Assembly Behavior of Hard and DNA-Programmable Colloidal Shapes into Complex Crystal Structures

by

Sangmin Lee

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Chemical Engineering)
in the University of Michigan
2021

Doctoral Committee:

Professor Sharon C. Glotzer, Chair
Professor Nicholas A. Kotov
Professor Ronald G. Larson
Associate Professor Xiaoming Mao

Sangmin Lee

lsmin@umich.edu

ORCID iD: 0000-0002-1145-4708

DEDICATION

To my parents for their love and support throughout my life.

And to my wife Yein for her love and support.

ACKNOWLEDGEMENTS

First and foremost, I have to thank my advisor Professor Sharon C. Glotzer. It was really lucky to work with you in your group. You taught me how to do a good science, how to write a paper, how to work with others, how to learn from mistakes and everything about how to be a good scientist. Thank you. I'd also like to thank my committee members: Professor Ronald G. Larson, Nicholas A. Kotov and Xiaoming Mao for taking the time to discuss my work and keep me on the right path.

Next, I should especially thank Professor Michael Engel (University of Erlangen-Nuremberg, Germany), who was a research scientist when I join the group. He guided me on how to do research and Science when I began my study in this group. I also thank Karen Coulter, who have supported everything happening in this group. Without her, it might not be possible to fully focus on my research.

I would like to thank everyone who is or has been a member of the Glotzer Group, especially Yein Lim and Kwanghwi Je, who have worked with me for many interesting projects. I'd also like to thank my external collaborators, Professor Chad Mirkin of Northwestern university and his group members, including Dr. Haixin Lin, Dr. Shunzhi Wang, Dr. Seungkyu Lee, Jingshan Du, Wenjie Zhou, and Professor David Pine of New York university and his group member Dr. Joonsuk Oh.

Last but not least, I would like to thank my parents, who have supported me everything throughout my life. And, to my wife Yein: You are the best present of my life. Thank you for supporting and loving me.

The research in this dissertation was made possible by the generous support of numerous contributions. Various aspects of this work were supported by the Center for Bio-Inspired Energy Science, an Energy Frontier Research Center funded by the U.S. Department of Energy, Office of Science, Basic Energy Sciences under Award # DE-SC0000989. This work used the Extreme Science and Engineering Discovery Environment (XSEDE), which is supported by National Science Foundation grant number ACI-1548562; XSEDE award DMR 140129. This research also used resources of the Oak Ridge Leadership Computing Facility (OLCF), which is a DOE Office of Science User Facility supported under Contract DE-AC05-00OR22725. This research was supported in part through computational resources and services supported by Advanced Research Computing at the University of Michigan, Ann Arbor.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABSTRACT

Compared to atomistic assemblies with angstrom (Å) length scale, colloidal crystals at nanoscale often exhibit emergent and unique electronic, optical, mechanical and magnetic properties. Recent significant efforts in extending obtainable anisotropic shape and developing surface functionalization methods have enabled the design of a colloidal building block that assembles into a variety of structures and phases. However, the particle design space is highly multi-dimensional and is as yet mostly unexplored. Therefore, developing an easy and efficient model to study the assembly behavior of these particles is important.

In this dissertation, using Monte Carlo and molecular dynamics simulations, we investigate the effect of particle shape and DNA functionalization of colloids on their assembly behavior. We first explore the assembly behavior of hard polyhedra, for which the entropy maximization is the sole driving force, with three projects. In the first project, we show two-step crystallization pathways via a metastable fluid-fluid phase transition in hard particle systems. We categorized our two-step crystallization pathways based on the dimension of the prenucleation motifs and discussed possible comparisons to other crystallization pathways. In the second project, we show entropy-driven self-assembly of five new open caged clathrate crystals stabilized by rotating guests. We demonstrate that crystallization of these crystals occurs via entropy compartmentalization, where the entropy of the system is decomposed into low entropy (host particles) and high entropy (guest particles) subsystems. In the third project, we show the formation and stabilization of fivefold and icosahedral twin clusters in a hard truncated tetrahedra system. We demonstrate that

the icosahedral twin cluster can be entropically stabilized within a dense fluid due to a strong fluid-crystal interfacial tension.

Next, we explore the assembly behavior of DNA-programmable colloids, in which the interaction between DNA ligands on the particle surface determines the assembly behavior. This part also covers three projects. In the first project, we show the formation of complex colloidal clathrate crystals using DNA-functionalized gold nano-bipyramids. The shape of bipyramid with 109.5^{o} edge angle directs the particles to form a local tetrahedral network by maximizing DNA-hybridization, resulting in the formation of clathrate cages. In the second project, we introduce a symmetry breaking mechanism of DNA-programmable colloids, using mobile electron equivalents. We show that the spatial distribution of the electron equivalents in colloidal crystals varies depending on the DNA grafting density, resulting in the formation of nine different crystal structures. In the third project, we study the formation of diverse superstructures of Janus colloids functionalized by DNA. We show that colloidal chains of Janus particles form through cooperative assembly, and the formation of colloidal rings is catalyzed by introducing substrate binding through depletion.

This dissertation shows that entropic bonds can be as effective as traditional chemical bonds in producing complex assembly behaviors and crystal structures, and the tunability of colloidal particles via DNA-functionalization allows us to find a rich set of new colloidal crystal applications.

CHAPTER 1

Introduction

Self-assembly is the interaction process by which particles spontaneously form an ordered structure. Unlike the crystallization of atoms or molecules through covalent bonds, the self-assembly of colloidal particles at nanoscale generally occurs through relatively weak interactions (such as hydrogen bonding and van der Waals interaction) or hard particle interactions (excluded volume effect, *i.e.* entropic bonding (*1*–*3*)). The self-assembly of colloids has been controlled in various ways using particle shape (*4*), particle size dispersity (*5*), DNA-programming (*6*) or interface design (*7*). Such efforts have enabled the assembly of nanostructures with numerous morphologies or properties. In this dissertation, I will explore the entropy-driven assembly behavior of hard particles (Chapters 2, 3, 4) and DNA-programmable assembly of colloids (Chapters 5, 6, 7) of a variety of shapes into complex crystal structures.

1.1 Self-assembly of hard particles

Hard particles interact with each other solely by excluded volume effect without energetic interactions. Thermodynamically, they minimize the Helmholtz free energy $F = -TS$ under the constant volume condition and minimize the Gibbs free energy $G = PV - TS$ under constant pressure. The most thermodynamically stable phase at a given volume is, therefore, the phase that maximizes the total entropy of the system. Interestingly, when systems of hard particles are sufficiently dense, an ordered state sometimes has higher entropy than a disordered state. This was

first demonstrated by Lars Onsager in the 1940s by the thin hard rod model (*8*), which shows a phase transition from a dense fluid to a liquid crystalline phase; it has now been well established that entropy maximization can drive a variety of assembly behaviors. Such entropy-driven self-assembly can be effectively predicted by Monte Carlo simulation using stochastic sampling (*9*), which has enabled study of the assembly behavior of shaped hard particles that are not yet obtainable in experiments. The work of Damasceno *et al.* (*4*) is particularly noteworthy. They used hard particle Monte Carlo simulations to study the self-assembly of 145 hard convex polyhedra and showed that simple measures of particle shape and local order in the fluid can be used to predict the assembly of a particle into a liquid crystal, plastic crystal, or crystal. Since the hard particle model is a good approximation for nanoparticles with predominantly repulsive interaction, these simulation studies have provided predictive guidance for building block design and synthesis in experiments.

1.2 Self-assembly of DNA-programmable colloids

Another approach to control the interaction between colloidal particles is to functionalize the surfaces of particles with various surface ligands such as DNA strands (*6*), polymer brushes (*10*), and oleic acid (*11*). In particular, the assembly of colloids through DNA-functionalization has received much attention because it enables access to a variety of colloidal crystals (*6*, *12*) due to the programmability of DNA ligands. Since the nucleobases of DNA strands can be designed with arbitrary sequences, DNA ligands on the particle surface can be designed to have a specific interaction in the desired way. For instance, designing DNA linkers with self-complementary sequences allows particles to form a crystal that can maximize DNA hybridizations (*13*). On the other hand, designing DNA linkers with complementary sequences (*i.e.*, A-type linkers only attract B-type linkers) allows particles to form a crystal structure maximizing A-B contacts without

allowing A-A or B-B contacts becomes more favorable. These approaches have been applied to nanoparticles with various shapes (*14*), and the structural diversity of nanostructures accessible by DNA-programmable assembly has been significantly expanded. Theoretical approaches have developed computer simulation models of DNA-functionalized colloids that successfully reproduce the experimental observations (*15*, *16*). These models generally adopt coarse-grained schemes because a full atomistic description of DNA strands is computationally demanding.

In this dissertation, using Monte Carlo (MC) and molecular dynamics (MD) simulations, we explore the assembly behavior of various systems of two categories: hard particle (Chapters 2, 3, 4) and DNA-programmable colloid (Chapters 5, 6, 7) systems. In Chapter 2, we demonstrate an entropy-driven two-step crystallization process that is mediated by a metastable fluid-fluid phase transition. In this work, we show that crystallization proceeds via a high-density precursor fluid phase with prenucleation motifs in the form of clusters, fibers and layers, and networks, respectively. In Chapter 3, we demonstrate entropy compartmentalization of single component hard particle systems. In this work, we show that crystallization of an open-structure host-guest clathrate crystal occurs via entropy compartmentalization, in which the entropy of the system is decomposed into low entropy (host particles) and high entropy (guest particles) subsystems. Chapter 4, 5 and 6 describe computational studies of self-assembly conducted in collaboration with experimental groups. In Chapter 4, we show entropy-driven formation of fivefold and icosahedral twin clusters. In this work, we show that fivefold and icosahedral twin clusters of hard truncated tetrahedra can be entropically stabilized within a dense fluid due to strong fluid-crystal interfacial tension. In Chapter 5, we study the formation of colloidal clathrate crystals using DNA-functionalized gold nanoparticles. In this work, we demonstrate that a highly complex crystal structure can be obtained by proper shape design of particles and DNA-functionalization. In

Chapter 6, we describe an approach to generating complex and low symmetry colloidal crystals using mobile electron equivalents (EEs). In this work, we show that, under appropriate conditions, the spatial distribution of the EEs breaks the symmetry of isotropic PAEs, akin to the anisotropic distribution of valence electrons or coordination sites around a metal atom. Finally, Chapter 7 demonstrates a simple scalable method to precisely vary the Janus balance over a wide range and selectively functionalize one patch with DNA. In this work, we observe the formation of diverse superstructures of colloidal micelles, chains, or bilayers, depending on the Janus balance.

CHAPTER 2

Entropic Colloidal Crystallization Pathways via Fluid-Fluid Transitions and Multidimensional Prenucleation Motifs

This chapter is reproduced from Ref. (*17*), a publication in *Proceedings of the National Academy of Sciences of the United States of America* authored in 2019 by Sangmin Lee, Erin G. Teich, Michael Engel[*] and Sharon C. Glotzer[*] ([*]corresponding authors).

2.1 Introduction

Entropically stabilized colloidal crystals of hard particles exhibit a range of structures as diverse and complex as traditional crystals (*3*, *4*). In these systems, entropic forces emerge as a consequence of a multiplicity of equivalent microscopic states, producing, effectively, entropic bonds (*1*, *18*). The resulting entropic valence increases with density in both strength and directionality. There is now a preponderance of evidence that entropy is as versatile as chemical bonds in producing structural complexity. However, all crystallization pathways that have been reported for hard particle systems are simple. For example, hard sphere crystallization follows predictions of classical nucleation theory in simulation and experiment (*19*, *20*). By contrast, fluid-to-solid transition scenarios well known for atomistic and molecular systems, including liquid-liquid phase separation (*21–28*), prenucleation cluster formation (*29*, *30*), and particle attachment (*31*), have not been reported for purely entropic systems. In particular, in the absence of depletion

interactions involving hard colloids with small molecule depletants, no fluid-fluid transitions are predicted or reported. Here, we present three hard particle systems that exhibit entropically driven two-step crystallization. Each two-step process features, notably, first the transformation from a homogeneous low-density fluid (LDF) to a LDF phase coexisting, temporarily, with a high-density fluid (HDF) with more pronounced prenucleation order. This fluid-fluid transition is followed by a transformation to the final crystalline structure. Nucleation of the crystal is catalyzed by the LDF-HDF interface (*32*, *33*), and the resultant crystal is an ordered variant of the HDF. We show that, despite the lack of explicit attractive forces, both the LDF-to-HDF transformation and the HDF-to-crystal transformation can be first order phase transitions accompanied by structural, thermodynamic, and kinetic discontinuities.

It is known that the geometry of phases involved in prenucleation depends on the specifics of the multi-step nucleation pathway. For instance, calcium carbonate ($CaCO_3$) solutions favor compact pre-nucleation clusters at the early stage of nucleation (*34*). Amyloid fibrils crystallize via a hierarchical assembly of β-strands and β-sheets (*35*). And in the crystallization of a gas hydrate, water molecules form amorphous frameworks around guest molecules before clathrate nucleation occurs (*36*, *37*). Given the diversity in these examples, it appears natural to categorize multi-stage nucleation pathways according to prenucleation motif dimension. We report here that many of the possibilities suggested by dimension categorization are realized in hard particle – that is, purely entropic – systems (Figure 2.1). Each of the three particle shapes in this study is a polyhedron whose geometry favors certain prenucleation motifs via the maximization of face-to-face contacts. Truncated tetrahedra (TTs) crystallize into a complex cubic crystal via a HDF containing cluster-type motifs. Pentagonal bipyramids (PBPs) crystallize into a layered crystal via a HDF containing fiber-type and layer-type motifs. And triangular bipyramids (TBPs) crystallize

into a clathrate crystal via a HDF containing network-type motifs. All three crystals are highly complex compared to other known hard particle crystals. Interestingly, the clathrate crystal is identical to that reported for TBPs linked together with DNA (*38*); here the crystal is obtained solely with entropy.

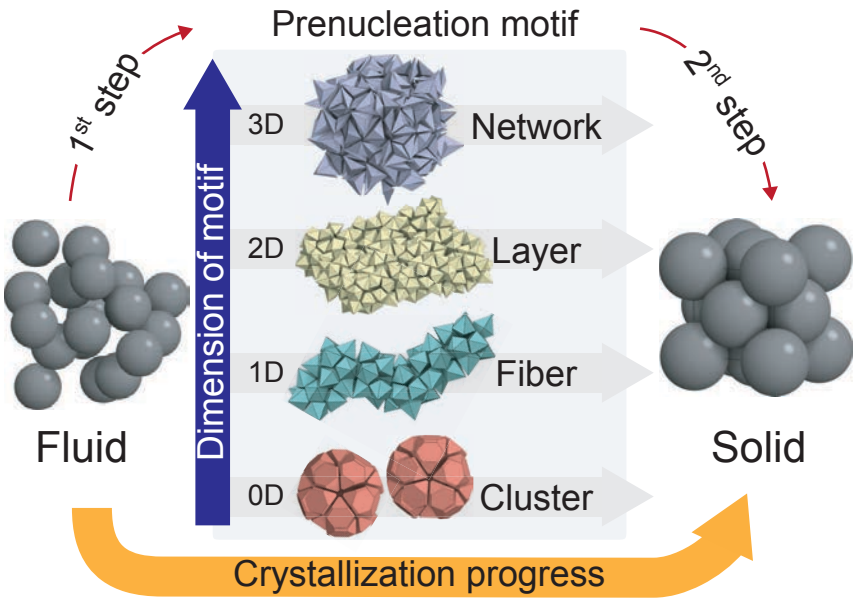


Figure 2.1: Categorization of two-step nucleation pathways. Two-step nucleation pathways are mediated by precursors that can be categorized by the dimension of the prenucleation motif. We distinguish in our simulations of three different systems clusters (0D motif), fibers (1D) and layers (2D), and networks (3D), respectively. The dimension of the motif is a direct consequence of geometrical frustration that builds up as prenucleation motifs grow (*39*).

2.2 Crystallization of hard truncated tetrahedra into *cF*432 crystal

The TT particle is obtained by truncating the vertices and edges of a regular tetrahedron to create adequate space for one TT to fit at the center of an assembled dodecahedron cluster of twenty TTs (*Section 2.7*). Our simulations demonstrate that the self-assembly of hard TTs is mediated by a metastable LDF-to-HDF phase transition and proceeds in two steps. In the first step, an isotropic fluid of hard TTs rapidly (within a few million Monte Carlo sweeps) separates into a

coexistence of LDF and HDF at constant packing density $\phi = 0.615$ (Figure 2.2a). At the beginning of HDF formation, individual clusters emerge from the LDF. As the number of clusters grows, they aggregate by sharing particles (Figure 2.2d). In the HDF, about 27% of the TTs form cluster-type motifs while the rest remain weakly ordered with weaker face-to-face contacts. Clusters and aggregates of clusters continuously change position and morphology via particle exchange with the weakly ordered region in the HDF (Figure 2.9a-2.9d). Clusters do not diffuse within the HDF as rigid units. In the second step, nucleation occurs and crystal growth proceeds isotropically. Unlike clusters in the HDF, clusters in the crystal are periodically arranged at larger separations without shared particles (Figure 2.2e). During crystal growth the number of non-shared clusters rapidly increases, while the number of shared clusters decreases (Figure 2.2i). After 150 million Monte Carlo sweeps, a single crystal coexists with the LDF fluid in equilibrium (Figure 2.2b).

Structural analysis with sites placed at the centroid of each particle reveals that the unit cell of the crystal has Pearson symbol *cF*432, i.e. it contains 432 particles (Figure 2.2c and 2.3) arranged with space group $Fd\bar{3}c$. This is the most complex crystal structure reported in any hard particle system to date. The crystal can be described as a body-centered cubic packing of large truncated octahedral clusters of three successive shells: a TT in the center, a dodecahedron inner shell, and the truncated octahedral outer shell (Figure 2.2f and 2.3d-2.3h). The large clusters tile space and describe the structure completely. The truncated octahedron is decorated with additional particles centered over its hexagonal faces and can be regarded as the union of a strongly distorted dodecahedron and a distorted icosahedron. If both of these polyhedral shells were regular, they would form a rhombic triacontahedron, which is a typical cluster shape found in metallic compounds, such as Bergman-type icosahedral quasicrystals and their approximants (*40*). The

crystal reported here is equivalent in its structure to that of the clathrate-type compound *cF*432-$Te_{7+x}Si_{20-x}$ (*41*).

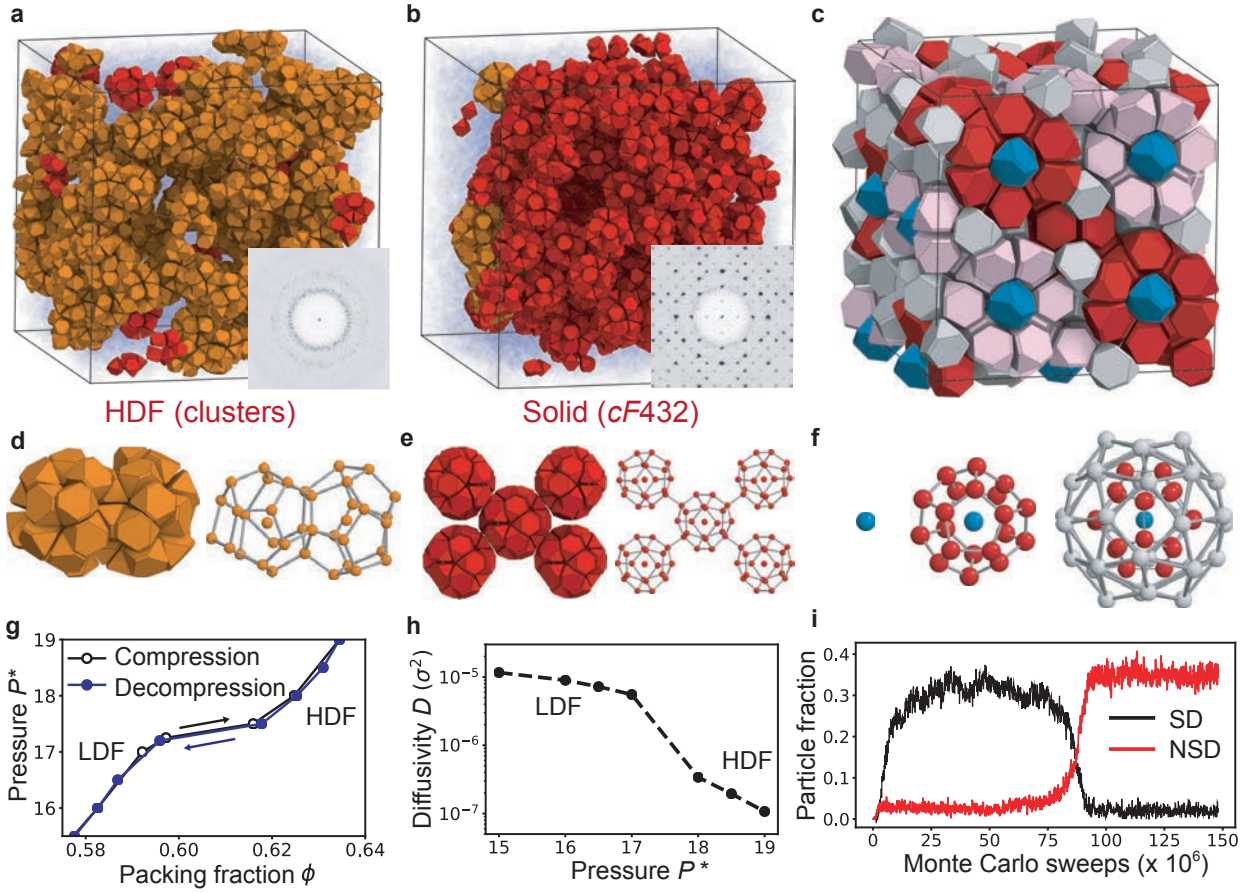


Figure 2.2: Crystallization of *cF*432 via cluster formation. Hard truncated tetrahedra (TTs) self-assemble into the *cF*432 crystal at constant packing density $\phi = 0.615$. **a**, In the early stage of crystallization, hard TTs separate into a low-density fluid (LDF) and a high-density fluid (HDF). The HDF consists of cluster aggregates that share particles (orange in **a** and **d**) and individual clusters that do not share particles (red in **a**). **b**, The crystal consists of clusters that are periodically arranged without sharing particles (red in **b** and **e**). **c**, A unit cell of the *cF*432 crystal consists of a body-centered cubic packing of three-shell truncated octahedron clusters (**f**). **g**, Equation of state of hard TTs. Each data point is obtained by a single isobaric Monte Carlo run initialized in a disordered starting configuration for compression (black line) and initialized in a HDF configuration for decompression (blue line). **h**, Diffusivity of the two fluids near the LDF-to-HDF transition. **i**, Evolution of the number of shared dodecahedron (SD, shown in **d**) motifs and non-shared dodecahedron (NSD, shown in **e**) motifs during the crystallization trajectory. SDs are dominant until 75×10^6 Monte Carlo sweeps, and NSDs rapidly increase as nucleation occurs.

