ROYAL SOCIETY
OF CHEMISTRY

PAPER

Effects of co-ordination number on the nucleation behaviour in many-component self-assembly

Aleks Reinhardt, Chon Pan Ho and Daan Frenkel*

We report canonical and grand-canonical lattice Monte Carlo simulations of the self-assembly of addressable structures comprising hundreds of distinct component types. The nucleation behaviour, in the form of free-energy barriers to nucleation, changes significantly as the co-ordination number of the building blocks is changed from 4 to 8 to 12. Unlike tetrahedral structures – which roughly correspond to DNA bricks that have been studied in experiments – the shapes of the free-energy barriers of higher co-ordination structures depend strongly on the supersaturation, and such structures require a very significant driving force for structure growth before nucleation becomes thermally accessible. Although growth at high supersaturation results in more defects during self-assembly, we show that high co-ordination number structures can still be assembled successfully in computer simulations and that they exhibit self-assembly behaviour analogous to DNA bricks. In particular, the self-assembly remains modular, enabling in principle a wide variety of nanostructures to be assembled, with a greater spatial resolution than is possible in low co-ordination structures.

1 Introduction

Materials that can be formed by self-assembly have over time become increasingly more complex.[1] Furthermore, in the last few years, the field has seen something of an explosion in the number of self-assembling materials which exhibit not only structural complexity, but which are 'addressably' complex,[2] in the sense that the individual building blocks making up these structures are all distinct. Such self-assembled materials are not only interesting from the point of view of funda-mental science, but are thought to hold considerable promise for applications in many aspects of nanotechnology,[3] especially since the addressable nature of the building blocks should allow the structures to be functionalised with sub-nano-metre-scale resolution.

Recent experiments have demonstrated that it is possible to assemble struc-tures comprising thousands of distinct modular building blocks into well-formed

Department of Chemistry, University of Cambridge, Lensfield Road, Cambridge, CB2 1EW, UK. E-mail: df246@cam.ac.uk

target structures by making use of single-stranded DNA molecules – termed 'DNA bricks' – designed to have an obligate set of hybridisation partners, in the sense that those parts of the DNA molecules that are designed to be bonded in the target structure have complementary sequences.[4–8] In the past few years, several theoretical and computational studies have also been undertaken, probing the intriguing self-assembly behaviour exhibited by such systems.[9–18]

We have previously shown that DNA brick self-assembly is made possible by the interplay between self-assembly and growth. In particular, as a system of DNA bricks is cooled, at some temperature the free-energy barrier to nucleation becomes small enough that nucleation can occur, but nucleation events remain sufficiently rare so that any clusters that do form do not interact significantly with one another, and monomers are not initially significantly depleted,[10,16] which enables these clusters to grow in an essentially error-free manner as the temperature is decreased. However, such behaviour only occurs over a very narrow window of temperatures: if the experiment is performed at a low temperature from the outset, misassembled aggregates dominate instead.[10] Nucleation thus plays an important role in enabling structures of this type to self-assemble successfully.

Whilst DNA bricks have been shown to self-assemble reliably, our previous theoretical work has indicated that the co-ordination number of the particles that form self-assembling structures determines their nucleation behaviour in both two and three dimensions.[16] In particular, the larger the co-ordination number, the more classical-looking the free-energy barrier to nucleation becomes. Yet one of the key aspects that seemed to enable the lower co-ordinate structures to form successfully was the non-classical nucleation barrier. Specifically, for tetrahedrally co-ordinated building blocks, the critical cluster size was found to be largely insensitive to the nature of the target structure, and the nucleation barrier was significant but surmountable at the point at which a large, nearly fully assembled cluster of the designed structure is thermodynamically stable. By contrast – and in agreement with the predictions of classical nucleation theory – for higher co-ordination number structures, the free-energy barrier to nucleation changes with temperature and is considerably larger than that for tetrahedral structures at the same supersaturation.[16] This suggests that, in order to overcome the free-energy barrier to nucleation, the driving force for growth must increase, for example by increasing the monomer concentration, reducing the temperature or increasing the bond strengths by choosing a different set of DNA sequences. Such approaches, however, would make competing structures in which monomers have not assembled as designed ever more stable, and our previous theoretical work thus suggests that, as the co-ordination number increases, the structures should become more and more difficult to form.

However, in order to create more varied target structures in an addressable way, we may well need to move to a system with a higher co-ordination number, as this should in principle allow us to construct structures with finer small-scale features due to the considerably greater spatial resolution of the system than we can achieve using tetrahedrally co-ordinated particles. Moreover, a greater degree of bonding can help to stabilise such structures, which may also be important in practical applications.

