-shaped” molecules are able to order into those fractals, and only if three tectons assemble either by hydrogen bonding (B(OH)<sub>2</sub> substitution) or by metal–organic coordination (for instance, –CN groups with Fe or Ni metal atoms). Here we have shown another possible scheme, together with a proposal on how this kind of model chemical compound could look like (cf. Fig. 1(c)). As already mentioned, if one would substitute it with active groups that would allow two molecules to interact with one another (–OH, COOH, or –Cl with Fe(II)) it

corresponds to the  $l = 0.36\sigma$  case, whereas interactions of three building blocks (–B(OH)<sub>2</sub>, or –Cl with Fe(III)) can reflect results for  $l = 0.44\sigma$ . On the other hand, the intermediate scenario where two phases occur which is most likely the case in experimental studies due to the oscillations of chemical bonds of active groups can be seen for  $l = 0.40\sigma$ . Therefore, we believe that our finding may be inspiring for experimentalists.

We have presented a variety of different structures and their analysis. Moreover, we have presented a molecular dynamics model and shown how a slight change of its parameters can drastically change the behavior of systems of interest. We have demonstrated that the possibility of a particular number of molecules that can assemble with one another can determine the formation of completely different structures. As already mentioned, multiple chemically active groups can display diverse behavior depending on the experimental conditions. Therefore, we hope that our observations can be very useful for the better understanding and possibility of selective control of the self-assembly in future experimental studies.

## Conflicts of interest

There are no conflicts of interests.

## Acknowledgements

This study has been supported by the Polish Ministry of Science and Higher Education, under Grant No. DI2017 001147 and the research has been supported by the Foundation for Polish Science (FNP).

## References

- 1 C. C. Ng, Y.-L. Cheng and P. S. Pennefather, Properties of a self-assembled phospholipid membrane supported on lipobeads, *Biophys. J.*, 2004, **87**, 323–331.
- 2 L. Nasalean, S. Baudrey, N. B. Leontis and L. Jaeger, Controlling RNA self-assembly to form filaments, *Nucleic Acids Res.*, 2006, **34**, 1381–1392.
- 3 J. Won, S. K. Chae, J. H. Kim, H. H. Park, Y. S. Kang and H. S. Kim, Self-assembled DNA composite membranes, *J. Membr. Sci.*, 2005, **249**, 113–117.
- 4 R. F. Garmann, A. M. Goldfain and V. N. Manoharan, Measurements of the self-assembly kinetics of individual viral capsids around their RNA genome, *Proc. Natl. Acad. Sci. U. S. A.*, 2019, **116**, 22485–22490.
- 5 S. R. White, N. R. Sottos, P. H. Geubelle, J. S. Moore, M. R. Kessler, S. R. Sriram, E. N. Brown and S. Viswanathan, Autonomic healing of polymer composites, *Nature*, 2001, **409**, 794–797.
- 6 V. L. Colvin, From opals to optics: Colloidal photonic crystals, *MRS Bull.*, 2001, **26**, 637–641.
- 7 H. Stoyanov, M. Kollosche, S. Risse, R. Waché and G. Kofod, Soft conductive elastomer materials for stretchable