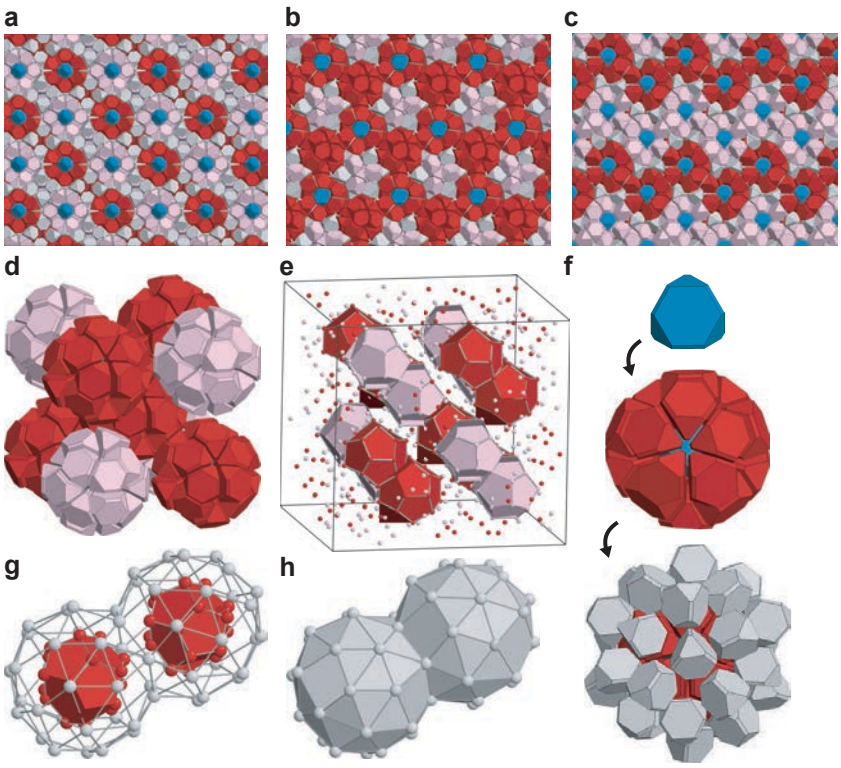


Figure 2.3: Crystal structure of the *cF*432 crystal. Ideal structure model of *cF*432 constructed with hard TTs. Red and pink TTs represent dodecahedron clusters with different orientations. Blue TTs represent TTs at the cluster centers. All other TTs are colored grey. **a,** View along four-fold axis, **b**, three-fold axis and **c**, two-fold axis. **d**, Body-centered cubic packing of dodecahedron clusters extracted from the crystal. **e**, The arrangement of dodecahedron clusters in the crystal. TTs are displayed by their centers and dodecahedron clusters are shown in polyhedron representation. **f**, Three successive shells of the dodecahedron cluster: a single TT (blue), a dodecahedron inner shell (red), and a truncated octahedron outer shell (grey). **g**,**h**, Clusters are arranged on a body-centered cubic lattice by sharing a hexagonal face. The outer shells cover space without gaps.

2.3 Crystallization of hard pentagonal bipyramid into *oF*244 crystal

We next investigate the crystallization pathway of a system of hard PBPs. The aspect ratio of the PBP particle is set to 1.31, which is optimal for twelve PBPs to form a small stellated dodecahedron (*Section 2.7*, Figure 2.8c, 2.8d). We find that crystallization of hard PBPs is again mediated by a metastable LDF-to-HDF transition and follows two steps. In the first step, an isotropic fluid of hard PBPs separates into LDF and HDF phases at constant packing density $\phi =$

0.58 (Figure 2.4a). The HDF possesses fiber-like prenucleation motifs, where small stellated dodecahedra are connected through shared PBPs (Figure 2.4e) but cannot easily continue ordering perpendicular to these connections due to geometrical frustration (*39*). At the beginning of HDF formation, small fibers emerge and then grow into longer, interconnected fibers with various morphologies. The fiber-type motifs continuously change their morphologies and locations by exchanging particles with weakly ordered regions within the HDF (Figure 2.9e-2.9h). In the second step, a decagonal quasicrystal approximant with Pearson symbol *oF*244 and containing ten-fold symmetric columns (Figure 2.5) nucleates at the interface between the HDF and the LDF. Crystal growth proceeds into the LDF, reminiscent of the growth of certain protein crystals (*42*) (Figure 2.4b and Figure 2.10). The approximant crystal is composed of multiple layers, where highly ordered layers and weakly ordered layers alternate (Figure 2.4c). Highly ordered layers are fiber nets with periodic order within the layer (Figure 2.4f), and their growth proceeds outwards. Remarkably, the crystal grows in a zig-zag pattern within a layer (Figure 2.10).

The structure and dynamics of the crystal phase of PBPs is highly anisotropic. Structural differences between the two distinct layers are apparent in the radial distribution function $g(r)$ (Figure 2.4d). $g(r)$ in the weakly ordered layer is mostly similar to that of the LDF, and by contrast develops distinctive peaks in the highly ordered layer. Particle dynamics of the two layers are also distinct as evidenced by the particle density at a cross-section of the crystal averaged over 2×10^8 Monte Carlo sweeps (Figure 2.4g). Particle positions in the highly ordered layer are sharp, while particle positions in the weakly ordered layer are smeared out, indicating the more diffusive nature of their motion.

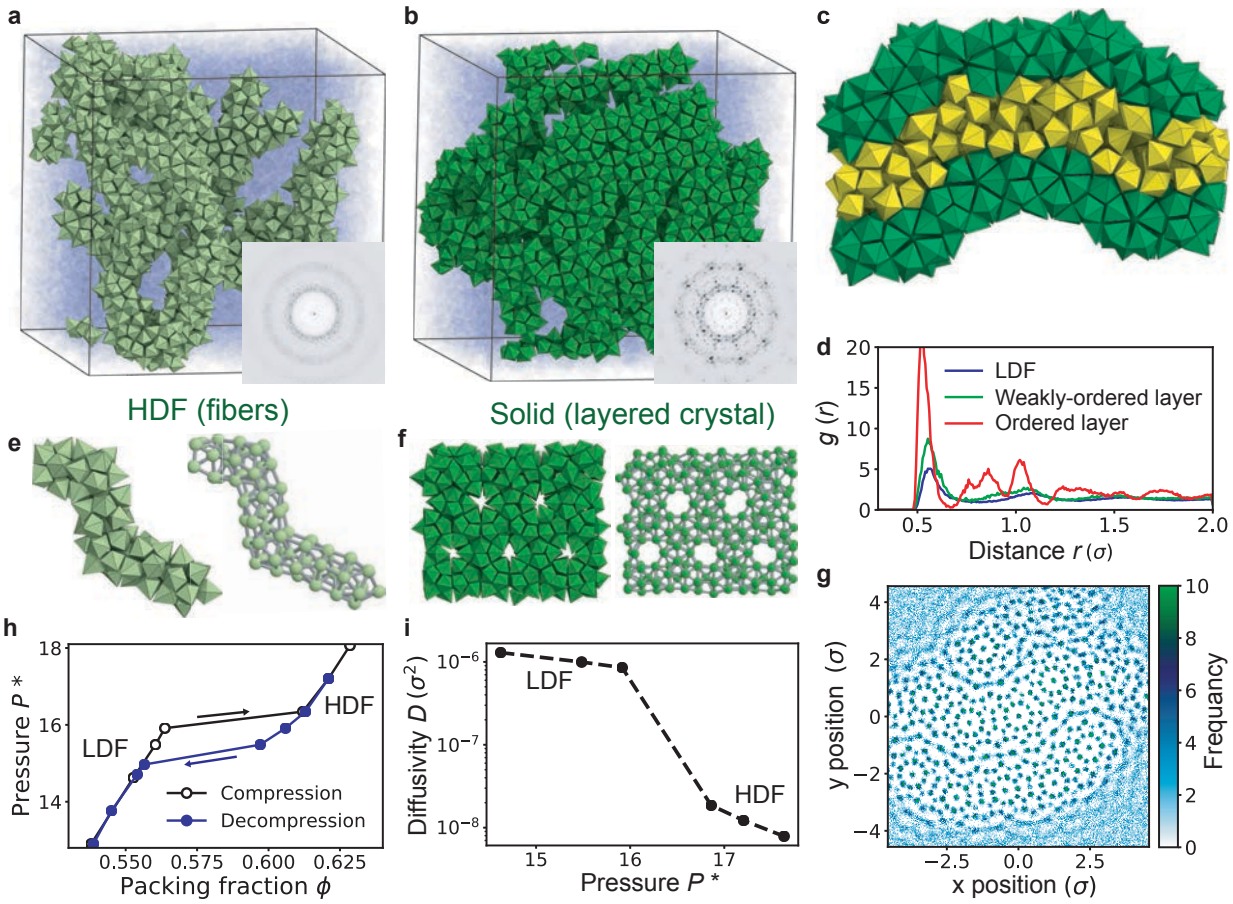


Figure 2.4: Crystallization of a layered crystal (*oF*244) via fiber formation. Hard pentagonal bipyramids (PBPs) self-assemble into a layered crystal (*oF*244) at constant packing density $\phi = 0.580$. **a**, In the early stage of crystallization, hard PBPs separate into a LDF (transparent blue) and a HDF that consists of fiber-like motifs (light green). **b**, The obtained crystal consists of layers of parallel fibers (dark green). **c**, A cross-section of the crystal phase shows that a crystal layer (dark green) and a weakly ordered layer (yellow) alternate within the crystal. **d**, The radial distribution function $g(r)$ of the crystal layers, the weakly ordered layers and the LDF. **e**, A fiber-type prenucleation motif (left) found in the HDF. Centers of PBPs and bonds (grey) are shown on the right. **f**, Part of a crystal layer with PBP representation (left) and its centers (right). **g**, A density plot of particle centers for a cross-section of the crystal averaged over 200×10^6 Monte Carlo sweeps. **h**, Equation of state of hard PBPs. Each data point is obtained by a single isobaric Monte Carlo run initialized in a disordered starting configuration for compression (black line) and initialized in a HDF configuration for decompression (blue line). **i**, Diffusivity of the two fluids near the LDF-to-HDF transition.

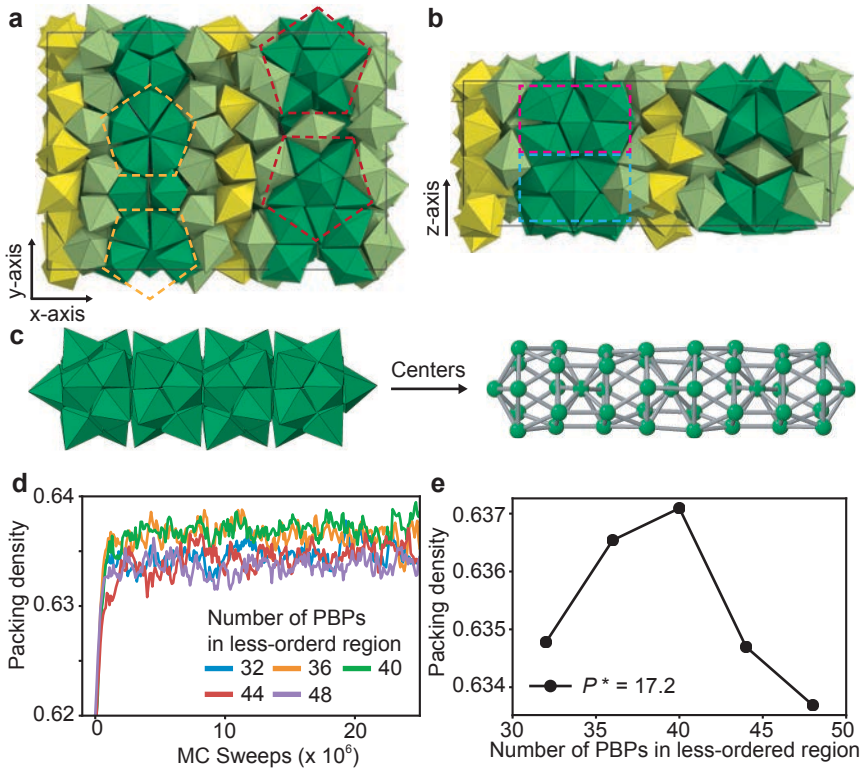


Figure 2.5: Crystal structure of the *oF*244 crystal. Ideal structure model of the *oF*244 crystal constructed with hard PBPs. **a**, View along the z-axis and **b**, y-axis of the orthorhombic unit cell. Dashed lines indicate different orientations and layers of the ten-fold symmetric column (**c)** within the unit cell. Dark green PBPs (120 particles in the unit cell) represent highly ordered layers of the crystal. Yellow PBPs (40 particles) represent weakly ordered layers of the crystal. PBPs existing between the two layers (84 particles) are colored in light green. The configuration of the dark green and the light green PBPs is well defined in the unit cell, while the configuration of yellow PBPs cannot be determined unambiguously due to their diffusive motion. **d**, To determine the most probable number of PBPs in the weakly ordered region, we compared packing density at constant pressure ($P^* = 17.2$) as a function of the number of the weakly ordered PBPs in the unit cell. **e**, The unit cell with 40 weakly ordered PBPs has the highest packing density, suggesting it corresponds to the most stable structure.

2.4 Crystallization of hard triangular bipyramid into clathrate type-1 crystal

Finally, we investigate the crystallization pathway of a system of hard TBPs. The edge angle of the TBP particles is set to 109.5°, which allows groups of four TBPs to arrange into tetramers and eventually a tetrahedral network (*38*) (Figure 2.8e-2.8g). We observe that

crystallization of hard TBPs is similar to the previous examples, except that the HDF of TBPs now has a network-like prenucleation motif. In the first crystallization step, an isotropic fluid of hard TBPs separates into a LDF and a HDF at constant packing density $\phi = 0.50$ (Figure 2.6a and 2.7a). TBPs in the HDF prefer to form tetramers that are interconnected without long-range order (Figure 2.6d). Centers of the tetramers can be mapped onto the nodes of a tetrahedral network. The prenucleation network spans the HDF isotropically and incorporates most of the HDF particles (Figure 2.9i-2.9k). Few particles can exchange between the network and the non-network particles in the HDF (Figure 2.9l), resulting in much slower dynamics in the HDF of TBPs compared to the dynamics in the HDF of TTs or PBPs. In the second, slower step, a clathrate crystal with Pearson symbol *cP*92 nucleates and grows (Figure 2.6b). We identified the clathrate structure as clathrate type I (Figure 2.6e), where the tetramer centers (network nodes) are arranged in two types of clathrate cages, 5^{12} and $5^{12}6^2$ (Figure 2.7b, 2.6c). Here, the notation 5^n6^m indicates that a cage has *n* pentagons and *m* hexagons as faces (*38*).

The snapshots in Figure 2.6c cover the complete crystallization process in the network representation obtained by connecting tetramer centers. A disordered network (HDF) is dominant in the early stage. The formation of the HDF from the LDF occurs spontaneously and, at early times, follows the prediction of the Cahn-Hilliard theory of spinodal decomposition (*43*), namely exponential growth of the first peak in the structure factor at constant wavevector (Figure 2.11). Two-step nucleation and growth in which the first step occurs through spinodal decomposition was recently reported in simulation of an aqueous NaCl solution (*44*). Clathrate nucleation occurs at the LDF-HDF interface and growth proceeds into the LDF. Depletion of LDF particles during clathrate growth is compensated by reduction of the HDF phase in directions distinct from the growth direction of the nucleus. The HDF region assumes a gradually more spherical shape during

clathrate growth suggesting an increase in surface tension during the second step of crystallization, as recently predicted (*45*). Two crystallization stages can be clearly distinguished by following the time evolution of appropriate order parameters (Figure 2.6f, 2.7d, 2.7e and *Section 2.7.6*). In the HDF formation stage, a density order parameter rapidly increases while the clathrate order parameter remains small. Subsequently, the clathrate order parameter increases during clathrate crystal growth and finally reaches equilibrium.

The crystallization behavior of the TBP system mimics aspects of clathrate formation in water, for which an amorphous network has been identified as a crystallization precursor (*36*, *46*). The nucleation behavior of TBPs is also similar to that found in simulation studies of some tetrahedral systems in the sense that a metastable liquid-liquid transition can occur prior to crystallization (*23*, *24*, *26*). In contrast, while the precursor of the TBP system is a high-density phase, precursors of many tetrahedral crystals are low-density phases (*47*). The tetrahedral networks of the molecular systems are open structures, while TBPs fill most of the space when forming a tetrahedral network (Figure 2.8g). In addition, the location of the TBP clathrate nucleus at the LDF-HDF interface and its growth direction into the LDF are different from those observed in molecular models of water (*47*), where nucleation takes place within the low-density liquid. Such details of the nucleation process are affected by the diffusivity of the precursor phase, as discussed for protein crystallization (*22*, *42*).

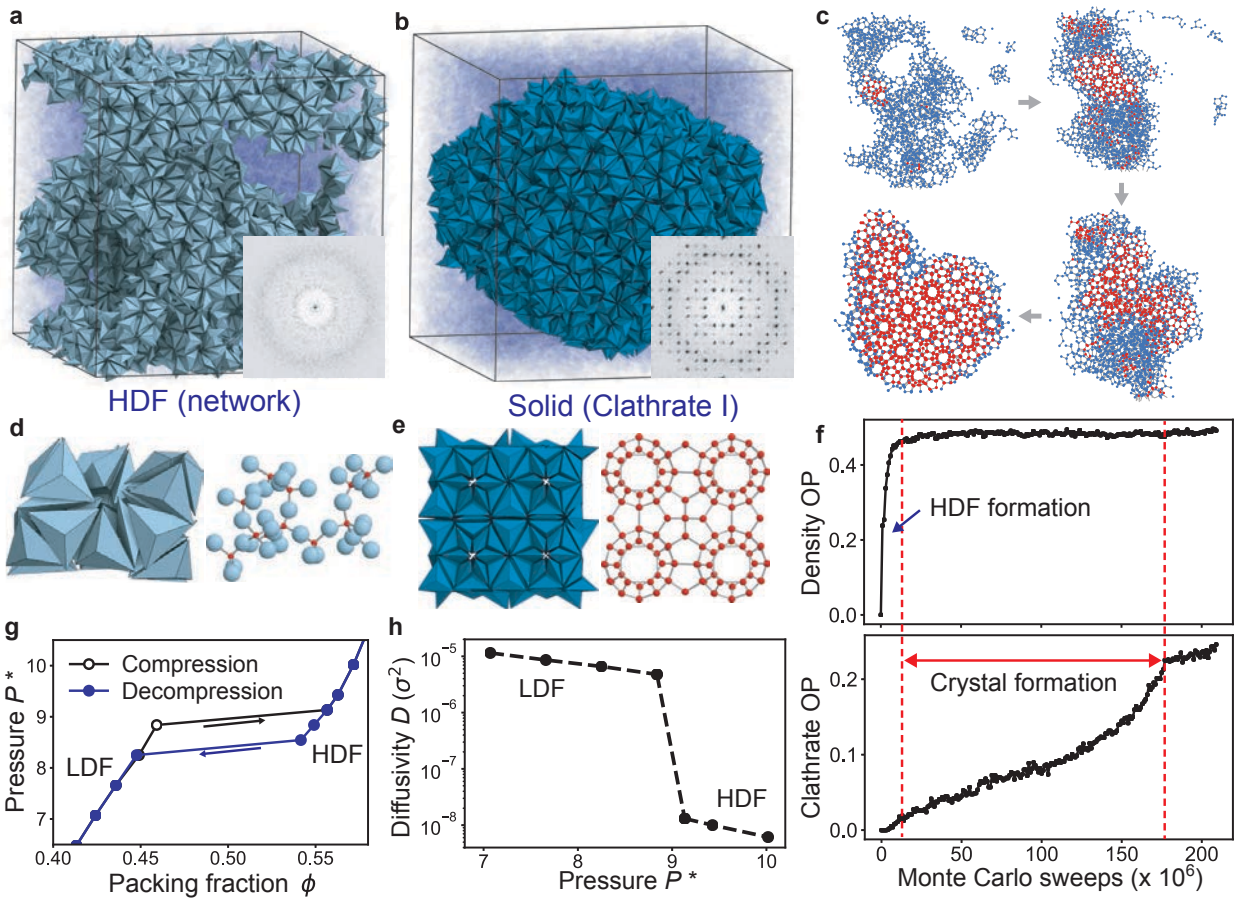


Figure 2.6: Crystallization of clathrate I (cP92) via network formation. Hard triangular bipyramids (TBPs) self-assemble into clathrate I at constant packing density $\phi = 0.500$. a, A snapshot from the early stage of TBP crystallization, where a LDF (transparent blue) and a HDF (light blue) coexist. b, A snapshot of the final stage of TBP crystallization, where the TBP clathrate (dark blue) coexists with the LDF (transparent blue). Diffraction patterns are shown as insets. c, Snapshots of the crystallization process using a network representation that connects TBP tetramer centers. The HDF is colored in blue and the clathrate is colored in red. d, Local motifs found in the HDF (left). Centers of tetramers (red spheres on the right) can be mapped onto nodes of a tetrahedral network. e, Unit cell of the TBP clathrate crystal (left). Centers of tetramers can be mapped onto an atomic clathrate I structure (red spheres on the right). f, Evolution of the density order parameter (top) and a clathrate order parameter (below). Definitions for order parameters are found in Section 2.7.6. g, Equation of state of hard TBPs. Each data point is obtained by a single isobaric Monte Carlo run initialized in a disordered starting configuration for compression (black line) and initialized in a HDF configuration for decompression (blue line). h, Diffusivity of the two fluids near the LDF-to-HDF transition.