Although DNA bricks are tetrahedrally co-ordinated,[4] there are many possible ways in which addressable structures with higher co-ordination numbers might be experimentally realised. For example, one can envisage that colloidal particles with carefully positioned DNA strands grafted onto the particle in the correct

geometry might be possible to assemble in the near future, perhaps similar to the experiments of Wang *et al.*[19] or Lu *et al.*,[20] but with each particle functionalised with a unique set of DNA strands. Alternatively, DNA Holliday junctions and multi-arm motifs[21] can be synthesised to correspond to high co-ordination number structures. Of course in practice, producing structures of this type in experiments may be non-trivial because, in our examples, each colloidal particle would have to be created with a unique set of grafted DNA strands, and each DNA junction with a different sequence would have to be pre-assembled. It is therefore important that future experiments focus on strategies that are likely to be successful. It is with this in mind that we have carried out the simulations presented here: if structures of this kind cannot be assembled on a computer with a toy model, then it may be risky to attempt to do so experimentally in the light of the significant cost and effort likely to be involved.

2 Simulation methods

We perform canonical ensemble simulations on a lattice, with periodic boundary conditions, using a Metropolis Monte Carlo[22] scheme. To determine the free-energy barriers as a function of the size of the largest crystalline cluster in the system, we use umbrella sampling with adaptive weights[23] in a time-step separated[24] Monte Carlo scheme. We use 'virtual moves'[25] to allow for realistic dynamics of cluster motion. In our simulations, clusters are randomly translated or rotated on a lattice, with 24 permissible orientations per particle, corresponding to all the possible neighbour interactions on a cubic lattice.[10]

Each particle in the system is hard in the sense that dual occupancy of lattice sites is not permitted, and each particle has n 'patches', where n is the co-ordination number. Every patch is assigned a DNA sequence such that, in the fully assembled target structure, adjacent patches have a complementary sequence, but otherwise these sequences are randomly assigned (subject to the rules identified by Wei *et al.*[5]).† Particles that are adjacent to each other interact with a slightly repulsive energy $\varepsilon_{\text{init}}/k_{\text{B}} = 100$ K,[10] to which we add the hybridisation free energy of the longest complementary sequence match between the nearest pair of patches, calculated using a standard thermodynamic model.[26] As in the experimental work of Ke *et al.*,[4] the outermost particles in the target structure are assigned a poly-T sequence to minimise any misbonding.

Particles which have 4 and 8 patches have a minimum interparticle distance of $a\sqrt{3}$, where a is the lattice parameter,[10] whilst particles with 12 patches have a minimum interparticle distance of $a\sqrt{2}$ to be able to accommodate the additional neighbours. This means that the effective densities are not strictly comparable, as the lower co-ordination structures have a greater excluded volume.

In grand-canonical simulations, we introduce particle addition and removal moves in addition to the canonical (virtual move) translations and rotations. Particles to be added or removed are chosen at random. A particle addition move

---

† It is by no means essential for particles in our system to interact *via* DNA hybridisation; it is sufficient that they have specific, designed interactions. In practice, however, we anticipate at this stage that DNA is the most likely candidate for an experimental realisation of such systems, and we have chosen to parameterise our model accordingly.
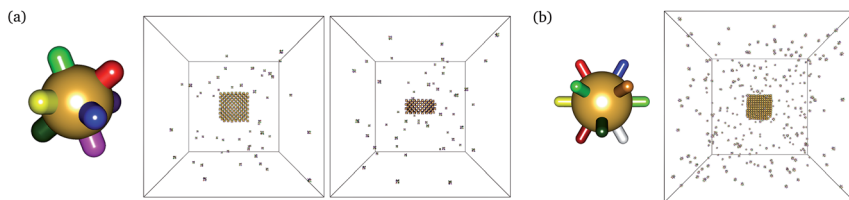
Fig. 1 A single monomer and snapshots towards the end of the nucleation process of some structures assembled from a vapour of monomers for the (a) 8 (332 K) and (b) 12 (345 K) co-ordinate monomers. The target structures in each case were simple rectangular parallelepipeds. In the simulation snapshots, correctly bonded clusters are shown in the same colour, but each particle, and each patch, is in fact distinct.

in which a particle of type i has been placed at a random position and with a random orientation in the simulation box is accepted with probability‡[27]

$$P_{acc}^{add} = \min\left[1, \frac{V z_i}{N_i + 1} \exp\left(-\Delta E / k_B T\right)\right], \tag{1}$$

where V is the volume of the simulation box, N_i is the current number of particles of type i in the system, ΔE is the trial change in the system's potential energy, and z_i is the fugacity of particles of type i, $i.e.$ $z_i = \exp(\mu_i / k_B T)$, where μ_i is the particle's chemical potential. The ideal chemical potential is given by $\mu_{id} = k_B T \ln \rho$, where ρ is the number density; in the absence of interactions, the fugacity thus determines the target number density. An analogous acceptance probability holds for particle removals,

$$P_{acc}^{rem} = \min\left[1, \frac{N_i}{V z_i N_{types}} \exp\left(\Delta E / k_B T\right)\right], \tag{2}$$

where N_{types} is the number of types of particle in the system. This accounts for the fact that when we add a particle, we choose its type uniformly at random, whereas when we remove a particle, we choose the particle at random: in order to obey detailed balance, we must account for the probability of choosing a particle of each type.