- electronics and voltage controlled artificial muscles, *Adv. Mater.*, 2013, **25**, 578–583.
- 8 M. D. Lima, N. Li, M. Jung de Andrade, S. Fang, J. Oh, G. M. Spinks, M. E. Kozlov, C. S. Haines, D. Suh, J. Foroughi, S. J. Kim, Y. Chen, T. Ware, M. K. Shin, L. D. Machado, A. F. Fonseca, J. D. W. Madden, W. E. Voit, D. S. Galvão and R. H. Baughman, Electrically, chemically, and photonically powered torsional and tensile actuation of hybrid carbon nanotube yarn muscles, *Science*, 2012, **338**, 928–932.
  - 9 C. Zhang, B.-H. Wu, M.-Q. Ma, Z. Wang and Z.-K. Xu, Ultrathin metal/covalent-organic framework membranes towards ultimate separation, *Chem. Soc. Rev.*, 2019, **48**, 3811–3841.
  - 10 S. Wang, Q. Wang, P. Shao, Y. Han, X. Gao, L. Ma, S. Yuan, X. Ma, J. Zhou, X. Feng and B. Wang, Exfoliation of covalent organic frameworks into few-layer redox-active nanosheets as cathode materials for lithium-ion batteries, *J. Am. Chem. Soc.*, 2017, **139**, 4258–4261.
  - 11 G. Li, K. Zhang and T. Tsuru, Two-dimensional covalent organic framework (COF) membranes fabricated via the assembly of exfoliated COF nanosheets, *ACS Appl. Mater. Interfaces*, 2017, **9**, 8433–8436.
  - 12 L. Wang, R. Liu, J. Gu, B. Song, H. Wang, X. Jiang, K. Zhang, X. Han, X.-Q. Hao, S. Bai, M. Wang, X. Li, B. Xu and X. Li, Self-assembly of supramolecular fractals from generation 1 to 5, *J. Am. Chem. Soc.*, 2018, **140**, 14087–14096.
  - 13 Z. Jiang, D. Liu, M. Chen, J. Wang, H. Zhao, Y. Li, Z. Zhang, T. Xie, F. Wang, X. Li, G. R. Newkome and P. Wang, Assembling shape-persistent high-order Sierpiński triangular fractals, *iScience*, 2020, **23**(5), 101064.
  - 14 D. Nieckarz and P. Szabelski, Simulation of the self-assembly of simple molecular bricks into Sierpiński triangles, *Chem. Commun.*, 2014, **50**, 6843–6845.
  - 15 J. Shang, Y. Wang, M. Chen, J. Dai, X. Zhou, J. Kuttner, G. Hilt, X. Shao, J. M. Gottfried and K. Wu, Assembling molecular Sierpiński triangle fractals, *Nat. Chem.*, 2015, **7**, 389–393.
  - 16 Y. Mo, T. Chen, J. Dai, K. Wu and D. Wang, On-surface synthesis of highly ordered covalent Sierpiński triangle fractals, *J. Am. Chem. Soc.*, 2019, **141**, 11378–11382.
  - 17 C. Li, X. Zhang, N. Li, Y. Wang, J. Yang, G. Gu, Y. Zhang, S. Hou, L. Peng, K. Wu, D. Nieckarz, P. Szabelski, H. Tang and Y. Wang, Construction of Sierpiński triangles up to the fifth order, *J. Am. Chem. Soc.*, 2017, **139**, 13749–13753.
  - 18 Y. Zhang, X. Zhang, Y. Li, S. Zhao, S. Hou, K. Wu and Y. Wang, Packing Sierpiński triangles into two-dimensional crystals, *J. Am. Chem. Soc.*, 2020, **142**, 17928–17932.
  - 19 D. Nieckarz and P. Szabelski, Chiral and fractal: from simple design rules to complex supramolecular constructs, *Chem. Commun.*, 2016, **52**, 11642–11645.
  - 20 A. Rochefort and J. D. Wuest, Interaction of substituted aromatic compounds with graphene, *Langmuir*, 2009, **25**, 210–215.
  - 21 A. Ibenskas and E. E. Tornau, Modeling of ribbon and oblique structures of benzene-1,3,5-triyl-tribenzoic acid, *J. Phys. Chem. C*, 2020, **124**, 18650–18659.
  - 22 C. Karner, C. Dellago and E. Bianchi, Design of patchy rhombi: From close-packed tilings to open lattices, *Nano Lett.*, 2019, **19**, 7806–7815.
  - 23 C. Karner, C. Dellago and E. Bianchi, Hierarchical self-assembly of patchy colloidal platelets, *Soft Matter*, 2020, **16**, 2774–2785.
  - 24 A. Ibenskas, M. Šimėnas, K. J. Kizlaitis and E. E. Tornau, Pinwheel structures of deprotonated trimesic acid on Ag(111): Model and simulations, *J. Phys. Chem. C*, 2020, **124**, 11212–11220.
  - 25 A. I. Fadeeva, V. A. Gorbunov, P. V. Stishenko, S. S. Akimenko and A. V. Myshlyavtsev, Melting of Feterephthalate layers on Cu(100) surface with randomly distributed point defects, *Appl. Surf. Sci.*, 2021, **545**, 148989.
  - 26 J. Lisiecki and P. Szabelski, Designing 2d covalent networks with lattice Monte Carlo simulations: precursor self-assembly, *Phys. Chem. Chem. Phys.*, 2021, **23**, 5780–5796.
  - 27 C.-A. Palma, P. Samori and M. Cecchini, Atomistic simulations of 2d bicomponent self-assembly: From molecular recognition to self-healing, *J. Am. Chem. Soc.*, 2010, **132**, 17880–17885.
  - 28 Y. Zhao and J. Wang, How to obtain high-quality and high stability interfacial organic layer: Insights from the ptdca self-assembly, *J. Phys. Chem. C*, 2017, **121**, 4488–4495.
  - 29 Ł. Baran and W. Rzyśko, Application of a coarse-grained model for the design of complex supramolecular networks, *Mol. Syst. Des. Eng.*, 2020, **5**, 484–492.
  - 30 Ł. Baran, D. Nieckarz, P. Szabelski and W. Rzyśko, Controlling of the 2d self-assembly process by the variation of molecular geometry, *J. Phys. Chem. C*, 2019, **123**, 19549–19556.
  - 31 Ł. Baran, W. Rzyśko and S. Szajnar, Archimedean tessellation found by the variation of building blocks' and linkers' geometry: In silico investigations, *J. Phys. Chem. C*, 2020, **124**, 20101–20108.
  - 32 S. Xing, Z. Zhang, X. Fei, W. Zhao, R. Zhang, T. Lin, D. Zhao, H. Ju, H. Xu, J. Fan, J. Zhu, Y.-q. Ma and Z. Shi, Selective on surface covalent coupling based on metal-organic coordination template, *Nat. Commun.*, 2019, **10**, 70.
  - 33 S. Xing, Z. Zhang, H. Liang, B. Sun, H. Xu, J. Fan, Y.-q. Ma and Z. Shi, On-surface cascade reaction based on successive debromination via metal-organic coordination template, *Langmuir*, 2020, **36**, 6286–6291.
  - 34 W. Rzyśko, S. Sokołowski and T. Staszewski, Monte Carlo simulations of a model two-dimensional, two-patch colloidal particles, *J. Chem. Phys.*, 2015, **143**, 064509.
  - 35 Y. V. Kalyuzhnyi and P. T. Cummings, Two-patch colloidal model with re-entrant phase behaviour, *J. Chem. Phys.*, 2013, **139**, 104905.
  - 36 D. Frenkel and R. Eppenga, Evidence for algebraic orientational order in a two-dimensional hard-core nematic, *Phys. Rev. A: At., Mol., Opt. Phys.*, 1985, **31**, 1776–1787.
  - 37 D. N. Theodorou and U. W. Suter, Shape of unperturbed linear polymers: polypropylene, *Macromolecules*, 1985, **18**, 1206–1214.
  - 38 P. Bolhuis and D. Frenkel, Tracing the phase boundaries of hard spherocylinders, *J. Chem. Phys.*, 1997, **106**, 666–687.