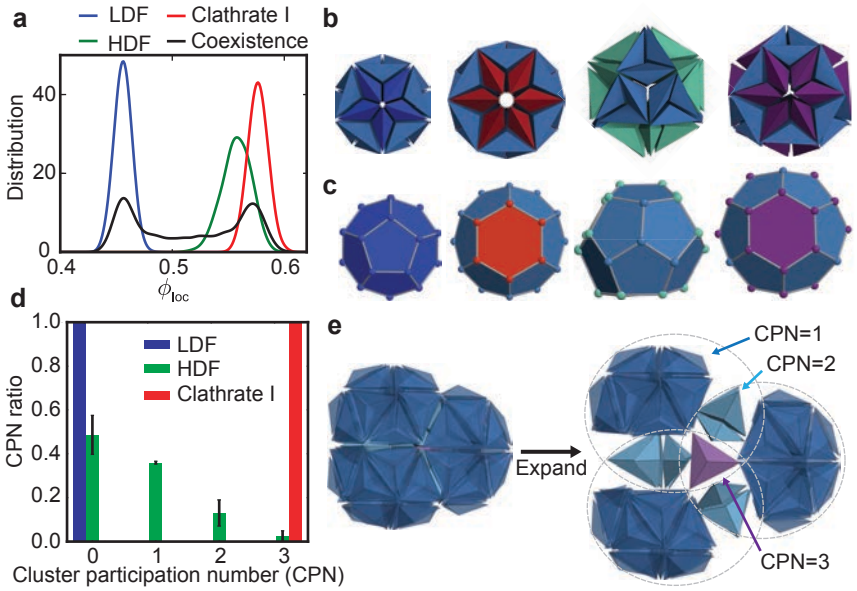


Figure 2.7: Order parameters of the TBP system. **a,** Distribution of the local packing density ϕ_{loc} at four subsequent stages. The LDF ($P^* = 8.84$), HDF ($P^* = 9.43$) and clathrate I ($P^* = 9.43$) are calculated under isobaric conditions, and the coexistence phase is calculated under isochoric conditions ($\phi = 0.50$). **b**, 30, 36, 39 and 42 TBPs form clathrate clusters that can be mapped onto **c**, clathrate cage A (5^{12}), cage B ($5^{12}6^2$), cage C ($5^{12}6^3$) and cage D ($5^{12}6^4$), respectively, by connecting the TBP tetramer centers (Figure 2.8g). **d**, Distribution of the values for the cluster participation number (CPN) in the three observed phases of the TBP system. **e**, The CPN is defined as the number of clathrate clusters a TBP participates in.

2.5 Metastable fluid-fluid phase transition

We find that the LDF-to-HDF transitions of TTs, PBPs and TBPs are accompanied by thermodynamic and dynamic discontinuities. The thermodynamic first-order nature of the three transitions is confirmed by the equations of state calculated under isobaric condition (Figure 2.2g, 2.4h and 2.6g). A small increase of pressure around the phase transition point results in an abrupt density increase of 5% (TT), 8% (PBP) and 16% (TBP) during the formation of the HDF. A dynamic discontinuity is identified via analysis of the diffusivity extracted as the slope of the mean-squared displacement (Figure 2.12). The diffusivity drops by around 1.5 (TT), 2.0 (PBP) and 2.5 (TBP) orders of magnitude across the LDF-to-HDF transition (Figure 2.2h, 2.4i and 2.6h).

Although diffusivity of the HDF is comparably low, it remains non-zero, demonstrating that the HDF is a fluid rather than an amorphous solid in all three cases. We observe the density and diffusivity changes across the fluid-fluid transition are more pronounced if the prenucleation motif in the HDF has a higher dimension. Thus, the density and diffusivity change from the cluster formation within the HDF (TT) are smaller than the density and diffusivity change from the fiber formation within the HDF (PBP), which in turn are smaller than the density and diffusivity change from the network formation within the HDF (TBP). This suggests that particles in a motif with higher dimension are more strongly correlated and as a result produce a denser and less diffusive HDF. The effect of motif dimension on the LDF-to-HDF transition is reminiscent of micellization of amphiphiles. where the character of the phase transition depends in a similar fashion on the size and morphology of aggregates (*48*, *49*).

2.6 Conclusions

Multi-step nucleation and a second liquid phase have been reported in a variety of systems and explained by unique system properties including ion-binding (*29*, *34*, *50*), molecular geometry (*35*, *51*) and tetrahedral bonding (*23*, *26*, *52*). The fact that we observe two-step nucleation mechanisms via a metastable fluid-fluid phase transition in our three model systems of anisotropic hard particles suggests that local order and the geometry of the nearest neighbor shell play the central role in complex phase behavior and kinetics. Whether this geometry is a result of enthalpic bonds or entropic bonds is less important. In this work, we categorized two-step nucleation pathways according to prenucleation motif dimension. It may be fruitful to apply this categorization to analyze the structure of precursors involved in other crystallization processes. If these ideas hold generally, then it may be possible to rationalize the appearance of liquid-liquid phase transitions and multi-step nucleation pathways dominated by connectivity and strain in local

particle clusters – two quantities that contribute to motif dimensionality. Independent of this, our results add to the growing body of evidence that entropic bonds can be as effective as traditional chemical bonds in creating the necessary valence for producing non-trivial crystal structures (*3*, *4*, *18*).

2.7 Method and supporting information

2.7.1 Particle geometry

The TT shape (Figure 2.8a) is a member of the spherical triangle invariant 323 family, with truncation parameters $(a, b, c) = (1.10, 1.00, 1.44)$ according to previous convention (*53*). The PBP used in this work has aspect ratio 1.31 (Figure 2.8c). The TBP has edge angle 109.5° (Figure 2.8e) and aspect ratio 0.35. The aspect ratio of the two bipyramids is defined by the ratio of the height of the bipyramid to the diameter of the circumscribed circle of the base polygon (a regular pentagon for PBP and a regular triangle for TBP). The vertices of the bipyramids are classified as equatorial tips composing the base polygon and polar tips perpendicular to the base polygon. This classification is used when detecting local motifs of the bipyramids.

2.7.2 Cluster-type prenucleation motif of TTs

Twenty TTs form a dodecahedron cluster where an additional TT is located at the center of the cluster (Figure 2.8b). The cluster is detected by the number of nearest neighbors (NNs) of the central particle. At high enough pressure, when the TT system forms cluster motifs, most TTs prefer face-to-face contacts that allow them to have at most four NNs. On the other hand, TTs at the cluster center have twenty NNs that compose a dodecahedron cluster shell (Figure 2.8h). After detecting all clusters in the system, we checked if a cluster shares TTs with other clusters to distinguish between shared clusters (Figure 2.2d) and non-shared clusters (Figure 2.2e). The number of each type of cluster is an order parameter for the crystallization of TTs (Figure 2.2i).

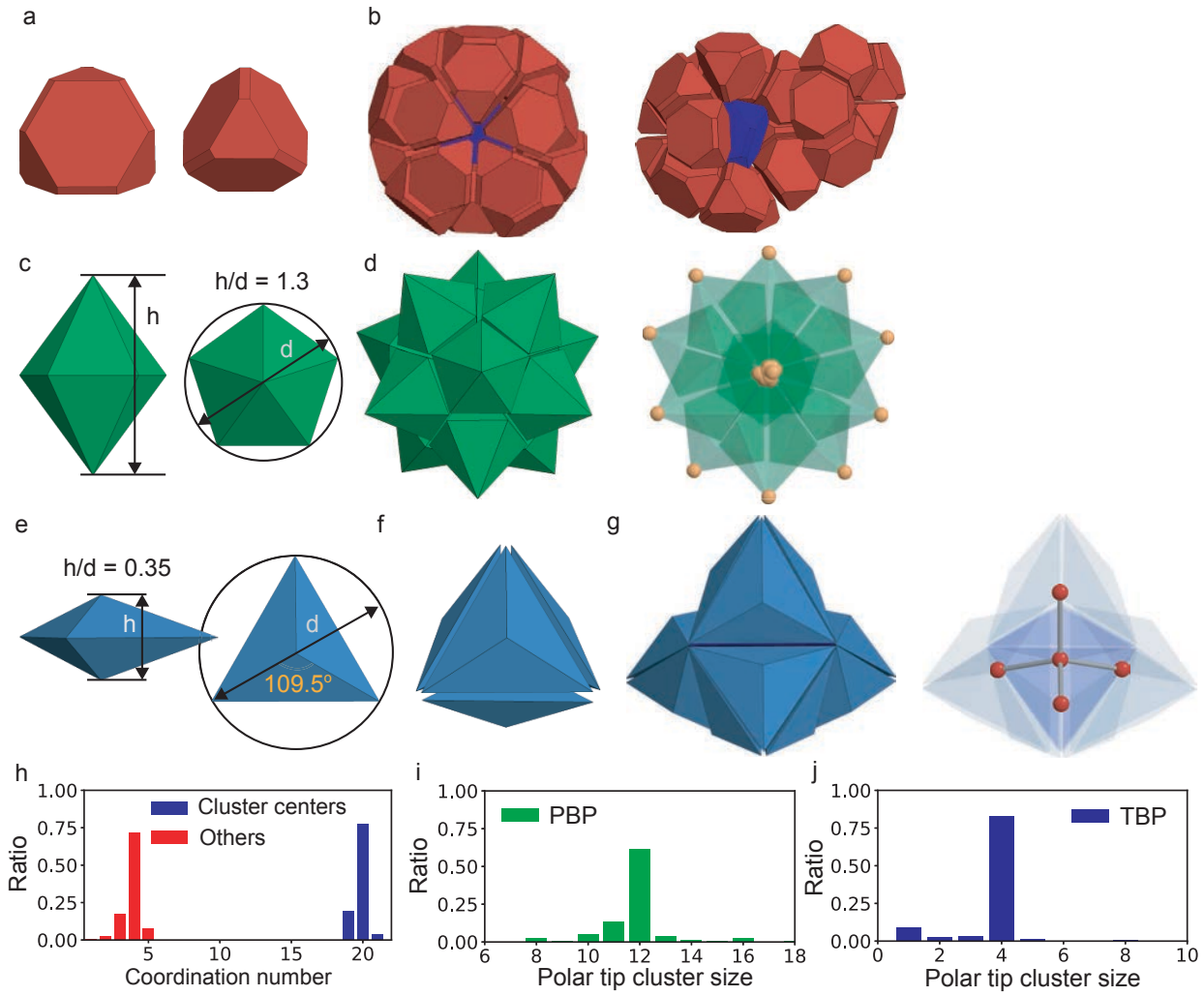


Figure 2.8: Particle geometry and local motifs. a, A single TT with truncation parameters(*53*) $(a, b, c) = (1.10, 1.00, 1.44)$ from side view (left) and top view (right). **b,** A cluster-type motif of $N = 20$ TTs (red) with one additional TT in the center (blue). **c,** A single PBP with aspect ratio 1.3 from side view (left) and top view (right). **d,** A small stellated dodecahedron (SSD) of $N = 12$ PBPs (green), a unit of fiber-like local motifs in the PBP system. One polar tip (yellow sphere) of each PBP points toward the SSD center. **e,** A single TBP with 109.5° edge angle from side view (left) and top view (right). **f,** A tetramer of $N = 4$ TBPs that is a node of the network-type motif in the TBP system. **g,** When five tetramers assemble (left), a unit of the tetrahedral network can be mapped by connecting the locations of polar tip groups (red spheres on the right). **h**, The coordination number distribution of the TT cluster centers (blue shown in b) and the other TTs (red shown in b) calculated from nearest neighbors in the HDF phase ($P^* = 18.0$). **i**, Distribution of

polar tip cluster size of PBPs in the HDF phase ($P^* = 16.35$), which shows a peak at 12. **j**, Distribution of polar tip cluster size of TBPs in the HDF phase ($P^* = 9.13$), which shows a peak at 4.

2.7.3 Fiber-type prenucleation motif of PBPs

Twelve PBPs form a small stellated dodecahedron (SSD) that is a unit of fiber-like motifs (Figure 2.4e and 2.8d). When PBPs align face-to-face, their polar tips come together at the center of the SSD (Figure 2.8d). The distribution of polar tip cluster size in the HDF ($P^* = 16.34$) shows a peak at 12 (Figure 2.8i), indicating that a large number of SSDs exist within the phase. We used this structural characteristic to detect the SSDs in PBP systems. From the arrangement of the SSDs, we identified the fiber-like motifs (Figure 2.4e) and the layered crystal (Figure 2.4f).

2.7.4 Network-type prenucleation motif of TBPs

Four TBPs form a tetramer without gaps by arranging in a face-to-face manner (Figure 2.8f). Each TBP arranged into a tetramer in this way can be part of two tetramers, a prerequisite for the formation of a tetrahedral network (Figure 2.8g). Because the four TBPs of a tetramer share a polar tip in the tetramer center, we can identify tetramers by detecting four-member polar tip clusters. The distribution of TBP polar tip cluster sizes in the HDF ($P^* = 9.13$) exhibits a peak at cluster size 4 (Figure 2.8j). We describe a configuration of many TBPs that are exclusively arranged face-to-face by connecting the centers of tetramers (red spheres as nodes in Figure 2.6d, 2.6e and 2.8g). We call the resulting representation the network representation (Figure 2.6c).

2.7.5 Highly Ordered and weakly ordered regions in the HDF

When the HDF is formed, locally highly ordered prenucleation motifs appear within the HDF and the rest of the HDF remains weakly ordered (Figure 2.9a, 2.9e, 2.9i). The relative number of particles in each of the two HDF regions depends on the dimension of the motif structure (cluster, fiber and network). At a similar degree of supercompression $P^*/P_m^* \sim 1.05$, the HDFs of TTs,

PBPs and TBPs contain 27%, 33% and 91% of particles in the highly ordered region, respectively (Figure 2.9b, 2.9f, 2.9j). Structural distinction between the highly ordered and the weakly ordered region is confirmed by the coordination number distribution calculated from nearest neighbors (Figure 2.9c, 2.9g, 2.9k). The coordination number distributions of TTs, PBPs and TBPs in the highly ordered region show a sharp peak at 4, 10 and 6, respectively, corresponding to the number of faces of each particle. This indicates that the particles in the highly ordered region have greater local order with more pronounced face-to-face contacts than those in the weakly ordered region.

The highly ordered and the weakly ordered region in the HDF continuously exchange particles, and therefore the location where prenucleation motifs appear continuously changes within the HDF. The particle exchange between the two regions is verified by tracking the location of five randomly selected particles during 30×10^6 Monte Carlo sweeps (Figure 2.9d, 2.9h, 2.9l). In all three systems, the selected particles move between the highly ordered and the weakly ordered regions. The TBP system shows a less frequent exchange than the other systems due to the relative sparseness of its weakly ordered region.

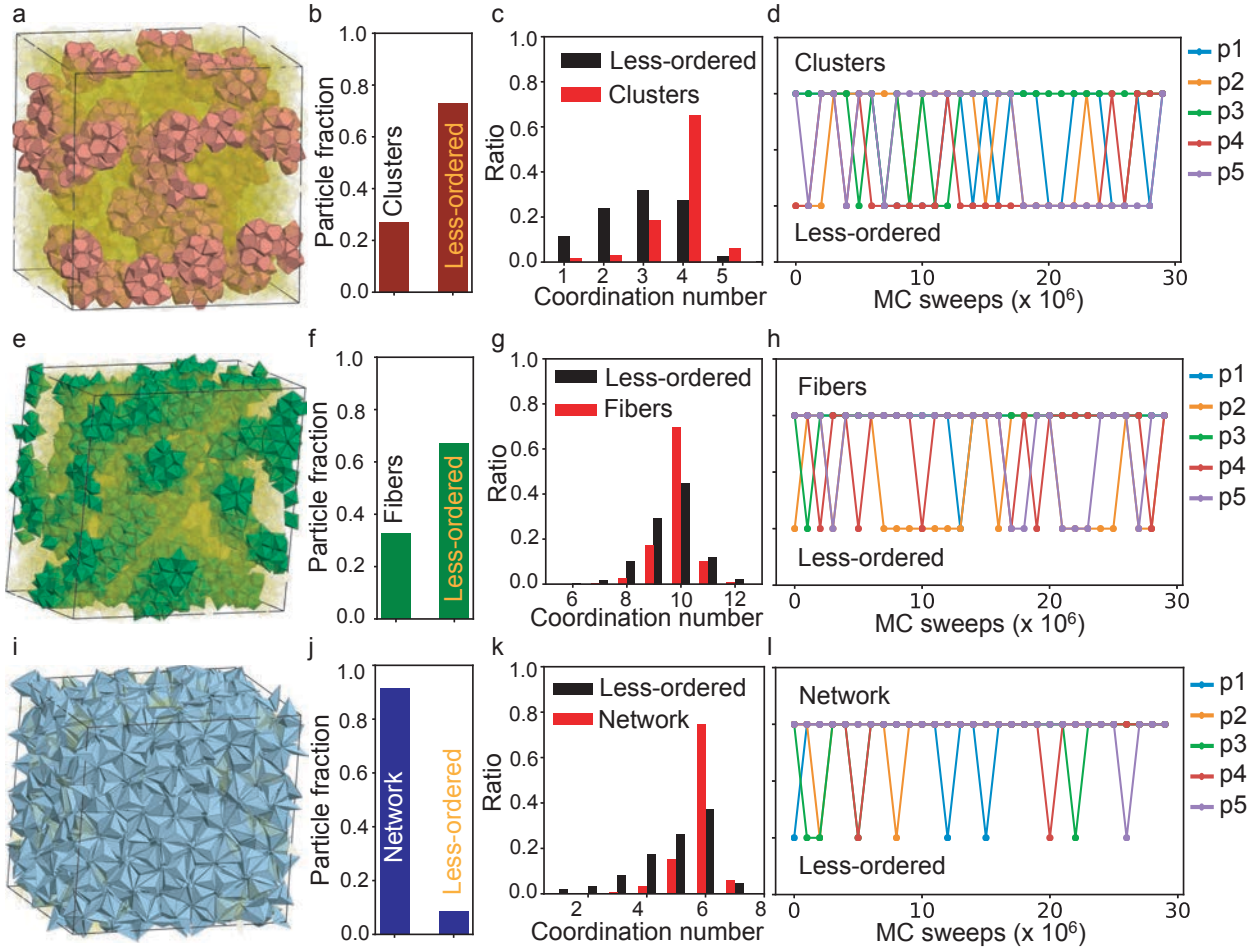


Figure 2.9: Highly ordered and weakly ordered regions of the HDF. **a, e, i,** Snapshots of the HDF prepared under isobaric conditions at $P^* = 18.0$ and $N = 3{,}456$ for TTs (a), $P^* = 16.35$ and $N = 4{,}000$ for PBPs (e), and $P^* = 9.13$ and $N = 4{,}000$ for TBPs (i). Particles in prenucleation motifs are colored pink for cluster-type motifs (a), dark green for fiber-type motifs (e), and light blue for network-type motifs (i). Particles in the weakly ordered region are transparent yellow. **b, f, j**, The fraction of particles that participate in the prenucleation motifs is 27% for TTs (b), 33% for PBPs (f) and 91% for TBPs (j). **c, g, k**, The coordination number distribution of particles in the prenucleation motif (red) shows a peak at 4 for TTs (c), 10 for PBPs (g) and 6 for TBPs (k), while that of the weakly ordered region (black) shows a relatively broad distribution. **d, h, i**, When tracking five randomly selected particles (p1 – p5) in the HDF during 30×10^6 Monte Carlo sweeps, all three systems show particle exchange between the highly ordered region and the weakly ordered region.