3   Results and discussion

We have previously considered tetrahedral co-ordination as applicable to DNA bricks. Here, we investigate the self-assembly behaviour of structures with co-ordination numbers of 8 (giving bcc-like target structures) and 12 (giving fcc-like target structures). The corresponding building blocks and sample target structures are shown in Fig. 1. The sequences associated with each patch for the structures we have studied are provided as supporting data.§ In the simulations reported here, the number of distinct particles in the target structures was 396 for the 4-, 403 for the 8- and 256 for the 12-co-ordinated structures.

---

‡ The de Broglie thermal wavelength is subsumed into the chemical potential, and cancels out in the case of an ideal chemical potential. For convenience, we have therefore set it to unity.

§ It is important to bear in mind that, if these sequences are chosen randomly, the temperatures at which nucleation and growth occur can change by a few degrees in identical conditions.[10] The temperatures we quote in the text refer to these specific DNA sequences. While the numerical values change with sequence choice, the qualitative behaviour does not.
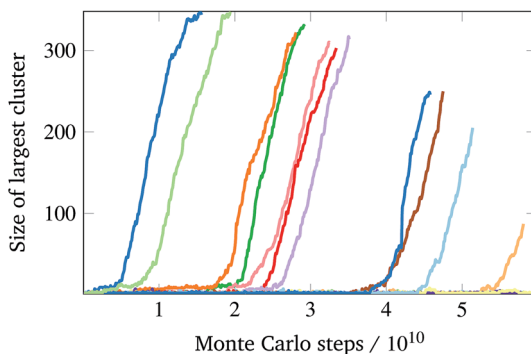
Fig. 2 The size of the largest cluster in the system as a function of Monte Carlo time for a canonical simulation with a total of 403 distinct particles with a co-ordination number of 8. $T = 332$ K, $\rho = 1/(62a)^3$. The different colours correspond to individual Monte Carlo trajectories starting from an equilibrated vapour of monomers. These trajectories were run for a fixed real-time; since virtual moves make simulations of larger clusters slower, simulations in which nucleation occurred later could run for a larger number of Monte Carlo steps.

Contrary to expectation,[16] brute-force simulations starting from a vapour of one copy of each of the monomers required to assemble a single target structure can, within a narrow temperature window, result in the successful self-assembly of the target structures shown in Fig. 1. Moreover, as evident from Fig. 2, this process is stochastic: under identical thermodynamic conditions, systems can exhibit drastically different lag times before any significant growth occurs. This is indicative of the presence of a free-energy barrier to nucleation, whereby a cluster of a sufficient size must form spontaneously before further growth is thermodynamically favoured. Since monomers coming together to form such a cluster lose a significant amount of translational and orientational entropy, this happens infrequently: there is a free-energy barrier associated with nucleation. Using umbrella sampling, we have calculated this free-energy barrier¶ for the two target structures shown in Fig. 1 at a number of temperatures, as shown in Fig. 3, where we also show a free-energy barrier for a reference tetrahedral system. Of course higher co-ordination structures are more stable at higher temperatures, since such structures entail many more bonds, and so the temperature scale at which nucleation occurs depends on the co-ordination number. Figs 2 and 3 indicate that the process is indeed nucleation-initiated for both the 8- and 12-co-ordinated target structures. However, whilst the process remains nucleation-initiated, there are significant differences in the systems' behaviours relative to the self-assembly of tetrahedral particles.

In particular, tetrahedrally co-ordinated structures, which include the experimentally studied DNA bricks, exhibit a free-energy barrier with a distinct jagged

---

¶ The order parameter used as a collective variable, *i.e.* the number of particles in the largest cluster, is a convenient choice consistent with classical nucleation theory. However, because each particle is different in these simulations, any *particular* cluster that forms can behave rather differently from this averaged behaviour. This is especially important if the cluster under consideration forms near a face or an edge of the target structure, where the average environments are different from those at the centre of the structure.
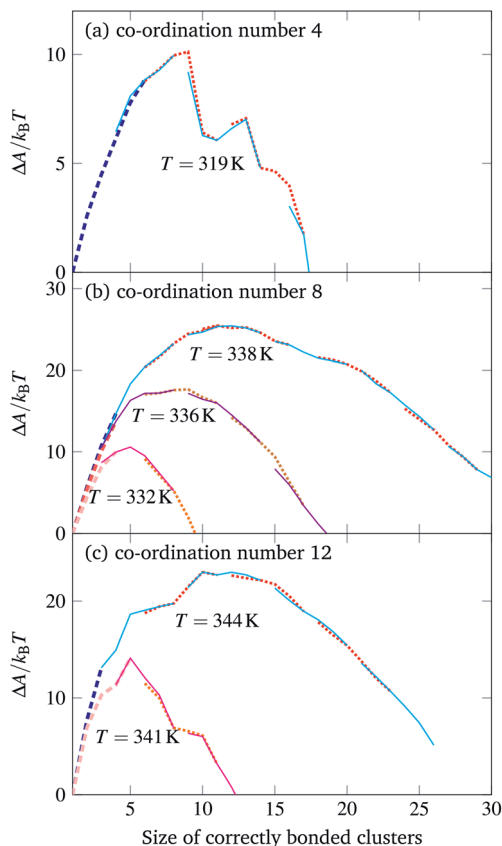
Fig. 3 The free-energy profile for cluster growth of particles with a co-ordination number of (a) 4, (b) 8 and (c) 12. Simulation results from different umbrella sampling windows are depicted in alternating styles to show their overlap. The thick dashed line corresponds to brute-force simulations.[10] In each case, there was one copy of each particle in a simulation box of dimensions $62a \times 62a \times 62a$.