2.7.6 Order parameter for TBPs

To identify each stage of TBP crystallization, we monitored two different order parameters: the local packing density (density order parameter shown in Figure 2.6f upper panel) and the cluster participation number (clathrate order parameter shown in Figure 2.6f lower panel). The former is used to identify phase separation between a low-density region (LDF) and a high-density region (HDF) in the first step. The latter is used to identify crystallization of clathrate from the HDF in the second step. For the second step, we could not use the density order parameter because the growth of the clathrate is accompanied by only a minor reorganization of the tetrahedral network and therefore is reflected weakly in local packing density (Figure 2.7a).

For the density order parameter, the local packing density ϕ_{loc} of a TBP is defined as the packing density around the TBP within a sphere of radius 1.5σ, where σ is the longest edge of a TBP. TBPs are identified as belonging to the low-density region (corresponding to the LDF) if $\phi_{\mathrm{loc}} \leq 0.52$ and to the high-density region (HDF or clathrate) if $\phi_{\mathrm{loc}} > 0.52$ (Figure 2.7a). Therefore, the particle fraction in the high-density region is used as the density order parameter.

For the clathrate order parameter, we developed the cluster participation number (CPN) that is defined as the number of clathrate clusters that a given TBP participates in with its three equatorial tips. The clathrate clusters are cage-type clusters that 30, 36, 39 and 42 TBPs can form (Figure 2.7b). They can be mapped onto clathrate cages by connecting polar tips (Figure 2.7c) (*38*). Each clathrate cluster can be identified by counting the number of equatorial tips at the center of the cluster because an equatorial tip of each TBP points toward the center when forming the cluster. Therefore, a TBP that has three equatorial tips can participate in up to three clathrate clusters. This property is used to define the CPN of each TBP (Figure 2.7e). From the distribution of CPN in each phase, we confirmed that TBPs in a clathrate have $\mathrm{CPN} = 3$ whereas TBPs in the HDF have

$\mathrm{CPN} < 3$ (Figure 2.7d). Therefore, the ratio of TBPs with $\mathrm{CPN} = 3$ is used as the clathrate order parameter.

2.7.7 Monte Carlo simulations

Simulations were performed using the hard particle Monte Carlo (HPMC) (*54*) simulation code implemented in the HOOMD-blue simulation package (*9*, *55*). The HOOMD-blue software suite is available at https://codeblue.umich.edu/hoomd-blue. HPMC is implemented on multiple CPUs or GPUs using MPI domain decomposition. The minimum system size with periodic boundary conditions depends on the size of the crystal unit cell. We used particle numbers of $N = 8{,}000$ for TT (Figure 2.2a, 2.2b), $N = 10{,}000$ for PBP (Figure 2.4a, 2.4b) and $N = 20{,}000$ for TBP (Figure 2.6a, 2.6b). Simulations with disordered starting configurations were initialized by placing particles randomly at very low packing density $\phi = Nv_0/V < 0.1\%$ and compressing until the desired thermodynamic condition (ϕ or reduced pressure $P^* = Pv_0/k_BT$) was reached. Here, v_0 and V are the volume of a particle and the volume of the simulation box, respectively. The unit length of all simulations is defined by σ. In this work, the particle volume of a TT is set to σ^3, the height of PBP is 0.9σ and the largest edge length of TBP is σ. After initialization, each run was continued in the isochoric ensemble at constant packing density ϕ or in the isobaric ensemble at constant pressure P^* until equilibration was reached. Total simulation times reached 1.5×10^8 Monte Carlo sweeps for TT, 11.0×10^8 Monte Carlo sweeps for PBP, and 1.5×10^8 Monte Carlo sweeps for TBP. Pressure or density was measured every 10^3 Monte Carlo sweeps to confirm equilibration. Although the duration of crystallization in the present systems is slightly longer than the duration of hard polyhedron crystallization observed in past works (a few million to tens of millions of Monte Carlo sweeps close to coexistence (*4*)), self-assembly is robust, repeatable, and comparably fast given the degree of structural complexity of the crystal structures.

Simulation parameters are listed in the text or figure caption. As an example of a typical compute time, a simulation with 20,000 TBPs at $\phi = 0.5$ run for 2.1×10^8 Monte Carlo sweeps required 20,570 CPU hours on Intel Xeon E5-2680v2 and E5-2680v3 CPUs.

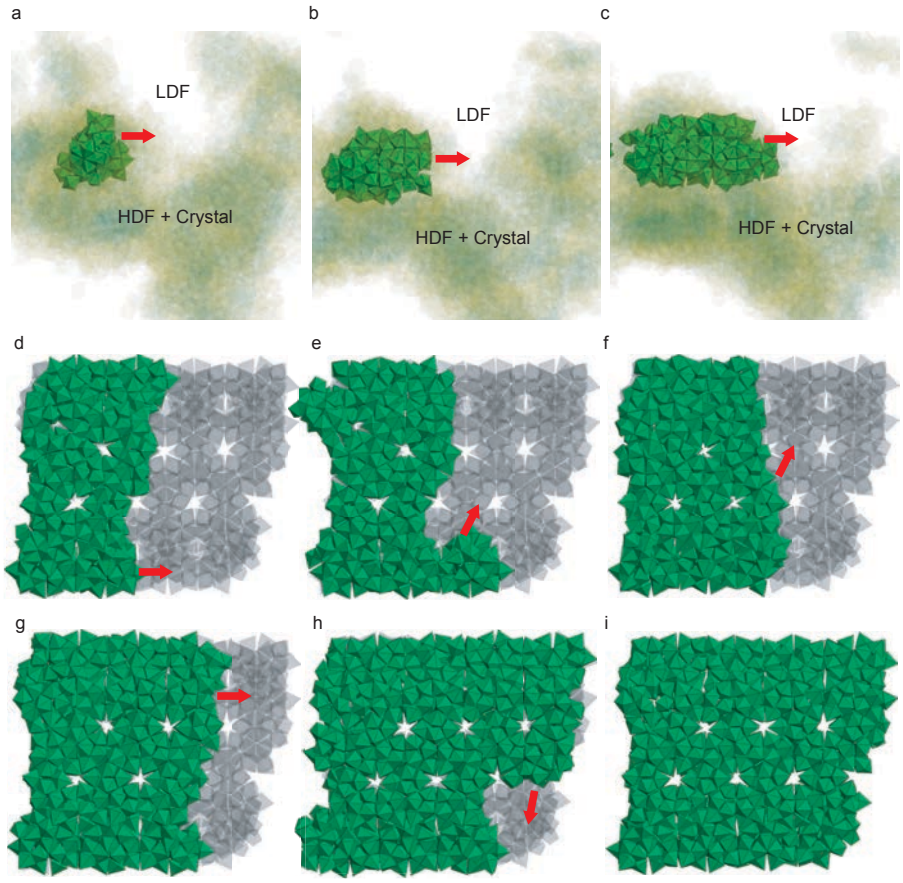


Figure 2.10: Growth of one layer of the PBP crystal. Simulation snapshots (side view of a crystal layer) of the PBP system after **a**, 115×10^6, **b**, 334×10^6 and **c**, 475×10^6 Monte Carlo sweeps. Only one layer of crystal is colored dark green, and particles in the dense phase (HDF and the other layers of crystal) are colored transparent yellow and green. Particles in the LDF are not shown. Red arrows indicate the growth direction. **d-i**, Growth of the crystal layer from top view. Transparent gray particles in the background are the final state of the crystal layer shown for reference. All snapshots correspond to the same system as that in Figure 2.3a, 2,3b.

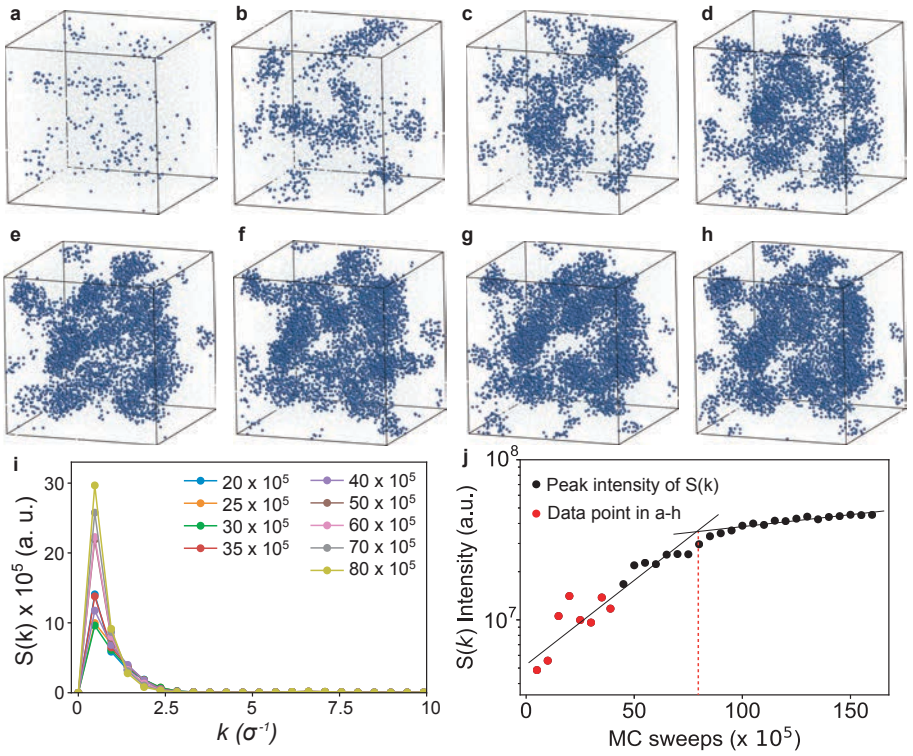


Figure 2.11: HDF precursor formation in the TBP system. Snapshots in the TBP center representation after **a**, 0.5×10^6, **b**, 1.0×10^6, **c**, 1.5×10^6, **d**, 2.0×10^6, **e**, 2.5×10^6, **f**, 3.0×10^6, **g**, 3.5×10^6, and **h**, 4.0×10^6 Monte Carlo sweeps. The HDF forms without delay and gradually coarsens. **i**, Time evolution of the first peak of the structure factor $S(k)$ between 2.0×10^6 and 8.0×10^6 Monte Carlo sweeps. **j**, The peak height increases exponentially at constant wavevector during HDF precursor formation (left side of the dashed red line). This observation agrees with behavior expected for the early stages of spinodal decomposition (*43*).

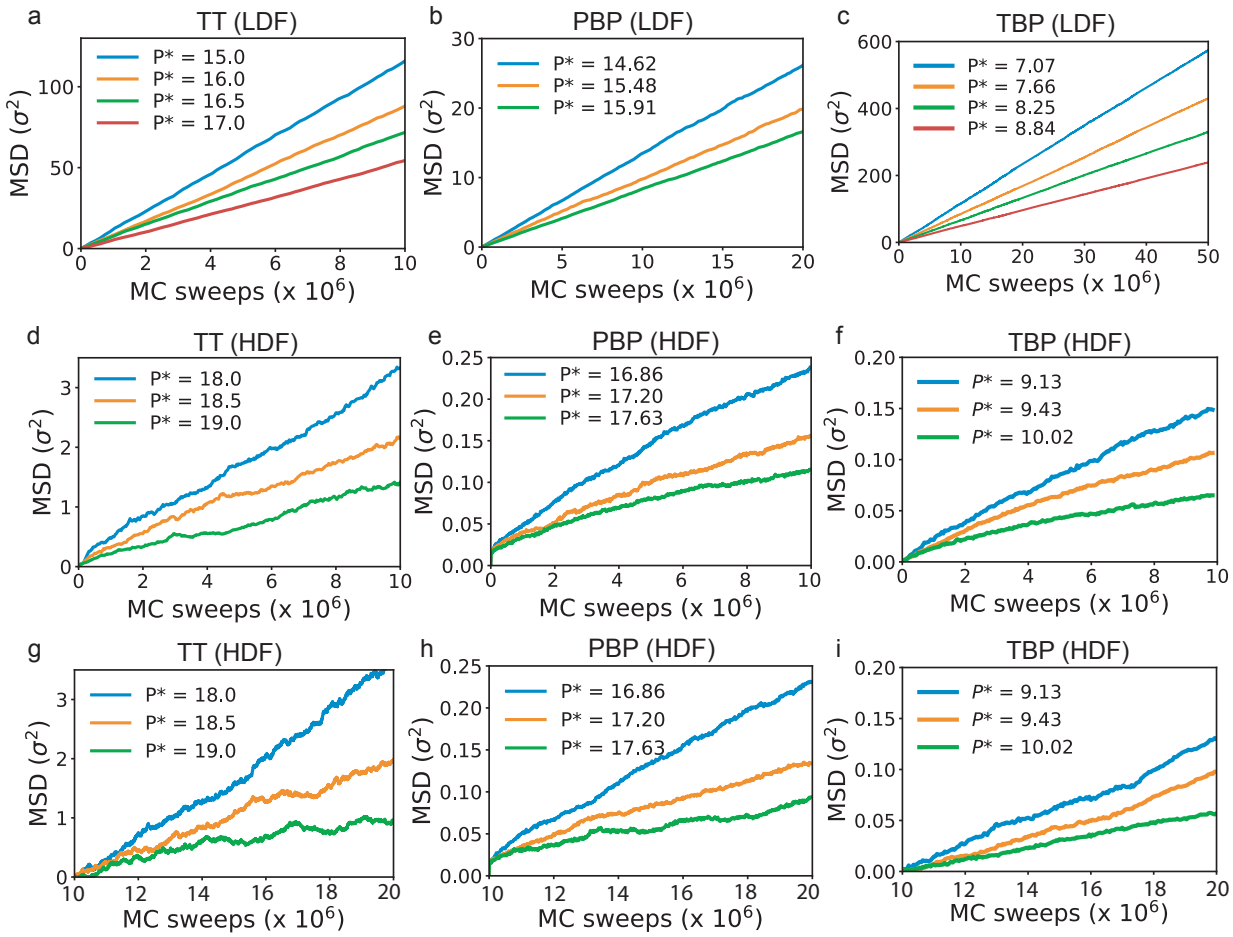


Figure 2.12: Mean-squared displacement (MSD). The MSD of TTs (**a**, **d, g**), PBPs (**b**, **e, h**), and TBPs (**c**, **f, i**) in LDF and HDF at different values of pressure. The MSD of HDF **d-f**, between 0 and 10×10^6 MC sweeps shows similar behavior with the MSD **g-i**, between 10×10^6 and 20×10^6 MC sweeps. The consistency of the MSD for different aging times provide further evidence that the HDF phase is equilibrated. We calculated the MSD of the HDF only when crystallization did not occur during the simulation. For all three systems, both fluids show diffusive motion.

CHAPTER 3

Entropic Compartmentalization in Open-Structure Colloidal Clathrates

This chapter is reproduced from a publication in preparation, authored by Sangmin Lee and Sharon C. Glotzer.

3.1 Introduction

Clathrate is usually found in gas hydrate systems where natural gas guest molecules are trapped inside ice cages (*56*, *57*) or in intermetallic systems where metal ion guests are encapsulated by frameworks of group 14 elements such as Si, Ge and Sn (*58*). The length scale of the cages found in the traditional clathrates has mostly been limited to one or two nanometers; thus, the recent discovery of colloidal clathrates (*59*) opens up the possibility of using the structure at mesoscale where biological or photonic applications are feasible. One of the most interesting properties of the molecular clathrate is the host–guest relation, which is often used for trapping natural gas (*57*) or storing hydrogen (*60*). Thus far, however, the reported colloidal clathrates are all guest-free, thus their potential applications are limited.

Compartmentalization, that is, a system has multiple components taking on different roles, is a common phenomenon in crystal phases. For instance, in various superlattices found in binary nanospheres systems (*61*), it is not surprising that each type of particle occupies a different lattice site that has different local environment. Among single-component systems, the crystal structure

of Uranium (*62*) and high-pressure Ga (*63*) have multiple Wyckoff sites and unique environments, and polymer spheres form the Frank–Kasper σ phase by adopting different packing geometries (*64*). In addition, considering various entropic crystals formed by a single type of hard polyhedral (*4*), it can be speculated that the compartmentalization can be driven by entropy alone. However, the compartmentalization of entropy has not been quantitatively studied and demonstrated yet.

Here, we report the formation of various colloidal clathrates that contain rotating guests via entropy compartmentalization in single-component hard particle systems. We show the self-assembly of five different host-guest colloidal clathrate crystals depending on the amount of truncation of hard triangular bipyramids (TBPs), using Monte Carlo (MC) simulations. The truncation of the TBP creates a cavity in a cage-like motif where a guest can be inserted (Figure 3.1), and the change in the guest/cavity size ratio affects the type of resulting clathrate crystal (Figure 3.2), similar to what has been suggested in gas hydrate systems (*57*, *65*). We show that the free energy of a rotating guest is significantly lower than that of a host particle, which indicates the entropy compartmentalization of the single-component hard particle system (Figure 3.3a-3.3c), where entropy is the only contribution to the free energy (*66*). Although there have been several studies that show multiple Wyckoff sites in an entropic colloidal crystal(*4*), the TBP clathrate is an extreme example where the host and the guest particles behave very differently in a crystal, which shows the entropy can be decomposed into "low entropy" and "high entropy" subsystems. Further, we confirmed the thermodynamic stability of the clathrates at a given amount of truncation by comparing the bulk free energy of various clathrate phases (Figure 3.3d-3.3e). We also show that the rotational motion of the guests is largely affected by the shape property such as the prolateness or the shape symmetry of the guest and the cavity (Figure 3.4). Lastly, we show that the TBP

clathrate cages can be occupied by various guests when the TBPs are mixed with other hard polyhedra (Figure 3.5).

3.2 Particle design for host-guest TBP clathrates

The size ratio of a guest to a clathrate cage is an important factor in determining the type of resultant clathrate structure in gas hydrate systems(*57*). We applied this idea to design the shape of the hard truncated TBP particles. For the regular hard TBPs, the interior of the cage-like motifs is almost fully filled with the body of the building blocks(*17*) so that there is no room for a guest to go into the cage (Figure 3.1b). However, a cavity of the cage can be made by truncating equatorial tips of the TBP (Figure 3.1a, 3.1b and 3.6). The truncation parameter ($0 \leq S \leq 1$) is determined by the distance between the tip of the regular TBP ($S = 0$) and the center of the truncated face (Figure 3.6). We calculated the average radius of the cavity (r_{cavity}) of four clathrate cages (A,B,C and D) from the self-assembly simulations (Table 3.1). The range of S that a cage can contain a guest particle can be estimated by α that is the size ratio between a guest and a cavity ($r_{\text{guest}}/r_{\text{cavity}}$). As increasing S, a monotonic decrease of α was observed for all four cages (Figure 3.1c), and within the entire S range, the relative cavity size between the four cages was D > C > B > A. Interestingly, we observed a critical value of α below which a single TBP guest goes into the cages ($\alpha_c \sim 1.2$, the red dotted line in Figure 3.1c). The S values to reach the α_c of cage A, B, C and D are 0.40, 0.46, 0.48 and 0.58, respectively. This indicates that once the S is given, the type of cage that can include a guest is determined. Then, the system selects the most thermodynamically stable combination of cages to form a certain type of clathrate structure, which will be confirmed by free energy calculation in the later part of this study. From self-assembly simulations, we confirmed that, for $S = 0.42$, the system forms clathrate type-II phase (Clath 2 – A0D1) where guest-free cage A and single-guest cage D are combined in 2:1 number ratio (Figure

3.1d). For $S = 0.52$, the system forms clathrate type-I phase (Clath 1 – A0B1) where guest-free cage A and single-guest cage B are combined in 1:3 number ratio (Figure 3.1e).