appearance. This is not an artefact of the simulation technique used or a lack of equilibration, but rather reflects the fact that as clusters grow, there is a competition between the entropy loss associated with monomers losing their translational and orientational degrees of freedom when they are attached to a larger cluster and the energy gain associated with the formation of 'designed' interactions, which are, by construction, highly favourable. Tetrahedral structures grow in a very predictable fashion, with steps at which clusters can form closed cycles, for which the entropic penalty is compensated by not one, but two designed bonds forming, having a considerably lower free energy than other steps do.[10,15,16] The critical cluster for tetrahedrally co-ordinated structures is typically bicyclic or tricyclic (adamantane-like) with a single particle missing,[10,15,16] *i.e.* the size of the critical cluster is typically 8 or 9, and this cluster size appears to be essentially temperature independent in the regime where nucleation can occur. By contrast, the free-energy profiles shown in Fig. 3, in agreement with our

theoretical prediction,[16] are considerably smoother, and the temperature greatly affects the size of the critical cluster. The reason for this behaviour is that there are a considerably larger number of possibilities of forming different clusters comprising the same number of building blocks;[16] this makes the nucleation considerably more classical, affecting both the smoothness and the dependence of the critical nucleus size on temperature.

However, despite this quite different behaviour at high temperatures, the systems behave in a less divergent manner at temperatures where the nucleation barrier is sufficiently small compared to the thermal energy that nucleation can reasonably be expected to occur. The degree of supercooling required in order to observe a nucleation event is not significantly different amongst the structures we have studied: if we deem the temperature at which a pre-formed target structure fully 'melts' to be an effective melting point, nucleation becomes sufficiently fast to observe in brute-force simulations at a supercooling of approximately 2% for all target structures considered. The point at which mass aggregation occurs is also similar, at roughly 4% supercooling. These results indicate that a more optimistic view of the possibility of assembling high co-ordination number structures is perhaps warranted.

Nevertheless, one difference in the behaviour observed is noteworthy. At temperatures at which there is a reasonably small free-energy barrier to nucleation, the driving force for growth is considerably larger for higher co-ordination number structures. One proxy for this is the gradient of the free-energy profile at post-critical cluster sizes: this gradient has roughly the same value ($\sim -1.1k_BT$ per particle) in the tetrahedral case where the critical free-energy barrier height is approximately $10k_BT$, and in the 8-co-ordinate structure at 338 K with a critical free-energy barrier height of $25k_BT$. As the temperature is decreased, the effective supersaturation increases: at 332 K, the large-cluster gradient of the free energy is already $-5k_BT$ per particle. This means that the conditions in which the 8- and 12-co-ordinate structures grow are considerably harsher than in the tetrahedral case, which is likely to lead to more mistakes during assembly.[11]

In simulations where only one particle of each component is present, the increased supersaturation may not interfere with correct self-assembly, since competing structures are less likely to occur. Of course, in experiments, many copies of each building block are present. To investigate whether higher co-ordination number structures can still form in circumstances where competition from additional monomers and clusters is possible, we have also simulated the self-assembly process in the grand-canonical ensemble. We have run simulations at a fugacity corresponding to the same ideal number density as in the canonical simulations, starting from an empty simulation box of various volumes, and we observe successful self-assembly to completion at a number of temperatures for both the 8- and 12-co-ordinate structures.‖ Correctly assembled clusters grow one-by-one in such simulations: at sufficiently high temperatures, nucleation remains a rare event and the clusters grow essentially to completion before additional clusters nucleate. This observation supports the conclusion from canonical simulations that nucleation helps to prevent cluster interactions. These

---

‖ In the grand ensemble, the stability of the target structure at temperatures at which nucleation occurs changes with the co-ordination number: for the tetrahedral structures, partially formed structures dominate, whilst for high co-ordination structures, essentially fully formed structures result at the end of the self-assembly process.

grand-canonical simulations also confirm that the lack of competition from monomers and clusters in solution is not the principal reason why self-assembly can succeed in the canonical ensemble, and the self-assembly process is surprisingly robust.