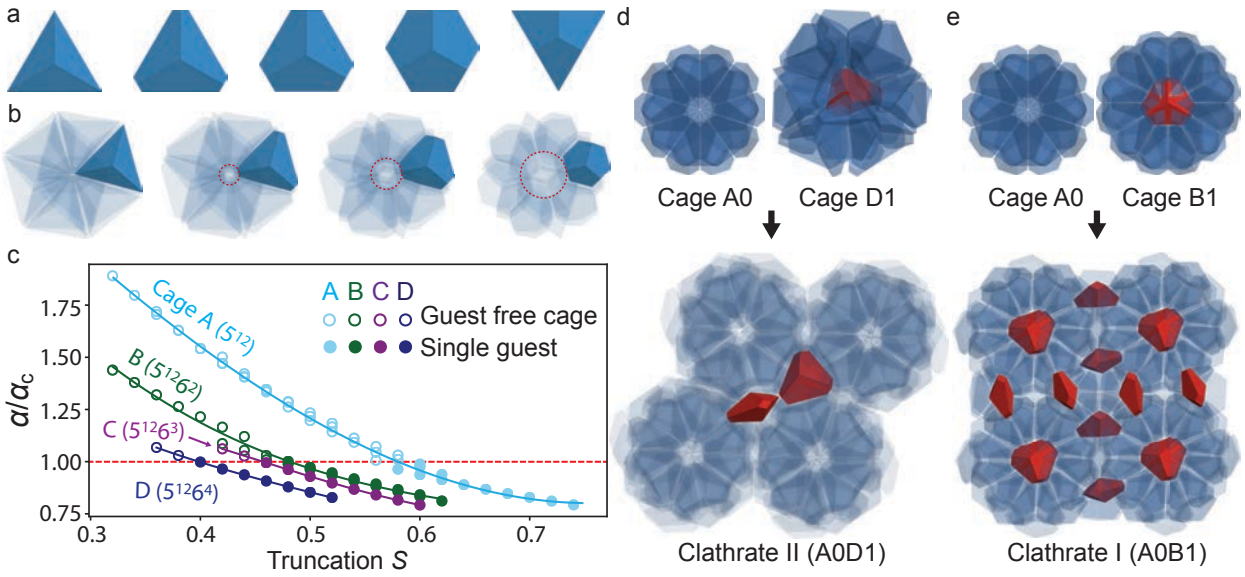


Figure 3.1: Particle design for host-guest colloidal clathrates. a, Triangular bipyramid (TBP) with truncation parameter $S = 0, 0.3, 0.5, 0.7$ and 1.0. b, 30 TBPs form a cage-type clathrate cluster. The truncation creates a cavity (red dotted circle) at the center of the cluster. The size of cavity increases as S increases. c, The change of α/α_c as a function of S for the four clathrate cages at a constant packing density, $\phi = 0.65$. The α is size ratios of a single TBP guest to a cavity ($r_{\text{guest}}/r_{\text{cavity}}$), and α_c is the critical value of α to have the guest. Data points are obtained from self-assembly simulations (solid lines are guide to the eye). Where the red dotted line meets each curve indicates the minimum S for the cage to have a single TBP guest. d, For instance, at $S = 0.42$ where cage A is guest free and cage D has a single guest (top), mixture of the two cages form Clath 2 – A0D1 (bottom). e, Similarly, at $S = 0.52$ where cage A is guest free and cage B has a single guest (top), mixture of the two cages form Clath 1 – A0B1 (bottom).

3.3 Self-assembly simulation

The self-assembly of hard truncated TBPs is studied using hard particle Monte Carlo (HPMC) simulation(*54*) that uses small local moves that mimic Brownian motion (*Section 3.8* and Table 3.1 for simulation condition details). In the range of $0.36 \leq S \leq 0.74$, we observed self-assembly of six different colloidal clathrate crystals (Figure 3.2). The type of clathrate was

confirmed by the cage identification method and the arrangement of the cages (*Section 3.8.3*). At $0.36 \leq S < 0.40$, a guest-free Clath 1 – A0B0 was obtained, which is the same type of clathrate found in the regular hard TBP system(*17*), but with a larger empty space inside the cages. At $0.40 \leq S < 0.46$, we obtained Clath 2 – A0D1 where the guest-free cage A and single-guest cage D are combined (Figure 3.2a-3.2c). In this phase, the cage A is arranged in a rhombohedral form and, in between, the cage D is tetrahedrally arranged. At $0.46 \leq S < 0.49$, we obtained Clath 3 – A0B1C1 where the guest-free cage A, single-guest cage B and single-guest cage C are combined (Figure 3.2f). In this structure, the arrangement of the cage B shows square and triangle tiles (Figure 3.2d, 3.2e), which is mappable to the Frank-Kasper σ phase(*67*). At $0.49 \leq S < 0.55$, we observed Clath 1 – A0B1 where guest-free cage A and single-guest cage B are combined (Figure 3.2i). The arrangement of the cage B shows square tiles, which is a unique feature of clathrate type-1 structure (Figure 3.2g, 3.2h). At $0.55 \leq S < 0.57$, Clath 1 – A1B1 was obtained, where both cage A and B are occupied by a single TBP guest. Lastly, at $0.57 \leq S < 0.71$, a clathrate with single-guest cage A and tetramer-guest cage D was identified (Figure 3.2l), and the cage A are arranged in a rhombohedral form (Figure 3.2j, 3.2k), which is Clath 2 – A1D4. For $S \geq 0.72$, crystallization was not observed during the simulation time. The phase diagram of the clathrate as a function of S observed in self-assembly simulations is shown in Figure 3.2m.

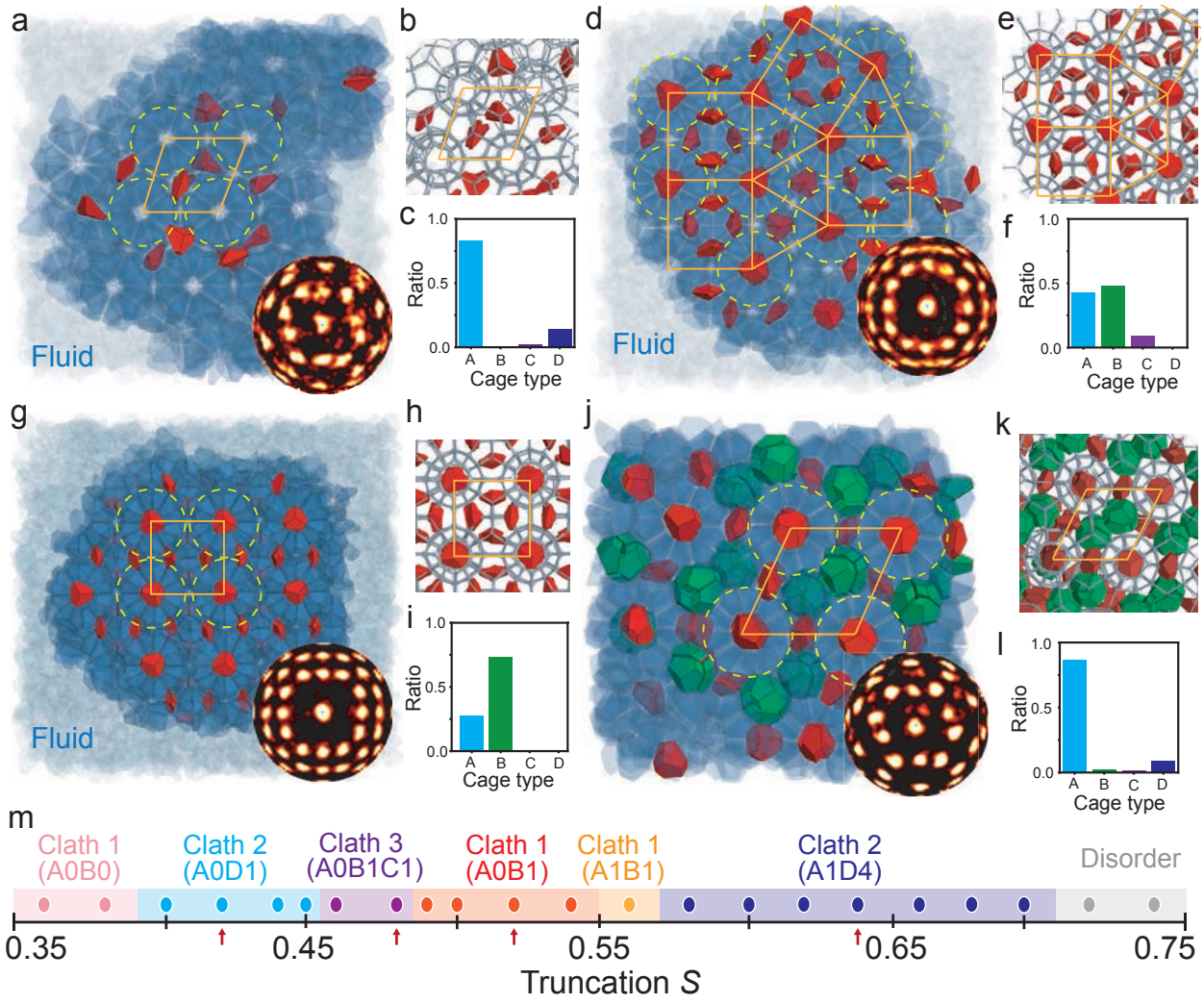


Figure 3.2: Self-assembly simulation results. Hard particle Monte Carlo simulation results of (**a-c**) Clath 2 – A0D1 at $S = 0.42$, (**d-f**) Clath 3 – A0B1C1 at $S = 0.48$, (**g-i**) Clath 1 – A0B1 at $S = 0.52$ and (**j-l**) Clath 2 – A1D4 at $S = 0.64$. Red and green TBPs are guest particles, blue TBPs are host particles. Each clathrate has a unique tiling shape (green dashed and yellow lines) and a bond order diagram (BOD) (bottom left). The BODs are a histogram of the directions of bonds connecting nearest neighbors. (**b,e,h,k**) Network representation of each clathrate (*Section 3.8.4*) for clear visualization of host frameworks. (**c,f,i,l**) Number ratio of cage identified in each clathrate (*Section 3.8.3*). **m**, Phase diagram as a function of S obtained from self-assembly simulation results. Colored dots are data points, and red arrows indicate the four systems shown in this figure.

3.4 Free energy calculation and entropy compartmentalization

Particles in an isotropic fluid phase of a single component system have an almost identical environment. Therefore, entropy, which is the only contribution to the free energy in hard particle systems(*3*, *66*), should be uniformly distributed to every particles in the hard truncated TBP systems before crystallization occurs (Figure 3.3a). On the other hand, the host and guest particles of the clathrate crystals have very different local environments and behavior. For instance, in the Clath 1 – A0B1 crystal ($S = 0.56$), both host and guest particles show a constant translational displacement ($\|\boldsymbol{r} - \boldsymbol{r_0}\|^2$, $\boldsymbol{r}$ is center of mass position) with a comparable amount of fluctuation, during the 10^7 MC sweeps (Figure 3.3b upper panel). This indicates a similar translational movement of the host and the guest. On the other hand, the guest particle shows a much larger oscillation in rotational displacement ($\|\boldsymbol{q} - \boldsymbol{q_0}\|^2$, $\boldsymbol{q}$ is unit quaternion describing orientation) than that of the host particle, indicating a very different rotational motion between host and guest particles (Figure 3.3b lower panel). This implies that the entropy of the crystal is compartmentalized into two different subsystems: host and guest particles.

The entropic compartmentalization is confirmed by free energy calculation. Based on the Frenkel-Ladd method(*68*, *69*), we developed a process of calculating the free energy of the host and guest particles in the crystal, respectively, by controlling the host and guest particles with different paths (Figure 3.8, 3.9 and *Section 3.8.5*). Using the method, we calculated that, for the Clath 1 – A0B1 system ($\phi = 0.60$ and $S = 0.56$), the Helmholtz free energy of a guest particle ($F_{Guest}/N_{Guest} \sim 0.72 k_B T$) is much smaller than that of a host particle ($F_{Host}/N_{Host} \sim 5.65 k_B T$). This is because the rotational motion of the guest particle increases its entropy, and thus decreases its free energy. Therefore, the free energy difference between the host and guest particles may depend on the degree of freedom of the rotational motion of the guest.

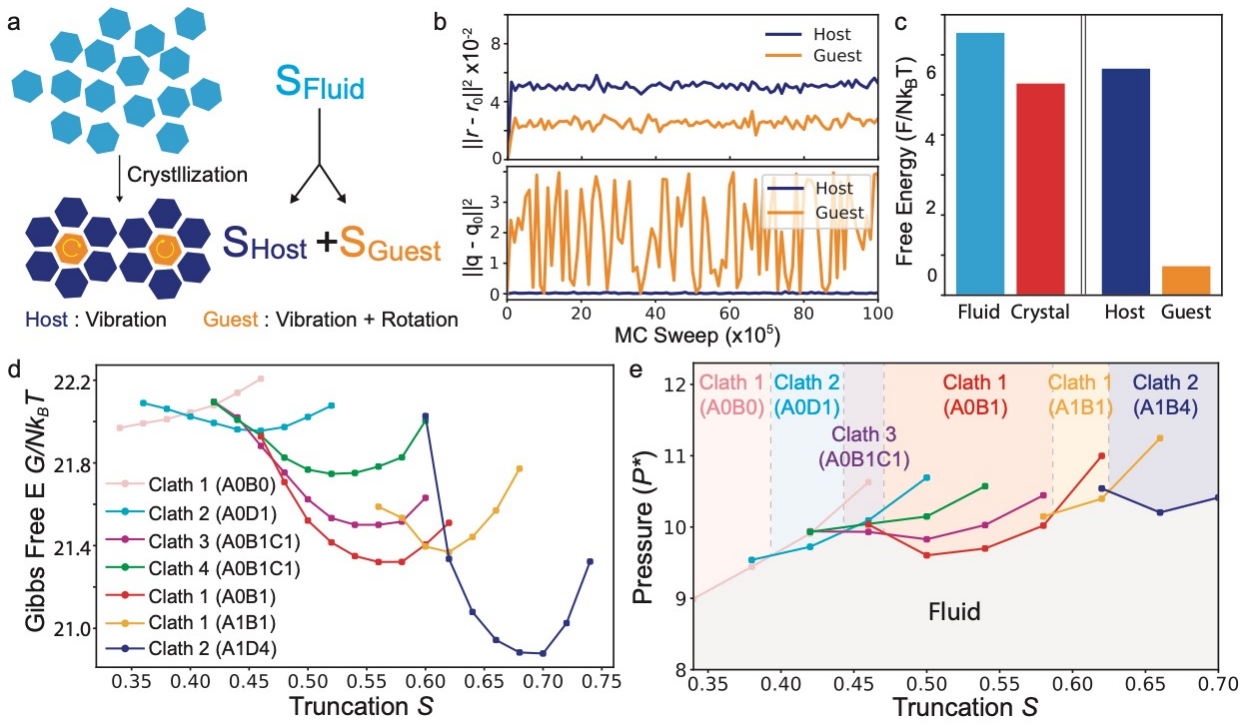


Figure 3.3: Entropy compartmentalization and free energy calculation. **a**, A schematic diagram of entropy compartmentalization of a single component system. **b**, (upper) Translational and (bottom) rotational displacement of the host and the guest TBPs of Clath 1 A0B1 ($S = 0.56$) at $P^* = 13$ for 10^7 MC sweeps. **c**, (left) Helmholtz free energy of $S = 0.56$ system at $\phi = 0.60$ in fluid and crystal (Clath 1 - A0B1) phases. (right) Helmholtz free energy of the host and the guest in the crystal. **d**, Gibbs free energy of clathrates at $P^* = 10.0$ in $0.35 \leq S \leq 0.75$ systems. **e**, Phase diagram of TBP clathrates (P^*vs S). Melting pressure is obtained from where the free energy of clathrate crosses the free energy of fluid as pressure increases (Figure 3.10). Plot legend is the same as **d**.

In addition, we compared the bulk free energy of the clathrate phases to confirm the relative thermodynamic stability between various clathrate types at a given S. We calculated the Gibbs free energy (G) of seven possible types of clathrates in a range of $0.35 \leq S \leq 0.75$ at a constant pressure $(P^* = 11.0)$ (Figure 3.3d and *Section 3.8.6*), using the Frenkel-Ladd method(*68*, *69*). From the calculation, we could accurately measure the range of S where the six clathrate crystals are the most stable (Figure 3.3d). Melting pressure (P_m^*) is another indicator to estimate the stability

of a crystal in hard particle systems. The P_m^* can be determined from where the free-energy of a fluid phase and a crystal phase cross (Figure 3.10a), and the free energy of a fluid can be calculated using thermodynamic integration method (*Section 3.8.6*). We confirmed that the P_m^* obtained from the free-energy calculation and that of the simulations are very similar (Figure 3.10b). Based on the P_m^* of each clathrate phase, a phase diagram (P^* vs S) was contructed in a range of $0.35 \leq S \leq 0.75$ (Figure 3.3e).

3.5 Rotational motion of guest particle

The motion of guest inside the clathrate cage can vary depending on the shape correlation between the guest and the cavity. To identify the rotational motion of guest in the clathrate, we first constructed ideal and periodic bulk crystals and tracked the motion of guests in equilibrium. We assigned a direction vector with a unit length to the guest (Figure 3.4a and 3.4d) and tracked the movement during 3×10^7 MC sweeps. The type of rotational motion can be characterized by mapping accumulated positions of the direction vector of a guest onto a unit sphere or spherical coordinates during the rotation. We identified three different rotational motions of the guest: free-rotation (Figure 3.4b), rotation around an axis (Figure 3.4c) and quantized rotation (Figure 3.4e). For the single TBP guest of cage D in the Clath 2 – A0D1 phase ($S = 0.42$), we assigned a direction vector from the center to a tip of the guest (Figure 3.4a) and observed that the accumulated positions of the direction vectors during rotation show a random distribution (Figure 3.4b), suggesting the free-rotation of the guest. For the single TBP guest of cage B in the Clath 1 – A0B1 phase ($S = 0.52$), the direction vectors were distributed along a continuous line (Figure 3.4c), indicating the guest is rotating around an axis. For the tetramer TBP guest of cage D in Clath 2 – A1D4 phase ($S = 0.64$), we assigned four direction vectors from center to each TBP (Figure 3.4d) and observed that the accumulated positions during rotation form several clusters (Figure

3.4e), indicating the guest has a preferred orientation in the cage. Each cluster has a mixture of the four direction vectors (Figure 3.4e right), indicating that the guest is rotating, not locking, in the cage. We named this type of rotational motion as a quantized rotation.

The shape correlation between the guest and the cavity largely affects the rotational motion of the guest. We characterized the shape of the cavity using the prolateness Pr (*Section 3.8.7*) and confirm that the cavity of cage D is almost spherical ($Pr{\sim}0$), while the cavity of cage B is oblate ($Pr{\sim}-0.02$) (Figure 3.4f). This suggests that the spherical cavity of cage D does not give any shape effect to the oblate single TBP guest (${\sim}0.4$ aspect ratio), resulting in the free-rotation of the guest (Figure 3.4f, right inset). On the other hand, the oblate cavity of cage B allows the oblate guest to be placed along the same circular symmetry axis and rotate around the axis (Figure 3.4f, left inset). The quantized rotation occurs when the guest shares symmetry elements with the cavity. For instance, both the TBP tetramer guest and the cavity in cage D has a tetrahedral symmetry (Figure 3.4g), and this allows the guest has a preferred orientation within the cavity when the orientation of the symmetry matches. Similar rotational behavior was reported for 2D colloidal crystals of polygons when particle symmetry matches the symmetry of its local environment(*70*).

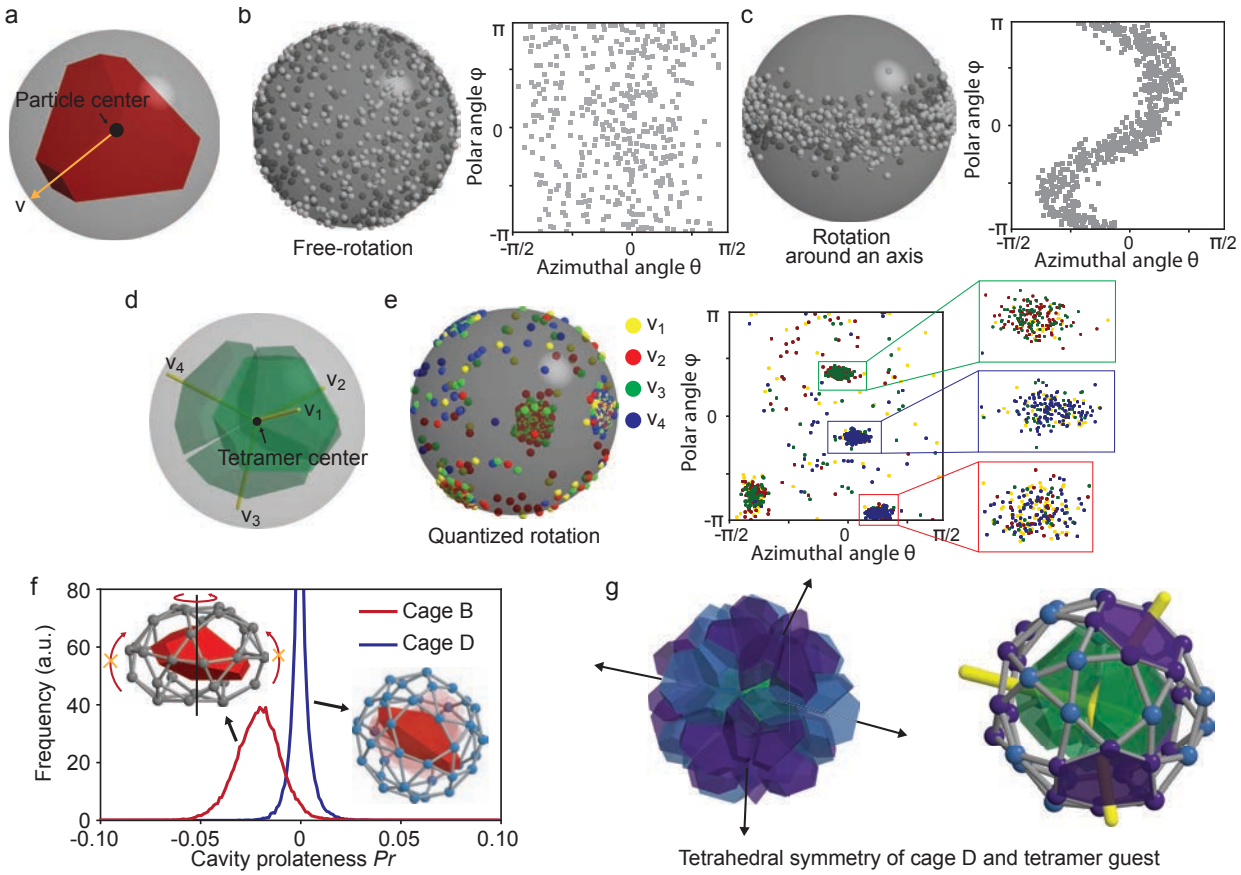


Figure 3.4: Rotational motion of guest in a clathrate cage. **a**, For a single TBP guest, a direction vector (v) with a unit length (σ) from the particle center to the center of a truncated face is chosen to describe the rotational motion of the guest. **b**, Accumulated points of v of a guest particle in cage D of Clath 2 – A0D1 crystal phase ($S = 0.42$) at a constant pressure ($P^* = 10.0$), represented on a unit sphere (left) and spherical coordinates ($r = 1.0\sigma$) (right). Randomly scattered points indicate the free-rotation of the guest. **c**, Accumulated points of v of a guest particle in cage B of Clath 1 – A0B1 crystal phase ($S = 0.52$) at $P^* = 12.5$ on a unit sphere (left) and spherical coordinates ($r = 1.0\sigma$) (right). The accumulated points form a continuous line, indicating a rotation of the guest around an axis. **d**, For a TBP tetramer guest, four direction vectors (v_1, v_2, v_3 and v_4) with a unit length (σ) were assigned from the tetramer center to the center of each TBP. **e**, Accumulated points of the direction vectors of a tetramer guest in cage D of Clath 2 – A1D4 crystal phase ($S = 0.64$) at a constant pressure $P^* = 15.0$ on a unit sphere (left) and spherical coordinates ($r = 1.0\sigma$) (right). The clustering of the points indicates that the guest has a preferred orientation in the cage. Multiple colors of the points in the clusters indicate that the guest is still rotating. **f**, Distribution of prolateness Pr of the cavity of cage B of Clath 1 – A0B1 (same system with **c**) and cage D of Clath 2 – A0D1 (same system with **b**). Negative Pr of cage B indicates an oblate shape of the cavity, and the $Pr \sim 0$

of cage D indicates a spherical shape of the cavity. **g**, Both cage D and the tetramer guest in Clath 2 – A1D4 (same system with **e**) have a tetrahedral symmetry, inducing the quantized rotation of the guest.

3.6 Other polyhedron guest in binary system

The cavity of the TBP clathrate cages can be filled by other types of particles, and this is confirmed by MC simulations of two binary hard particle systems: TBP-tetrahedra and TBP-dodecahedra systems (Figure 3.5 and Table B2). We set the size of the second polyhedron to fit into the cavity of cage A. For the first system, a fluid mixture of TBPs ($S = 0.52$) and tetrahedra (α_A~0.89) crystallize into Clath I at $\phi = 0.55$, and cage A and B are filled by single tetrahedron and TBP guest, respectively (Figure 3.5a-3.5c). For the second system, TBP with $S = 0.64$ forms Clath II at $\phi = 0.58$, and cage A and D are filled by single dodecahedron ($\alpha_A = 0.87$) and TBP tetramer guests, respectively (Figure 3.5d-3.5f). We found that the cage occupying ratio of the binary systems was lower than that of the TBP single component systems. However, we expect that the occupying ratio of guests will increase if a proper number ratio between TBPs and the other particle is given in the initial state.

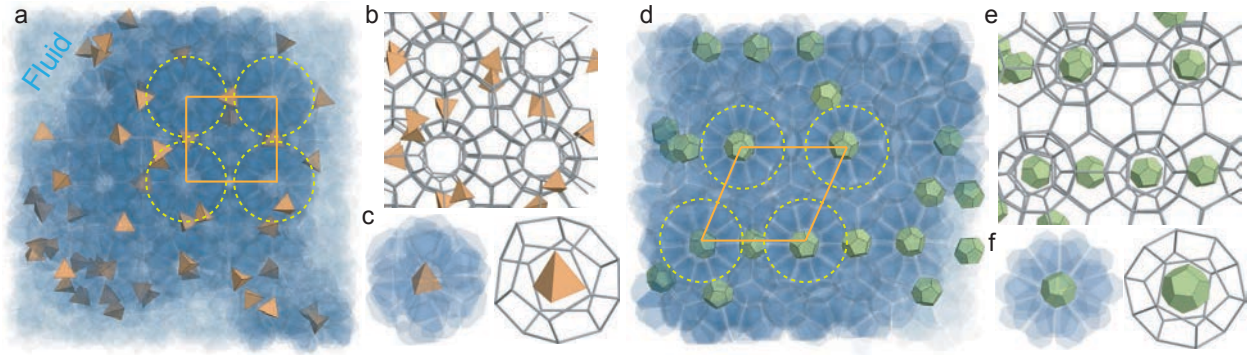


Figure 3.5: TBP clathrate with other polyhedra guests. Hard particle Monte Carlo self-assembly simulation results. **a-c**, Binary mixture of TBPs ($S = 0.52$) and tetrahedra (α_A~0.89) forms Clath 1 – A1B1. Cage A has a tetrahedron guest and cage B has a TBP guest. **d-f**, Binary mixture of TBPs ($S = 0.64$) and dodecahedra ($\alpha_A = 0.87$) forms Clath 2 – A1D4. Cage A has a dodecahedron guest and cage D has tetramer TBP guest.

3.7 Conclusions

This study shows entropic compartmentalization of a single component hard particle system when forming the host-guest clathrate crystals. We observed purely entropy-driven self-assembly of five different host-guest clathrate crystals by simply tuning the amount of truncation of the hard TBPs. The open clathrate structures of the host TBPs are stabilized by the entropy of the guest subsystem that exhibits three different rotational motions: free-rotation, rotating around an axis and quantized rotation. Using free-energy calculation, we demonstrated that the uniform entropy of a dense fluid phase is decomposed into low entropy (the host) and high entropy (the guest) subsystems when crystallization occurs. Our results demonstrate that the entropy is not uniformly distributed in a system and depends on the local environment. This provides another evidence that shows the role of entropy could be more diverse and complex than expected.

3.8 Method and supporting information

3.8.1 Monte Carlo simulation

Simulations were performed by the hard particle Monte Carlo (HPMC) (*54*) implemented in the HOOMD-blue simulation package (*9*, *55*), which is available at https://github.com/glotzerlab/hoomd-blue. The system size for the self-assembly simulations (Figure 2.2) are either $N = 4{,}000$ or $N = 10{,}000$, and all simulations were performed under periodic box condition. The self-assembly simulations were initialized from a dilute isotropic fluid $\phi = Nv_0/V < 0.1\%$ and compressed until the desired thermodynamic condition (ϕ or reduced pressure $P^* = Pv_0/k_BT$) was reached. Here, v_0 and V are the volume of a particle and the volume of the simulation box, respectively. The unit length of simulation is defined by σ, and the particle volume of a TBP is set to $v_0 = 1.0\sigma^3$ regardless of its truncation amount. After initialization, each run was continued in the isochoric ensemble (NVT) at constant packing density ϕ or in the isobaric

ensemble (NPT) at constant pressure P^* until equilibration was reached. For the most systems, crystallization occurs within 1.0×10^8 Monte Carlo sweeps. As suggested in the previous hard colloidal clathrate system (*17*), crystallization was favorable to occur when a solid cluster is surrounded by a fluid (i.e. a coexistence condition between the fluid and the solid). The simulation parameters for the Figure 3.2 are shown in Table 3.1.

3.8.2 Guest to cavity size ratio

The guest to cavity size ratio α is defined by $r_{\text{guest}}/r_{\text{cavity}}$. The r_{guest} is the radius of a circumscribed sphere of the guest. Sometimes, we used α_A which means the α of cage A (i.e. r_{cavity} is the r_{cavity} of cage A). The r_{cavity} is the average radius of the cavity obtained by the distribution of the distances from the center of the cavity to each vertex of the cavity (Figure 3.7). To get the cavity size distribution of four different clathrate cages in $0.36 \leq S \leq 0.74$, we constructed ideal clathrate structures of Clath 1 – A0B0, Clath 2 – A0D1, Clath 3 – A0B1C1, Clath 1 – A0B1 and Clath 2 – A1D4, and equilibrated those structures before measuring the size of their cavity. After measuring the cavity size of the four clathrate cages, we plot the α versus S in Figure 3.1c. Note that, if the shape of the guest and the cavity are both aspherical, a guest could be inserted in a cavity although r_{guest} is slightly larger than r_{cavity} because the r_{cavity} is an averaged value and could be underestimated.

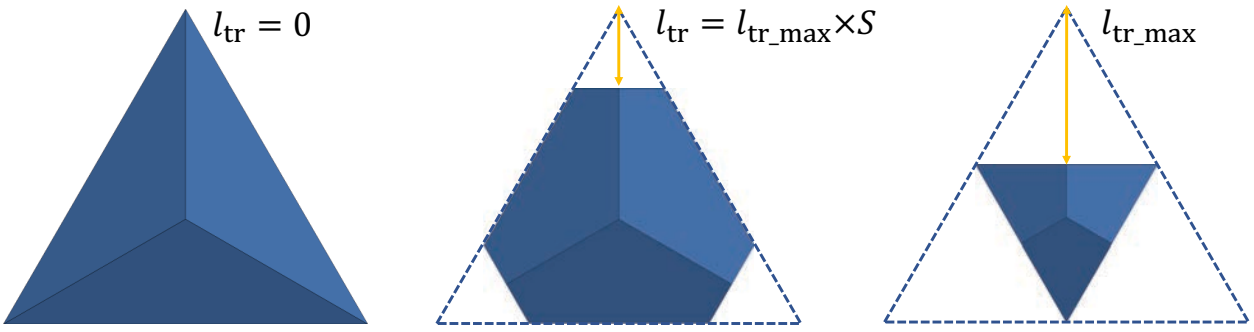


Figure 3.6: TBP truncation. The truncation parameter S is defined as the length ratio of the maximum truncation length (l_{tr_max}). The l_{tr_max} is defined as the maximum distance of the vertex truncation without overlapping the two truncation directions. The regular TBP ($S = 0.0$) vertices are: (1,1,1), (-1,-1,-1), (2,2,-4), (2,-4,2) and (-4,2,2). All the TBPs were rescaled to make the volume of particle be 1.0.

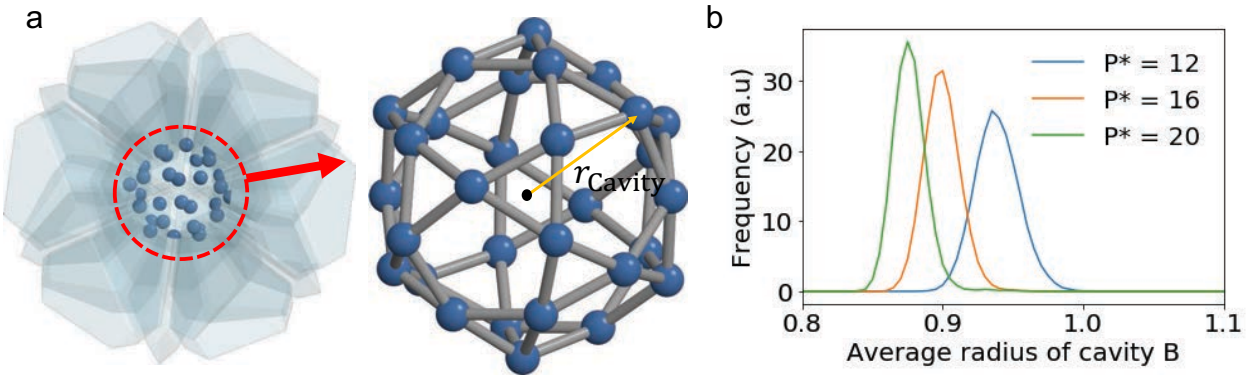


Figure 3.7: Cavity size measurement. **a**, The TBP tips (left) that points toward a cage center form a cavity (right). The radius of the cavity is determined by the average distance between the center of the cavity and the TBP tips. **b**, The distribution of the radius of the cavity B obtained from a bulk crystal phase of $S = 0.52$ system (Clath 1 – A0B1) at three different pressures.

3.8.3 Clathrate identification

The type of clathrate can be identified by the type of cages of the clathrate and the arrangement of the cages (*57*). We classified the cages of the TBP clathrate in four types (*38*) based on the number of TBPs of the cage (A: 30 TBPs, B: 36 TBPs, C: 39 TBPs, D: 42 TBPs) that are mappable to 5^{12}, $5^{12}6^2$, $5^{12}6^3$ and $5^{12}6^4$ cages of molecular clathrates, respectively. The notation 5^n6^m indicates that a cage has *n* pentagons and *m* hexagons as faces. To detect the number

and type of cages in the clathrate crystal, we converted the TPBs into their polar tips because the TBPs point their polar tips toward the cage center. Then, we detected clusters of the polar tips within a cut-off and counted the number of the polar tips of each cluster. The cluster sizes 30, 36, 39 and 42 correspond to the cage A, B, C and D, respectively. Each type of clathrate has its own unique arrangement of cages, which can be used to classify the type of clathrate. For the clathrate type-1, along the four-fold symmetry axis, the cage Bs are arranged in square tiles. For the clathrate type-2, along the two-fold symmetry axis, the cage As are arranged in rhombus tiles. For the clathrate type-3, along the four-fold axis, the cage Bs are arranged in the Frank-Kasper sigma tiles.

3.8.4 Network representation of clathrate

To simplify the visualization of the TBP clathrates, we sometimes represented the host TBPs into a network that connects their tetramer centers (Figure 3.2b, 3.2e, 3.2h, 3.2k, 3.5b and 3.5e). This representation is also shown in Ref. (*38*).

3.8.5 Free-energy calculation for the host and the guest

The free energy of the guest TBP and the host TBP were calculated by setting two paths (A and B) of the Frenkel-Ladd method (*68*, *71*). For each path, we set multiple states (1, 2, 3 …) between the reference state and the real state, which are connected by continuous and reversible harmonic potential (Figure 3.8), thus we can calculate the free energy difference between the states. From the reference state to the A1 state, we fully release the guest particles and remain the host particles same as the reference state. Then, we can get the F_{A1_host} $(= F_{Ref\_host})$ and F_{A1_guest}, respectively. In the same way, we can get the F_{B1_host} and F_{B1_guest} $(= F_{Ref\_guest})$ for the B1 state, by releasing the guest particles only from the reference state. For the A2 state, we partially release the host particle from the A1 state until the host of the A2 shows the same translational and rotational displacement with that of the B1 (Figure 3.9), thus $F_{A2\_host} \sim F_{B1\_host}$. In the same way,

for the B2 state, we partially release the guest particle from the B1 state until the guest of the B2 shows the same translational and rotational displacement with that of the A1, thus $F_{B2\_guest} \sim F_{A1\_guest}$. In this way, we can calculate the free energy of the guest and the host particles, respectively, for each state. And, as we increase the number of states *n*, the n^{th} state of the F_{guest} and the F_{host} will be very close values with the F_{real_guest} and the F_{real_host} that we want to calculate.

3.8.6 Free-energy calculation for bulk phases

The Gibbs free energy per particle G/Nk_BT of the bulk fluid and the bulk clathrate phases are calculated using a combination of the Frenkel-Ladd method (*68*) and thermodynamic integration (*72*). The method is described in detail in Ref. (*71*).

3.8.7 Motion of guest

The motion of guest was identified by accumulated positions of directional vectors of the guest, which are mapped on a unit sphere (Figure 3.4b-e). And their positions were converted to spherical coordinates (r, θ, φ), and plotted on the θ, φ plots where $r = 1$ (Figure 3.4f-i, upper panel). For a single TBP guest, the directional vectors are a set of vectors from the center of the TBP toward the three equatorial tips of the TBP (Figure 3.4b-d, bottom right). For a TBP tetramer guest, the directional vectors are a set of vectors from the center of the tetramer toward the center of each TBPs (Figure 3.4e, bottom right). For the free rotating guest, the scattered pattern of the directional vectors is isotropic. For the guest rotating around an axis, the directional vectors form a continuous line. For the rattling guest, the directional vectors form several clusters.

The mode of the guest motion is affected by the α and the prolateness of the cavity (Pr). The α is defined in the earlier part of this section. The Pr measures the prolateness (or the oblateness) of the cluster shape of a set of points from the elements of their gyration tensor (*73*).

$$S_{mn} = \frac{1}{N}\sum_{i=1}^{N} r_m^{(i)} r_n^{(i)}$$

where $r_m^{(i)}$ is the m^{th} cartesian coordinate of the position vector $r^{(i)}$ of the i^{th} particle. Since the gyration tensor is a symmetric 3×3 matrix, it can be diagonalized.

$$S = \begin{bmatrix} \lambda_x^2 & 0 & 0 \\ 0 & \lambda_y^2 & 0 \\ 0 & 0 & \lambda_z^2 \end{bmatrix}$$

where the axes are chosen to be $\lambda_x^2 \leq \lambda_y^2 \leq \lambda_z^2$. These diagonal elements are used to define the Pr of the cavity as shown in Ref. (*73*):

$$Pr = \frac{\prod_{i=1}^{3}(\lambda_i - \bar{\lambda})}{\bar{\lambda}^3} = 27\frac{\det \hat{\mathbf{S}}}{(\mathrm{Tr}\,\mathbf{S})^3}$$

At the limit of high prolateness, the $Pr = 2$, and at the limit of high oblateness, the $Pr = -0.25$.

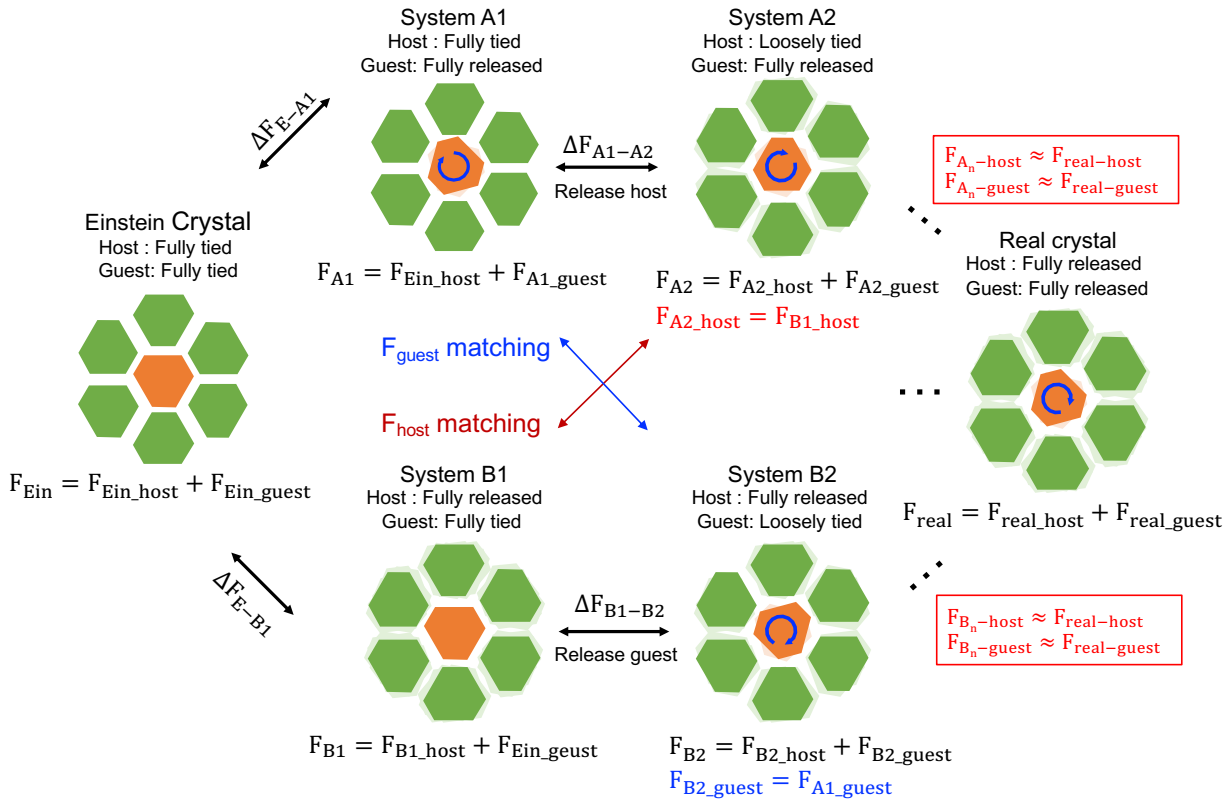


Figure 3.8: A schematic diagram of free energy calculation of the guest and the host in a single component system (*Section 3.8.5* for details).