Despite this apparent success, the prediction that the greater supersaturation leads to more defects does hold. If we compare the largest assembled structures in the grand ensemble at the highest temperature at which nucleation was found to occur for co-ordination numbers of 4 and 12 (319 K and 338 K, respectively), the high co-ordination number structures typically have one or two incorrect particles embedded in the structure, and one or two vacancies, whilst the tetrahedral structure is entirely error free. The error rates would, moreover, be expected to be higher still if we implemented a 'kinetic constraint' to prevent the change of state for any particle wholly within the solid structure to account for the relative slowness of the relaxation dynamics within a solid structure:[28] this would, in particular, prevent vacancies from being filled when the rest of the structure has already formed around them. While the number of defects in absolute terms is not large even for the high co-ordination number structures, it is worth bearing in mind that incorrect particles on the surface of the cluster can lead to additional undesired clustering as the temperature is lowered and the clusters are allowed to undergo diffusion for long periods of time.

One way in which the driving force for nucleation can be changed is by strengthening or weakening the average bond energy between particles. When using DNA bases, this can be achieved by varying the proportion of G and C bases at the expense of A and T: the larger the GC content, the stronger on average the hybridisation between two complementary strands will be.[26] We have therefore simulated the self-assembly of the same target structures, but with differently chosen patch sequences. These are still chosen randomly, but with an appropriate bias towards either GC or AT base pairs.** Because the DNA hybridisation free energy itself depends strongly on the temperature, changing the bond strengths in this way is not equivalent to simply shifting the temperature scale. We show in Fig. 4 some additional free-energy barriers calculated for a system with stronger average interactions. While the basic behaviour remains unchanged, the different temperatures at which nucleation becomes feasible do affect the driving force for growth and thus the likelihood of defects occurring during the process. For example, if we compare the curves corresponding to $T = 352$ K and $T = 344$ K in Fig. 4, the system with weaker bonds has a less negative large-cluster gradient of the free energy as a function of the largest cluster size ($\sim -1.1 k_B T$ per particle compared to $\sim -1.5 k_B T$ per particle) and thus has a weaker driving force for growth, even though the nucleation free-energy barrier is considerably smaller ($23 k_B T$ compared to $29.5 \ k_B T$). Moreover, the system with stronger bonds appears to grow with more defects in a grand-canonical simulation, with typically three or four incorrect particles bonded in the final structure. A judicious choice of DNA sequences can thus significantly affect the probability that high co-ordination number structures in particular can grow in a reasonably error-free manner.

One of the main advantages of the work on DNA bricks has been their modularity, in the sense that a large range of target structures have been

---

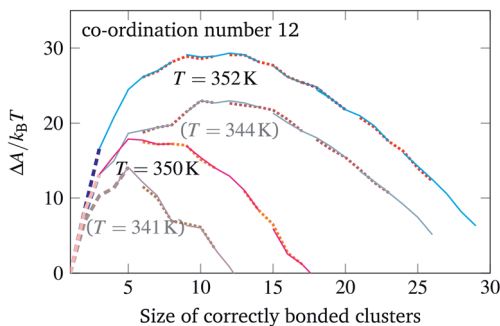** Terminal poly-T sequences are ignored in the GC content calculation.

Fig. 4 The free-energy profile for cluster growth of particles with a co-ordination number of 12 and a GC content of 68%. The free-energy profiles of Fig. 3 are reproduced in a greyed-out hue.

assembled from essentially the same building blocks: the cubic target structures considered so far can be thought of as a 'molecular canvas'.[4,5] It is possible in experiment to construct more intricate structures simply by excluding the undesired bricks from the assembly pot, although in practice, poly-T DNA strands were used at every non-bonded position to minimise undesired interactions. To verify that this modularity continues to be a feature of target structures with a higher co-ordination number, we have run grand-canonical simulations with certain building blocks simply missing. This results in the self-assembly of more complex target structures, exactly as expected. We show two structures that have formed in such conditions in Fig. 5: a 'top hat' style structure and a cube with a cavity. The self-assembly of these target structures from a cubic canvas confirms that the modularity of the building blocks remains a feature in these high-co-ordination number structures.