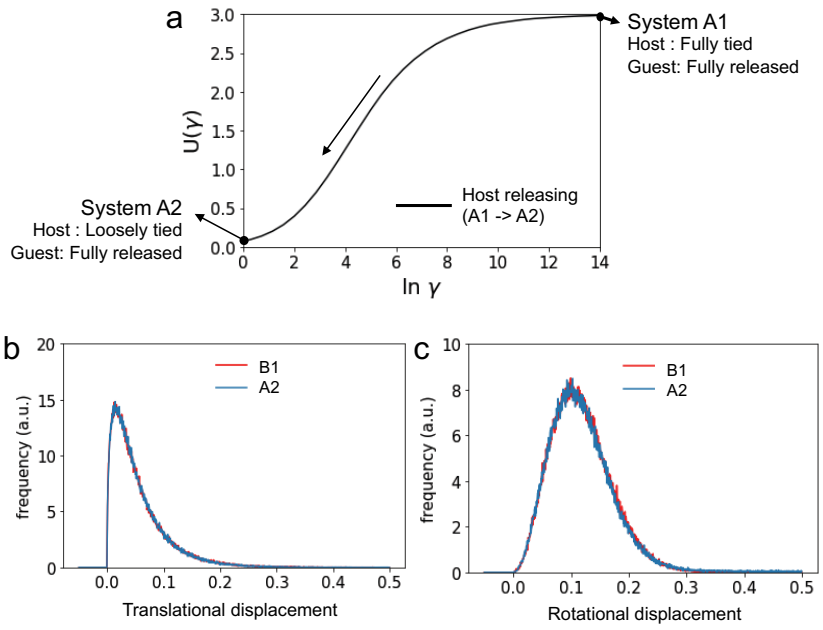


Figure 3.9: Free-energy matching between two states (B1, A2) represented in Figure 3.7. **a**, Potential energy $U(\gamma)$ change of between the A1 and the A2 states of $S = 0.56$ and $\phi = 0.60$. By gradually releasing

the spring constant γ of the harmonic potential tied on the host particles, we keep tracking the translational and the rotational displacements of the host particles. **b,c**, When the translational and the rotational displacements of the host particles of the A2 matches with that of B1, the releasing is stopped. Here, the γ of A2 is 1.0 ($\ln \gamma = 0$).

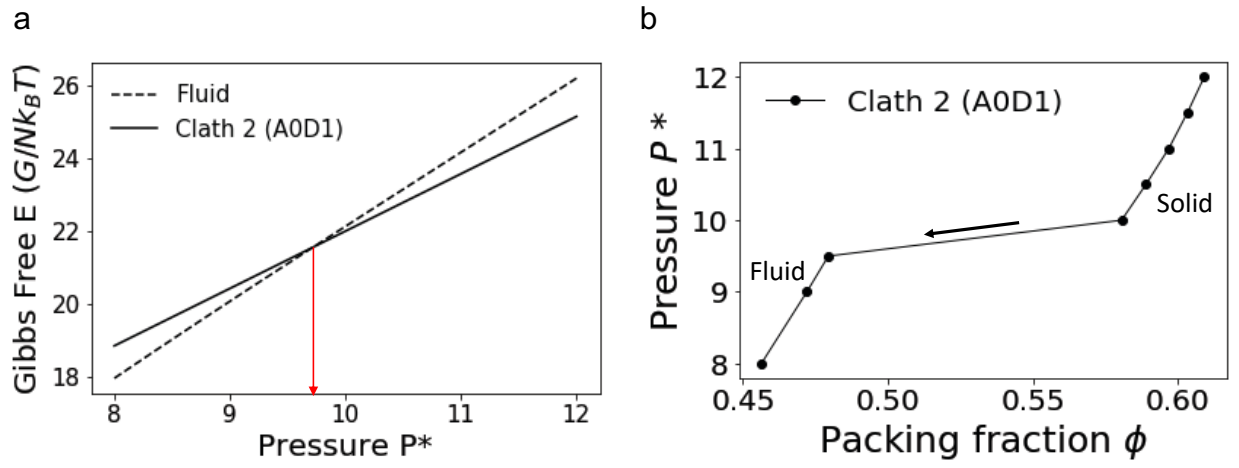


Figure 3.10: Melting pressure measurement. **a**, Gibbs free energy of the fluid and the solid phases of $S = 0.42$ system. The free energy of the fluid was calculated by thermodynamic integration and that of the solid was calculated by the Frankel-Ladd method (*Section 3.8.6*). The red arrow indicates the pressure where the free energy of the fluid and the solid crosses. **b**, The equation of state of $S = 0.42$ system, obtained by melting Clath 2 – A0D1. The crystal melts at P^*~9.7, which is a similar pressure obtained from **a**.

Table 3.1: Self-assembly simulation parameters of Figure 3.2.

S	Ensemble	ϕ	N	Resulting crystal
0.42	NVT	0.54	10,000	Clath 2 – A0D1 (Figure 3.2a)
0.48	NVT	0.55	10,000	Clath 3 – A0B1C1 (Figure 3.2b)
0.52	NVT	0.55	10,000	Clath 1 – A0B1 (Figure 3.2c)
0.64	NVT	0.59	4,000	Clath 2 – A1D4 (Figure 3.2d)

Table 3.2: Binary self-assembly simulation parameters of Figure 3.5.

TBP with	Ensemble	N_{TBP}	N_{Guest}	S_{TBP}	ϕ	Resulting crystal
Tetrahedron	NVT	9,000	1,000	0.52	0.55	Clath 1 – A1B1 (Figure 3.5a)
Dodecahedron	NVT	4,000	100	0.64	0.58	Clath 2 – A1D4 (Figure 3.5b)

CHAPTER 4

Entropically Engineered Fivefold and Icosahedral Twin Clusters of Colloidal Shapes

This chapter is reproduced from a publication in preparation, authored by Sangmin Lee and Sharon C. Glotzer.

4.1 Introduction

Twinning of crystals is a type of grain boundary where two separate crystals sharing the same lattice plane intergrow with a certain symmetry. Fivefold and icosahedral symmetries, which are generally known to be incompatible with long-range order, can be induced in relatively large-sized crystal clusters (with from hundreds to thousands of particles) via multiply twinning (*74*–*78*). For instance, materials that form the fcc (*74*, *75*) or the cubic diamond (*79*) crystals forms twin boundaries toward (111) or its equivalent lattice directions, and this allows to form $\sim 70°$ twin angle that can form multiply twinned structure with fivefold or icosahedral symmetry. This twinning behavior has been widely used to induce interesting properties of synthetic nanomaterials such as enhancing mechanical properties of nanowires (*80*) and synthesizing noble metal nanostructures with decahedral or icosahedral shape (*81*). Many studies have been conducted to control the growth mechanism (*75*, *82*) and stability of the multiply twinned structures, for example tailoring solution chemistry and changing compositions of materials. However, too many

variables are correlated and involved in these systems, thus it is hard to exclusively control a specific factor to understand its influence on the formation of the fivefold and the icosahedral twinned structures (*83*).

Compared to these systems, hard particle system is regarded as a relatively simple because the phase behaviors are driven only by the shape of particle and entropy maximization (*3*, *66*). The diversity and complexity of the entropy-driven phase behaviors, however, are enormously rich. The self-assembly of diverse complex crystals such as quasicrystals (*84*) and clathrates (*17*) have been observed, and complex nucleation dynamics such as multistep crystallization were reported in hard particle systems (*17*). In addition, the spontaneous formation of icosahedral twinned structures has been observed in the hard sphere system, both in experiments and computer simulations, when a proper system size and spherical confinement are given(*76*, *77*). The negligible free-energy difference between fcc and hcp crystals (*85*) minimizes the energy loss from the formation of twin boundaries that have the same local environment with the hcp (*74*), which makes the hard spheres favorable to form the icosahedral twinned crystal. However, the formation and stabilization of the fivefold and icosahedral twin crystals without the geometrical confinement of system have not been reported in any hard particle systems.

Here, using Monte Carlo (MC) simulation, we show that hard truncated tetrahedra (TTs), a purely entropy-driven system, form fivefold and icosahedral twin crystals at the coexistence state between the fluid and the solid (Figure 4.1). By controlling the edge and vertex truncation, the shape of hard TTs can be designed to have a negligible free-energy difference between cubic diamond and hexagonal diamond crystals. The stability of crystal boundaries between the cubic and hexagonal diamonds in fluid increases as the free-energy difference between the two crystals decreases (Figure 4.2). This indicates that the stability of twin boundaries in fluid can be

entropically controlled by a proper design of the shape of the TTs. Using the appropriately truncated hard TTs, we show the growth of fivefold and icosahedral twin crystals from either twinned or single-crystalline seeds. The formation of the fivefold twinned structure in melt follows an error-and-repair mechanism, increasing structural quality of the multiply twinned structure (Figure 4.3). We show that the icosahedral twinned structure of hard TTs is stable at the coexistence state with the fluid, unlike the hard spheres (Figure 4.4). We show that the fluid-solid interfacial free-energy of hard TTs is about two times higher than that of the hard spheres (Figure 4.4), indicating that a high fluid-solid surface tension can be entropically induced by an appropriate selection of particle shape.

4.2 Cubic and hexagonal diamond of hard truncated tetrahedra

Diamond crystal consisting of tetrahedral networks exists in two different symmetries in nature, cubic diamond and hexagonal diamond (or lonsdaleite). and the symmetry is determined by the conformation of the tetrahedral bonds. The tetrahedral network in cubic diamond has 100% staggered conformation, while that of hexagonal diamond has a 3:1 ratio between staggered and eclipsed conformations. For a system with tetrahedra particles, the two bond conformations can be mapped to the two types of pair contacts: staggered pair and eclipsed pair (Figure 4.1a). In previous studies, the formation of cubic diamond crystal in hard truncated tetrahedra was confirmed in computer simulation (*18*), where all pair contacts have staggered conformations (Figure 4.1b). However, the formation of hexagonal diamond crystal has not been reported yet. If one can control the ratio of the bond conformations in the tetrahedra system, the formation of the hexagonal diamond crystal can be induced (Figure 4.1c), and the formation of the twin boundary can also be induced because the cubic and hexagonal diamonds share an equal plane along (111) and (0001) directions, respectively (Figure 4.1b-4.1d).

The twinning of cubic crystal occurs towards (111) plane as well as its equivalent planes such as $(\bar{1}11)$, $(1\bar{1}1)$ and $(11\bar{1})$, thus when twinning occurs in multiple directions, two twin planes meet with a specific angle, $\cos^{-1}(1/3) = 70.5°$, which is close to $72°$ that enables the formation of multiply twinned structures with fivefold symmetry (*86*) (Figure 4.1e). The structure of the multiply twinned crystals made by tetrahedron particles can be analyzed by classifying particles based on the number of staggered and eclipsed pairs (*Section 4.7*), following the notation S_nE_m, $n + m = 4$. For instance, cubic diamond crystal consists of S_4E_0 particles (Figure 4.1a and 4.1b), and hexagonal diamond crystal consists of S_3E_1 particles (Figure 4.1a and 4.1c). In fivefold twin crystal, five tetrahedra clusters made by S_4E_0 particles form a pentagonal bipyramid supercluster, and five twin planes made by S_3E_1 particles are placed between faces of the tetrahedra clusters (Figure 4.1e). At the point where the five twin planes meet, S_2E_2 particles form a column where small pentagonal bipyramids (N=5) are stacked. In icosahedral twin crystal (Figure 4.1f), twenty tetrahedra clusters (S_4E_0 particles) form an icosahedron supercluster (Figure 4.1g), where the center of mass of tetrahedra clusters are placed at the vertices of a dodecahedron. Thirty twin planes (S_3E_1) exist between the faces of the tetrahedra clusters (Figure 4.1h), and twelve fivefold columns made by S_2E_2 particles are formed from the center to vertices of the icosahedral twin crystal (Figure 4.1i). At the center, there is a dodecahedron super-cluster ($N = 100$) consisting of three different shells with icosahedral symmetry: 60 S_2E_2 particles of the outer shell form a rhombicosidodecahedron, 20 S_4E_0 particles of the intermediate shell form a dodecahedron and 20 S_1E_3 particles of the smallest shell form a dodecahedron. This structure is equivalent to the super-cluster of water $(H_2O)_{100}$ when connecting oxygen atoms (*87*).

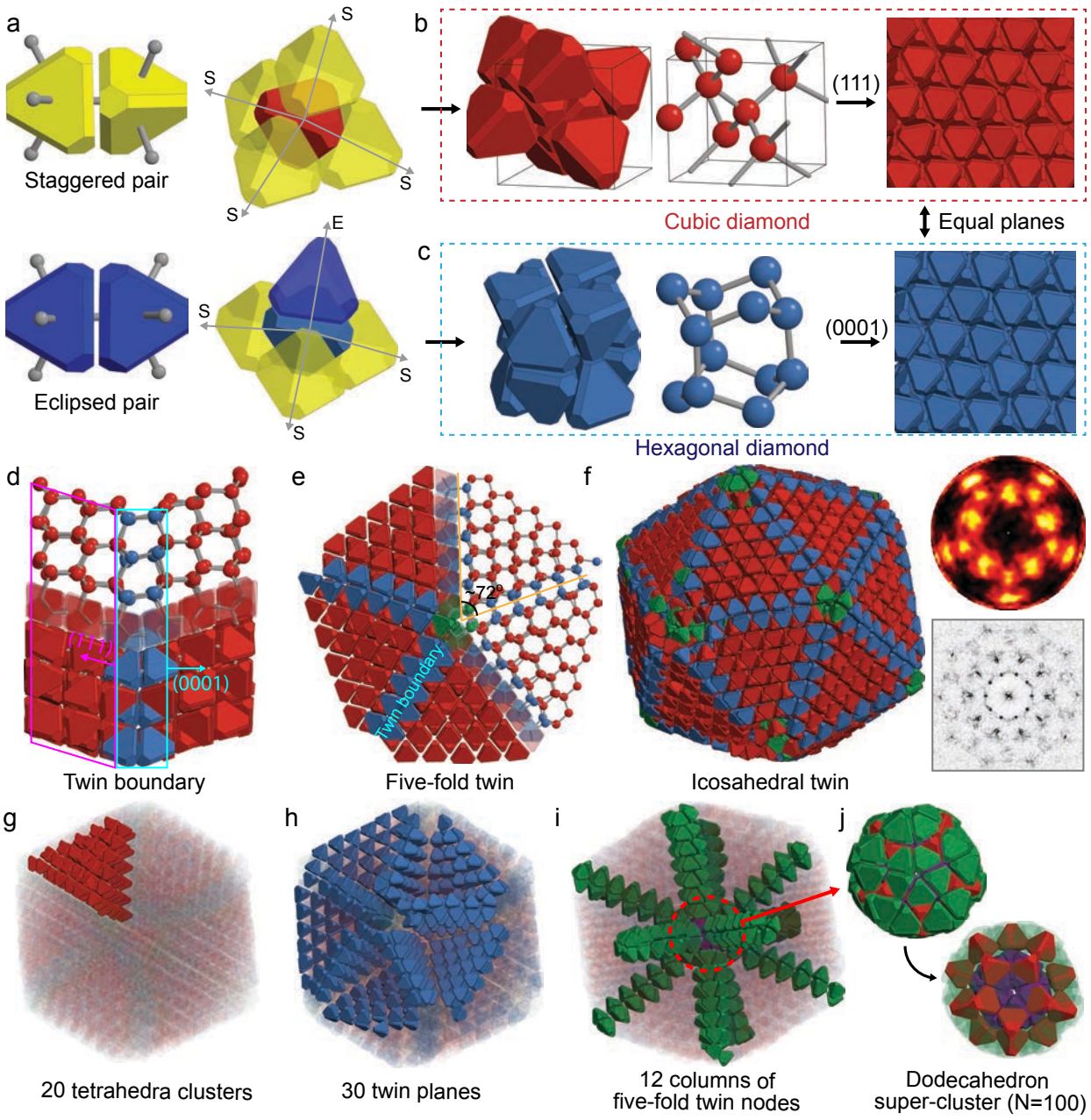


Figure 4.1: Twin crystals of truncated tetrahedra. **a**, Two types of pairwise contact (left) and the local environment of particles in cubic diamond (upper right) and hexagonal diamond (bottom right). **b-c**, Unit cells of cubic diamond (b, left and middle) and hexagonal diamond (c, left and middle) crystals. The (111) of the cubic diamond (b, right) and (0001) of the hexagonal diamond (c, right) have the same structure. **d-j**, Twinned structures of truncated tetrahedra. Red particles for S_4E_0, blue for S_3E_1, green for S_2E_2 and purple for S_1E_3. **d**, A twin boundary is formed when (111) of the cubic diamond (magenta arrow) and (0001) of the hexagonal diamond (cyan arrow) meet. **e**, Five-fold twinned structure of truncated tetrahedra. The angle between twin boundaries is ~72°. **f-j**, Structure of icosahedral twin crystal. The bond orientational

order diagram (**f**, upper right) and diffraction pattern (**f**, bottom right) along 10-fold symmetry axis. The ico-twin crystal consists of 20 cubic diamond clusters with tetrahedron shape (red in **f, g**), 30 twin planes existing between the tetrahedral clusters (blue in **f, h**) and 12 columns of node of five-fold twin structure (green in **f, i**). **j**, The center of the ico-twin is a dodecahedron super-cluster composed of a hundred of tetrahedra.

4.3 Stability control of twin boundary by particle shape design

Most particles in the multiply twinned crystals have either cubic (S_4E_0) or hexagonal (S_3E_1) diamond local environments, thus, in order to induce the formation of such structures, a method to control the preferences of the two phases is needed. Recent studies (*18*, *88*) have shown that various crystals, including the cubic diamond, can be formed in a tetrahedron shape space where the amount of edge and vertex truncation vary, but the hexagonal diamond has not been reported. However, given the fact that regular tetrahedron prefers the S_0E_4 and truncated tetrahedron forming the cubic diamond prefers the S_4E_0, we hypothesized that there should be a region that prefers the S_3E_1 forming the hexagonal diamond. Thus, we explored the shape space between the S_4E_0 and the S_0E_4 regions by adjusting the truncation parameters to a finer range than the previous studies. We constructed a free-energy landscape of the cubic and hexagonal diamonds in the tetrahedron shape space varying the edge ($1.14 \leq a \leq 1.30$) and vertex ($2.10 \leq c \leq 2.30$) truncation parameters at a constant packing fraction (Figure 4.2a), using Frenkel-Ladd free-energy calculation method (*71*, *85*). We confirmed a region where the hexagonal diamond is thermodynamically more stable than the cubic diamond, as shown in the phase diagram generated from the bottom view of the free-energy landscape (Figure 4.2b). We also confirmed the self-assembly of the hexagonal diamond. This demonstrates that the preference of pair conformation of truncated tetrahedra can be entropically controlled by designing the edge and vertex truncation of particles.

The cubic and hexagonal diamonds have an equal plane that enables the formation of twin boundaries or stacking faults, where the only difference between the two is spacing between stacking sequence changes. In general, crystal with these boundaries is thermodynamically less stable than a single-crystalline structure because the boundaries increase energy. However, the degree of instability of the boundaries can be changed by the free-energy difference between the cubic and hexagonal diamonds, and the formation of the boundaries could be induced if the free-energy difference is negligible. Using MC simulation (*Section 4.7*), we investigated the phase transition process of a single-crystalline cubic diamond cluster in the fluid when the system has different free-energy differences between cubic and hexagonal diamonds, $\Delta F = F_H - F_C = 0.20,\ 0.01,\ -0.10$ (Figure 4.2c). We prepared a large cubic diamond cluster fully surrounded by fluid at a coexistence packing fraction and equilibrated the system. We first distinguished particles in fluid and crystal based on the local density, which shows two peaks in distribution plots that indicate the two coexisting phases (Figure 4.2g). Then, we colored particles in crystal phase based on their pair contact conformations (S_nE_m) and analyzed simulation snapshots. When the ΔF is negligible ($\Delta F = 0.01$), the initial cubic diamond cluster tries to form stacking faults in multiple (111) equivalent directions at the early stage, and a certain predominant direction is selected and repeatedly forms the stacking faults (Figure 4.2d). This suggests that the formation of stacking faults is easily induced when the thermodynamic stability of the two crystals is comparable. On the other hand, when the hexagonal diamond is more stable ($\Delta F = -0.10$), a stacking fault is formed along (111) direction at the early stage, then a complete phase transformation into a single-crystalline hexagonal diamond was observed. Additional formation of stacking faults was rarely observed during the transformation. The difference between the two systems is more clearly observed from the change of the number ratio of S_4E_0 particles (N_{cubic}/N_{solid}) over MC steps

(Figure 4.2f). The change of the N_{cubic}/N_{solid} for the $\Delta F = -0.10$ system shows a more rapid decrease compared to that of the $\Delta F = 0.01$ system due to the frequent formation of stacking faults. This shows that the frequency of formation of stacking faults or twin boundaries in melt can be entropically controlled by the shape design of the particle.