Moreover, we have run simulations in which the target number density of the undesired building blocks is not set to zero, but rather to a finite but small number. In principle, one would expect that the undesired building blocks need not be completely absent from the reaction mixture, but must simply be vastly outnumbered by the correct building blocks. Our simulations suggest that this is indeed possible, but the fugacities (and hence the solution number densities) of the undesired particles must be set to very low values in order to form the target structure reproducibly. The precise value of the required fugacity depends on the environment of the undesired particles in the underlying canvas structure. For example, for structures with a co-ordination number of 12, if the target structure is a 'top hat' (Fig. 5(b)), most of the undesired particles are outlying particles with relatively few bonds connecting them to the remaining structure. It thus proves possible to form the desired target structure reliably when the undesired particle fugacities are set to approximately 0.5% of the desired particle fugacities of $z_{des} = 2/(78a)^3$ (where $78a$ is the length of the simulation box in lattice units). Larger 'undesired' fugacities result in considerable attachment of the undesired particles over time. However, if the target structure is the central cavity structure of Fig. 5(c), most of the undesired particles are at the centre of the cubic canvas and any undesired bonding that does occur is rather stable; therefore an even lower
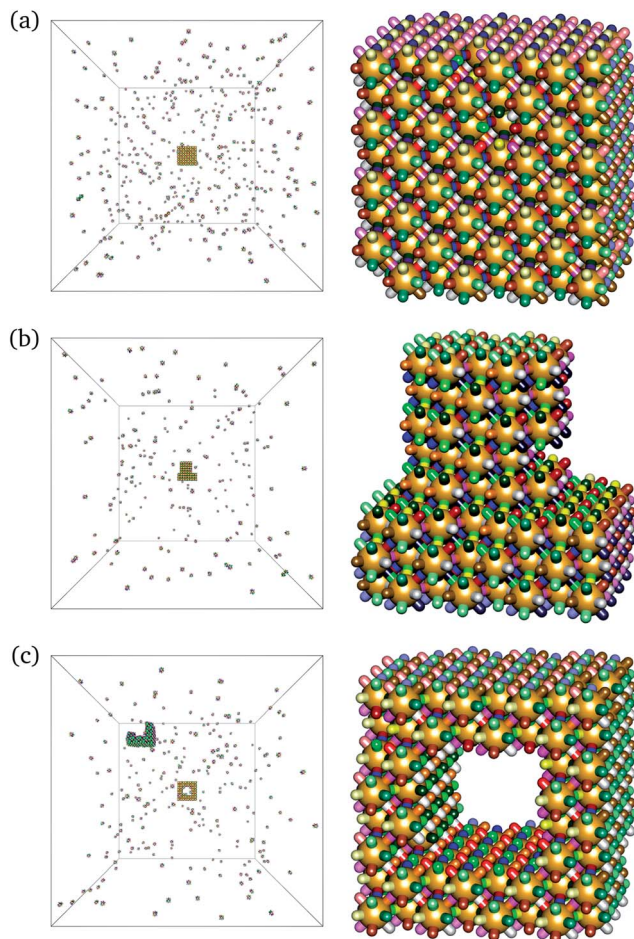
Fig. 5 Snapshots from grand-canonical ensemble simulations of 12-co-ordinate particles with a 68% GC content. $T = 344$ K. The fugacity of all 'desired' particle types is set to $z_{des} = 2/(78a)^3$, where $78a$ is the length of the simulation box in lattice units. All simulation snapshots shown here were obtained from the same building blocks, but with the fugacity of particles not part of the 'desired' structure set to zero. In each case, the whole simulation box and a close-up of the largest cluster are shown. (a) Original cubic target structure. (b) 'Top hat' structure. (c) Central cavity structure.

concentration of undesired particles is required in order to be able to self-assemble the target structure robustly.†† Whilst in theory, designed structures can form in a modular way even when the solution concentration of undesired particles is non-zero, if the target structures are not passivated as they are in experiment (with a poly-T sequence assigned to outlying non-bonding portions of the single-stranded DNA), there is always the chance that at least some undesired

†† In addition, such a structure is considerably more difficult to nucleate than the full cube, since the nucleus that forms must be near the edges of the target structure and has, of necessity, fewer bonds and is thus less stable.

particles will attach to the structure, either during growth or once the desired structure is already fully assembled. In this sense, the experimental strategy of passivating the outer surfaces appears to be very important and permits the desired structure to be assembled even in slightly unclean environments.

4 Conclusions

We have shown that, using a simplified computational model for addressable self-assembly, we are able to self-assemble structures with co-ordination numbers as high as 12. This was a somewhat unexpected result, because we had previously predicted that such structures will exhibit free-energy barriers to nucleation very different – and less conducive to self-assembly – from those previously determined for tetrahedrally co-ordinated structures. Our theoretical work suggested that the nucleation barriers would be less jagged in appearance and much more classical in shape. We predicted that this indicated that self-assembly would be considerably more challenging, because the supersaturation required for nucleation free-energy barriers to be surmountable would need to be greater: so great, we hypothesised, that competition from misassembled structures would dominate and it would be impossible for high co-ordination number structures to be assembled spontaneously in high yield.[16]

Indeed, the theoretical predictions we made about the free-energy barrier are borne out in simulations, but the hypothesis that such structures would be impossible to form is not. We have shown that the free-energy barriers do indeed become less jagged, the critical cluster size is considerably more temperature-dependent and it is more difficult to find mild conditions under which error-free self-assembly can occur. However, we have shown that despite this, it is still possible to find conditions under which the nucleation free-energy barrier is large enough that nucleation is rare, but sufficiently small that it can nonetheless sometimes occur, in conditions under which the stable structure lies along the pathway towards the formation of a fully assembled and designed target structure. This is very good news, because it gives us some confidence that higher co-ordination number structures, which are expected to be of considerable interest in nanotechnology, may indeed be possible to assemble using only a simple protocol.

We have also shown that the design process is modular in much the same way as it is for DNA bricks and that the designed structures self-assemble reproducibly in computer simulations. However, it is necessary to qualify these successes of the simulation method. The computational model we have used to study these effects is very crude and neglects a number of aspects that are likely to be important in any experimental realisation. Notwithstanding the molecular-level mechanisms of DNA hybridisation that have been coarse-grained away, one particular limitation of the model we have used is that it is a lattice model, which over-constrains the geometry of the growing structures and favours their successful assembly. This geometric constraint may be a significant issue in experimental work, perhaps especially so if DNA multi-arm motifs rather than coated colloidal particles were used in the assembly process, as they are themselves not very stiff, and the resulting poor geometry of the growing cluster may significantly retard the growth process. Such additional geometric considerations may cause difficulties not only during the nucleation stage itself, where the additional loss of

entropy of the monomers required to form a compact structure would likely increase the height of the free-energy barrier, but because of the time involved in the reorganisation of the monomer structure when bonding to the growing clusters, they may also reduce the ratio of the rate of cluster growth relative to cluster diffusion. This may make it more likely for different clusters in the system to meet and interact, frustrating their correct assembly. It would be useful in future work therefore to characterise more fully the effect of the cooling protocol on addressable self-assembly.

These considerations may mean that not all possible experimental approaches to many-component building blocks may result in successful self-assembly, and so experimental success is far from guaranteed. It is likely to be the case that an experimental realisation of such building blocks may involve a significant investment of time, effort and not least money. Nevertheless, since we have shown that high co-ordination number self-assembly is computationally feasible, this indicates that the underlying physics does not preclude such structures from self-assembling: we hope this will help to stimulate experimental efforts to achieve similar complexity.

Acknowledgements

References

1 S. Whitelam and R. L. Jack, The statistical mechanics of dynamic pathways to self-assembly, *Annu. Rev. Phys. Chem.*, 2015, **66**, 143–163, DOI: 10.1146/annurev-physchem-040214-121215.

2 D. Frenkel, Order through entropy, *Nat. Mater.*, 2015, **14**, 9–12, DOI: 10.1038/nmat4178.

3 (*a*) V. Linko and H. Dietz, The enabled state of DNA nanotechnology, *Curr. Opin. Biotechnol.*, 2013, **24**, 555–561, DOI: 10.1016/j.copbio.2013.02.001; (*b*) J.-P. Sobczak and H. Dietz, DNA–nanoparticle crystals: exploiting shape complementarity, *Nat. Mater.*, 2015, **14**, 752–754, DOI: 10.1038/nmat4372.

4 Y. Ke, L. L. Ong, W. M. Shih and P. Yin, Three-dimensional structures self-assembled from DNA bricks, *Science*, 2012, **338**, 1177–1183, DOI: 10.1126/science.1227268.

5 B. Wei, M. Dai and P. Yin, Complex shapes self-assembled from single-stranded DNA tiles, *Nature*, 2012, **485**, 623–626, DOI: 10.1038/nature11075.

6 Z. Zhang, J. Song, F. Besenbacher, M. Dong and K. V. Gothelf, Self-assembly of DNA origami and single-stranded tile structures at room temperature, *Angew. Chem., Int. Ed.*, 2013, **52**, 9219–9223, DOI: 10.1002/anie.201303611.

7 B. Wei, M. Dai, C. Myhrvold, Y. Ke, R. Jungmann and P. Yin, Design space for complex DNA structures, *J. Am. Chem. Soc.*, 2013, **135**, 18080–18088, DOI: 10.1021/ja4062294.

8 Y. Ke, L. L. Ong, W. Sun, J. Song, M. Dong, W. M. Shih and P. Yin, DNA brick crystals with prescribed depths, *Nat. Chem.*, 2014, **6**, 994–1002, DOI: 10.1038/nchem.2083.

9 J. D. Halverson and A. V. Tkachenko, DNA-programmed mesoscopic architecture, *Phys. Rev. E: Stat., Nonlinear, Soft Matter Phys.*, 2013, **87**, 062310, DOI: 10.1103/physreve.87.062310.

10 A. Reinhardt and D. Frenkel, Numerical evidence for nucleated self-assembly of DNA brick structures, *Phys. Rev. Lett.*, 2014, **112**, 238103, DOI: 10.1103/physrevlett.112.238103.

11 L. O. Hedges, R. V. Mannige and S. Whitelam, Growth of equilibrium structures built from a large number of distinct component types, *Soft Matter*, 2014, **10**, 6404–6416, DOI: 10.1039/c4sm01021c.

12 J. Zenk and R. Schulman, An assembly funnel makes biomolecular complex assembly efficient, *PLoS One*, 2014, **9**, e111233, DOI: 10.1371/journal.pone.0111233.

13 Z. Zeravcic, V. N. Manoharan and M. P. Brenner, Size limits of self-assembled colloidal structures made using specific interactions, *Proc. Natl. Acad. Sci. U. S. A.*, 2014, **111**, 15918–15923, DOI: 10.1073/pnas.1411765111.

14 A. Murugan, Z. Zeravcic, M. P. Brenner and S. Leibler, Multifarious assembly mixtures: systems allowing retrieval of diverse stored structures, *Proc. Natl. Acad. Sci. U. S. A.*, 2015, **112**, 54–59, DOI: 10.1073/pnas.1413941112.

15 W. M. Jacobs, A. Reinhardt and D. Frenkel, Communication: theoretical prediction of free-energy landscapes for complex self-assembly, *J. Chem. Phys.*, 2015, **142**, 021101, DOI: 10.1063/1.4905670.

16 W. M. Jacobs, A. Reinhardt and D. Frenkel, Rational design of self-assembly pathways for complex multicomponent structures, *Proc. Natl. Acad. Sci. U. S. A.*, 2015, **112**, 6313–6318, DOI: 10.1073/pnas.1502210112.

17 S. Whitelam, Hierarchical assembly may be a way to make large information-rich structures, *Soft Matter*, 2015, **11**, 8225–8235.

18 J. Madge and M. A. Miller, Design strategies for self-assembly of discrete targets, *J. Chem. Phys.*, 2015, **143**, 044905, DOI: 10.1063/1.4927671.

19 Y. Wang, Y. Wang, D. R. Breed, V. N. Manoharan, L. Feng, A. D. Hollingsworth, M. Weck and D. J. Pine, Colloids with valence and specific directional bonding, *Nature*, 2012, **491**, 51–55, DOI: 10.1038/nature11564.

20 F. Lu, K. G. Yager, Y. Zhang, H. Xin and O. Gang, Superlattices assembled through shape-induced directional binding, *Nat. Commun.*, 2015, **6**, 6912, DOI: 10.1038/ncomms7912.

21 (*a*) R. Holliday, A mechanism for gene conversion in fungi, *Genet. Res.*, 1964, **5**, 282–304, DOI: 10.1017/s0016672300001233; (*b*) D. M. J. Lilley, Structures of helical junctions in nucleic acids, *Q. Rev. Biophys.*, 2000, **33**, 109–159; (*c*) C. Zhang, M. Su, Y. He, X. Zhao, P. Fang, A. E. Ribbe, W. Jiang and C. Mao, Conformational flexibility facilitates self-assembly of complex DNA nanostructures, *Proc. Natl. Acad. Sci. U. S. A.*, 2008, **105**, 10665–10669, DOI: 10.1073/pnas.0803841105.

22 N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller and E. Teller, Equation of state calculations by fast computing machines, *J. Chem. Phys.*, 1953, **21**, 1087–1092, DOI: 10.1063/1.1699114.

23 (*a*) G. M. Torrie and J. P. Valleau, Nonphysical sampling distributions in Monte Carlo free-energy estimation: umbrella sampling, *J. Comput. Phys.*, 1977, **23**, 187–199, DOI: 10.1016/0021-9991(77)90121-8; (*b*) M. Mezei, Adaptive umbrella sampling: self-consistent determination of the non-Boltzmann bias, *J. Comput. Phys.*, 1987, **68**, 237–248, DOI: 10.1016/0021-9991(87)90054-4.

24 B. Hetényi, K. Bernacki and B. J. Berne, Multiple "time step" Monte Carlo, *J. Chem. Phys.*, 2002, **117**, 8203–8207, DOI: 10.1063/1.1512645.

25 (*a*) S. Whitelam and P. L. Geissler, Avoiding unphysical kinetic traps in Monte Carlo simulations of strongly attractive particles, *J. Chem. Phys.*, 2007, **127**, 154101, DOI: 10.1063/1.2790421; (*b*) S. Whitelam and P. L. Geissler, Erratum: "Avoiding unphysical kinetic traps in Monte Carlo simulations of strongly attractive particles", *J. Chem. Phys.*, 2008, **128**, 219901, DOI: 10.1063/1.2931944.

26 J. SantaLucia Jr and D. Hicks, The thermodynamics of DNA structural motifs, *Annu. Rev. Biophys. Biomol. Struct.*, 2004, **33**, 415–440, DOI: 10.1146/annurev.biophys.32.110601.141800.

27 D. Frenkel and B. Smit, *Understanding molecular simulation: from algorithms to applications*, Elsevier Academic Press, London, 2nd edn, 2002.

28 R. V. Mannige and S. Whitelam, Predicting the outcome of the growth of binary solids far from equilibrium, 2015, e-print: arXiv:1507.06267.