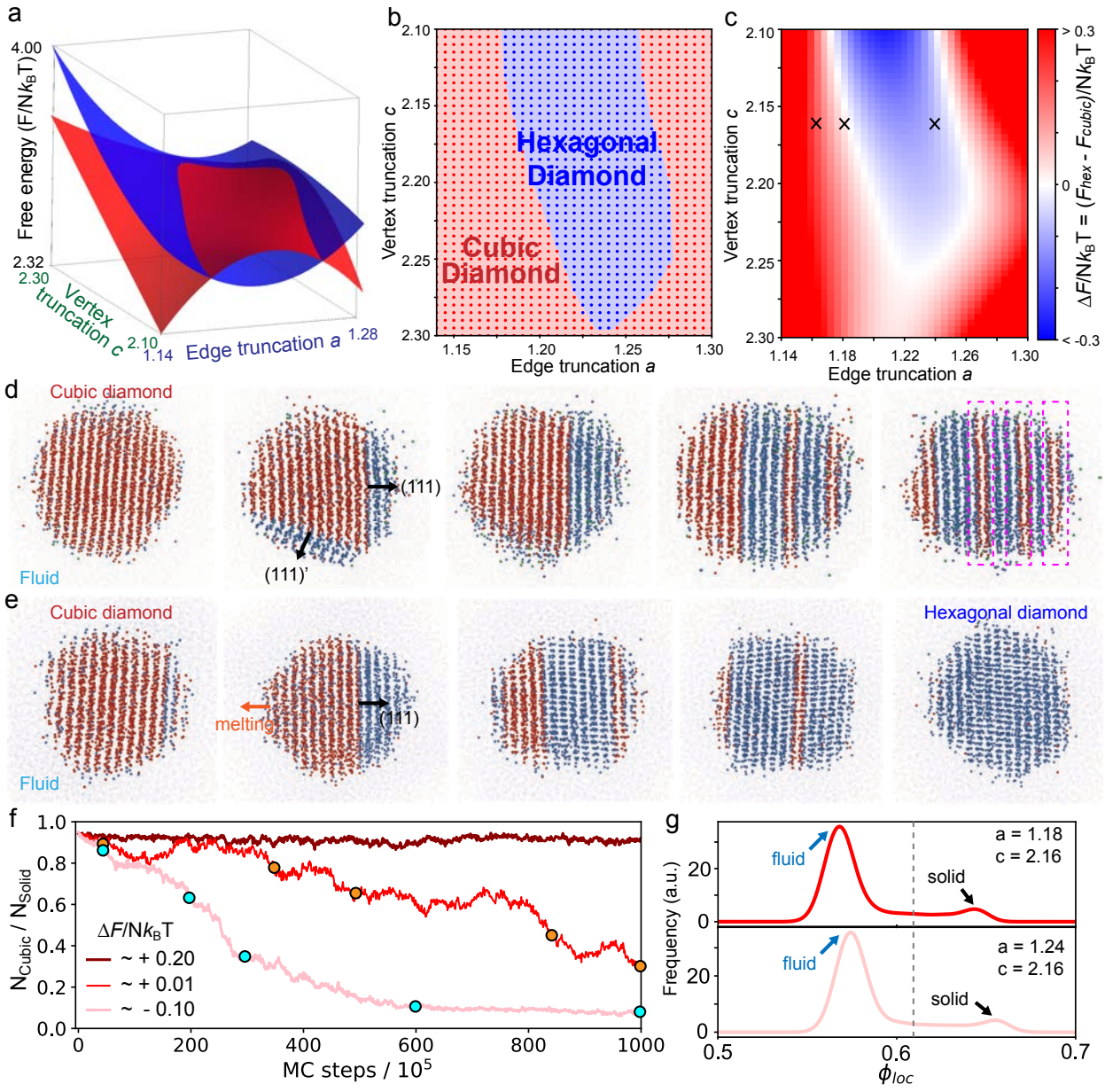


Figure 4.2: Stability control of diamond crystals via particle shape design. **a**, Helmholtz free-energy (F/Nk_BT) plot of cubic (red) and hexagonal (blue) diamond crystals of hard-TTs as a function of vertex

and edge truncation parameters at constant packing density ($\phi = 0.58$). **b**, Phase diagram of hard-TTs in the shape space determined by the free-energy calculation. **c**, Free-energy difference between the cubic and hexagonal diamond ($\Delta F = F_{\mathrm{H}} - F_{\mathrm{C}}$) in the shape space. **d**, When the TT is designed ($a = 1.18, c = 2.16$) to have negligible $\Delta F \sim 0.01$, the initial cubic diamond (left) single-crystalline cluster forms multiple stacking faults (right) in melt, suggesting that the energy loss from the twin boundaries is small. **e**, When the TT is designed ($a = 1.22, c = 2.16$) to have large enough $\Delta F \sim -0.10$, the initial cubic diamond cluster (left) completely transform into single-crystalline hexagonal diamond (right) cluster in a relatively short simulation time. **f**, The change of number ratio of particle in cubic diamond over time for three different $\Delta F = +0.20, +0.01$ and -0.10 (black X markers in **c**). Orange and cyan circles indicate each snapshot in **d** and **e**, respectively. **g**, Local density distribution plots calculated from the last snapshot from **d** (upper) and **e** (lower). Two peaks in the distribution show that the system is in coexistence phase between fluid and solid.

4.4 Formation of fivefold and icosahedral twin clusters

Based on the finding, we hypothesized that the formation of a multiply twinned structure might be induced when the ΔF is negligible because the formation of twin boundaries toward multiple directions will easily occur. We conducted seed-assisted growth simulation (*Section 4.7*) in melt with a system where the ΔF is negligible ($a = 1.18, c = 2.16$). We placed a small spherical cubic diamond seed ($N = 500$) in fluid and grew the seed at a coexistence state, $\phi = 0.575$, which is a state point that has enough fluid to fully surround the solid phase in equilibrium. In the early stage of growth, we observed the formation of twin boundaries in multiple directions (Figure 4.3a1-4.3a2). Through the subsequent multiply twinning, a fivefold center was formed (Figure 4.3a3), but an error can occur because growth occurs independently in multiple directions, which can cause a mismatch in the stacking sequence between two adjacent directions (Figure 4.3a4). In this case, however, we observed a repair mechanism, where one of the growing directions with the error is re-melted toward its opposite direction (Figure 4.3a5) until the boundary with its adjacent plane matches. Once the boundary matches (Figure 4.3a6), the fivefold twin crystal fully grows and is stabilized in fluid with truncated pentagonal dipyramid crystal shape (Figure 4.3a7-4.3a8).

This is additional evidence of the formation of fivefold twin crystals via the error-and-repair mechanism, but a different process from the rearrangement of particles observed in the nanosphere systems in solution (*82*).

The icosahedral twin cluster is known to be formed by repeatedly inducing the twinning from the fivefold twin crystal (*75*), and this is well applied to the entropy-driven system. We performed a growth simulation of the fivefold twinned crystal seed at a fluid-solid coexisting state (Figure 4.3b). We kept the ΔF negligible ($\Delta F \sim 0.01$ at $a = 1.18, c = 2.16$) for the twinning to be easily induced. We observed the growth of twinning from the surface of the fivefold twinned crystal seed and eventually formed an icosahedral twinned structure (Figure 4.3b), which reminds the formation process of a twinned icosahedron of gold atoms (*75*). We separated the fluid and solid phases from the local density calculation (Figure 4.3d) and confirmed that the solid cluster was fully surrounded by fluid during the formation of the ico-twin crystal. After screening particles of the fluid-solid interface (Figure 4.3d) that have unclear structure, we confirmed that all the twenty surfaces of the ico-twin cluster are (111) plane of the cubic diamond (Figure 4.3c). This suggests that the (111) plane of the cubic diamond of hard truncated tetrahedra system has the lowest fluid-solid interfacial free-energy compared to the other directions.

The stability of the ico-twin cluster in melt is evaluated by tracking the change of reduced pressure ($P^* = Pv_0/k_BT$) of the system over MC steps (Figure 4.3e). During the formation process ($\leq 50 \times 10^5$ steps), we observed a rapid decrease of pressure. After the growth, the pressure was stabilized around a constant value ($P^* \sim 12.6$) with small fluctuation, indicating that the ico-twin structure is stabilized in melt. From the simulation snapshots and diffraction patterns at 100×10^5 and 175×10^5 steps, we confirmed that the icosahedral symmetry of the cluster is maintained (Figure 4.3e, insets). Structurally, the icosahedral twin crystal has internal and surface

strain (*89*), thus the effect of compensating the strain energy is required to stabilize the cluster. For instance, the hard sphere system requires a spherical confinement effect to entropically stabilize the icosahedral cluster (*76–78*). Then, why the icosahedral cluster of the hard truncated tetrahedra is stable in melt without any external confinement effect?

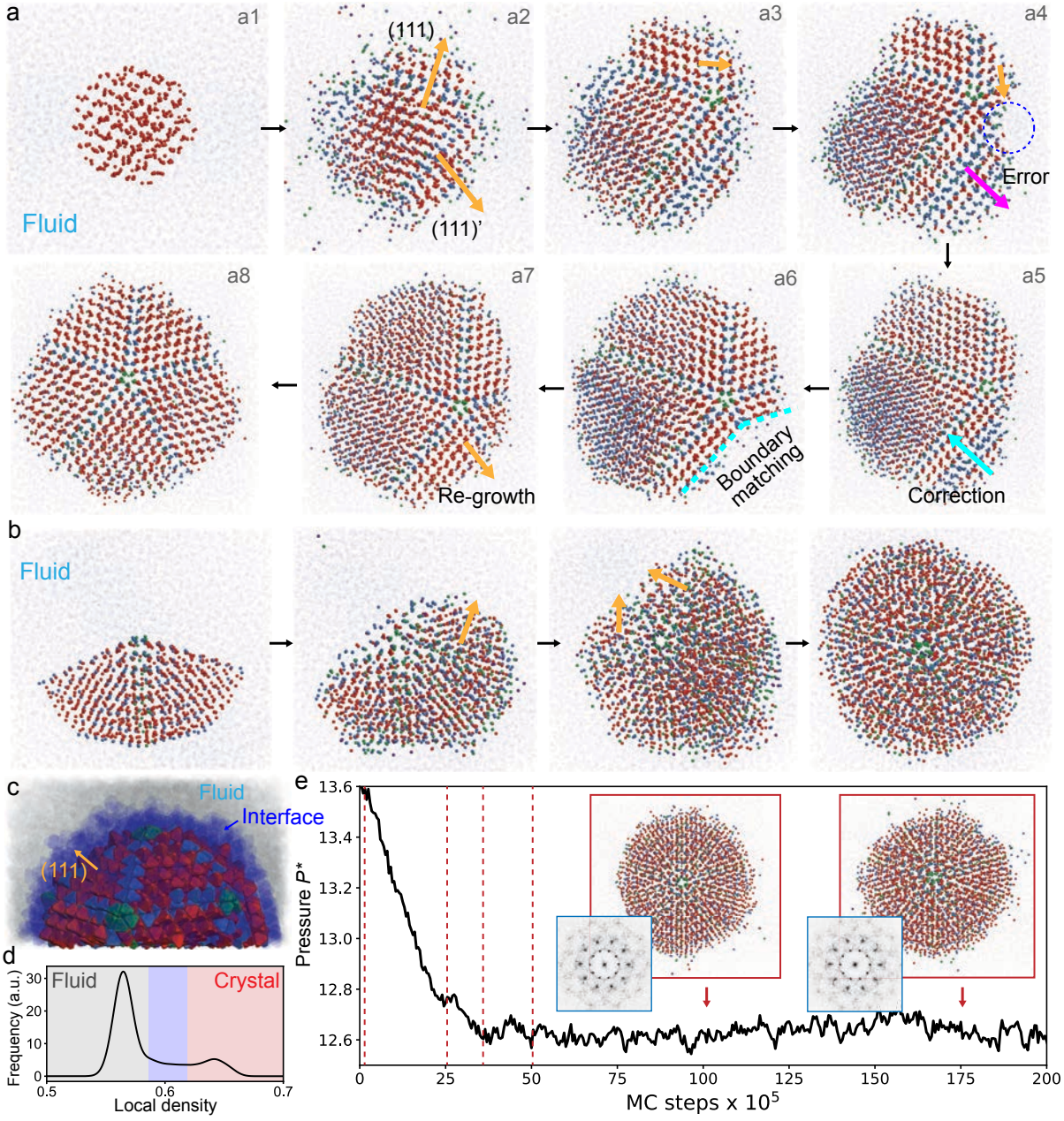


Figure 4.3: Growth process of fivefold and icosahedral twin crystals from seed in melt. Seed assisted growth of the twin crystals of hard-TTs in a coexistence phase between crystal and fluid ($\phi = 0.58$). **a**,

Simulation snapshots showing the error-and-correction formation mechanism of the five-fold twin crystal. All particles are represented by spheres showing center of masses. (a1) When a spherical seed of cubic diamond crystal (N=1,000) grows, (a2) twin boundaries are formed along (111) or its equivalent plane directions. Because the growth of each direction is independent, (a3-4) an error can occur when stacking sequences of two growth directions mismatch. (a5) The growing direction with the error is re-melted toward its opposite direction (a6) until the boundary with its adjacent plane matches. Once the boundary matches, (a7-8) the five-fold twin crystal grows to fully form the five-fold twin crystal with truncated pentagonal dipyramid (PD) crystal shape. The final crystal is fully surrounded by fluid and stabilized. **b**, Simulation snapshots of icosahedral twin crystal formation from PD seed. All particles are represented by spheres showing center of masses. **c**, The icosahedral twin crystal exposes (111) surfaces of cubic diamond crystal when stabilized in melt. **d**, Local density distribution plot of the coexistence system between the ico-twin and the fluid shows a bimodal shape at $\phi = 0.57$ and 0.62. **e**, The change of pressure over time for the ico-twin crystal forming system. Pressure decreases during the growth of the ico-twin crystal and is stabilized on a constant value (P^*~12.6) after the growth, indicating that the ico-twin crystal is stable in melt. The red dashed lines at the early stage indicate the simulation time when the snapshots in **b** are taken. The inset snapshots and diffraction patterns show that the ico-twin crystal structure surrounding by fluid is maintained during the pressure is stabilized.

4.5 The effect of fluid-solid interfacial energy on the stability of icosahedral twin cluster

To confirm the difference between the two systems, we first checked if the icosahedral twin cluster of hard spheres can be stabilized in melt without spherical confinement. We conducted MC simulation of hard spheres with an ideally constructed icosahedral twin cluster of face-centered cubic (FCC) crystal (*77*) which is fully surrounded by the fluid at a fluid-solid coexistence packing fraction ($\phi = 0.515$). From the simulation snapshots, we confirmed that the icosahedral cluster is rapidly destabilized in melt within 10^5 MC steps (Figure 4.4a). This suggests that the hard spheres have a much weaker fluid-solid surface tension than that of the hard truncated tetrahedra. This is because stronger surface tension makes the solid cluster more compact in fluid, which induces a similar effect to the spherical confinement that enables entropic stabilization of

the icosahedral cluster. The interfacial energy of (111) plane is particularly important (Figure 4.4b), because it is the largest surface exposed to the fluid.

The fluid-solid interfacial free energy can be calculated by the fluctuation of the interface using the capillary fluctuation method (*90*, *91*). Briefly, we expose a specific lattice plane of a crystal toward the fluid phase in a thin slab (z-axis in Figure 4.4c-4.4d) at a coexistence state point and monitor the fluctuation of the interface in equilibrium to calculate interfacial stiffness ($\tilde{\gamma}$) of the plane. After obtaining the $\tilde{\gamma}$ from multiple lattice directions, we can measure the interfacial free energy (γ) of each lattice plane. The interfacial profile normalized by the circumscribed radius of the particle ($h - h_0/r_{\mathrm{circ}}$) shows that the fluctuation of (111) plane of the hard sphere system is larger than that of the hard truncated tetrahedra system. Quantitatively, the interfacial stiffness of the FCC of hard spheres, $\tilde{\gamma}_{\mathrm{HS\_(111)}} = 0.684\ k_B T\sigma^{-2}$, is 56% lower than that of the cubic diamond of the hard truncated tetrahedra, $\tilde{\gamma}_{\mathrm{TT\_(111)}} = 1.547\ k_B T\sigma^{-2}$. From the interfacial free energy calculation, we confirmed that the fluid-solid interfacial free energy of the (111) plane of the hard truncated tetrahedra system is ~2.4 times higher than that of the hard sphere system, $\gamma_{\mathrm{TT\_(111)}} = 1.347\ k_B T\sigma^{-2}$ and $\gamma_{\mathrm{HS\_(111)}} = 0.560\ k_B T\sigma^{-2}$. This indicates that the icosahedral twin cluster of the hard tetrahedra receives higher interfacial tension in fluid, which induces a similar effect to the spherical confinement effect. The strong self-confinement effect allows the icosahedral cluster to be stable in melt despite the internal strain of the structure. The relationship between the particle shape and the fluid-solid interfacial tension is not clear yet, but we speculate that one reason is that the flat (111) plane of the cubic diamond of truncated tetrahedra (Figure 4.4b) gives a larger structural difference with fluid. Therefore, it will be an interesting future study to investigate the relationship between particle shape and fluid-solid interfacial tension.

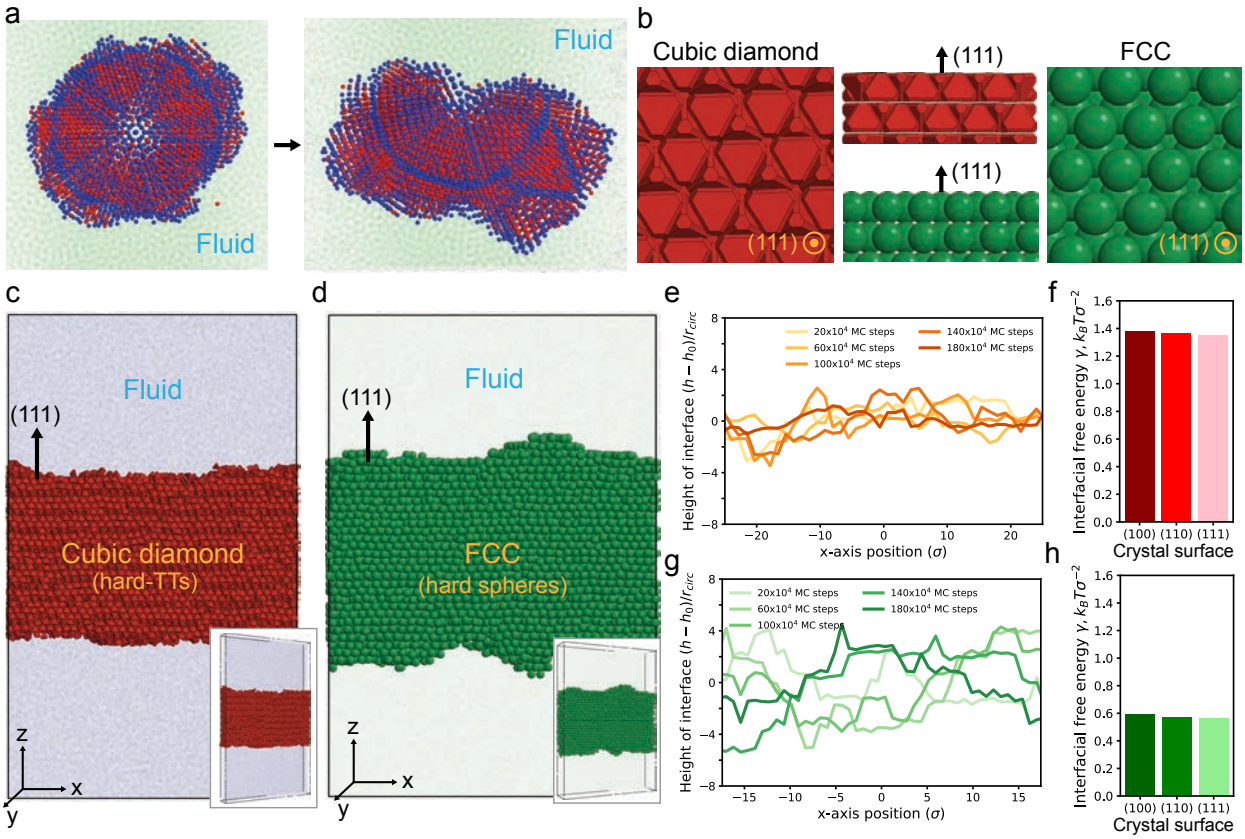


Figure 4.4: Fluid-solid interfacial energy calculation. **a**, Icosahedral twin crystal of FCC of hard spheres is quickly destabilized in melt ($< 10^6$ MC steps). **b**, (111) surfaces of the cubic diamond of hard-TTs (left) and the FCC of hard spheres (right), and their side views (middle). **c-d**, Simulation setup of the capillary fluctuation method for calculating the fluid-solid interfacial stiffness ($\tilde{\gamma}$) of (**c**) the cubic diamond and (**d**) the FCC. **e-g**, The change of interfacial profiles of (111) direction of (**e**) the cubic diamond and (**g**) the FCC. **f-h**, Fluid-solid interfacial free-energy (γ) of (**f**) the cubic diamond and (**h**) the FCC. For all three lattice directions, the hard-TT system has more than two times higher γ than that of the hard sphere system, indicating its stronger fluid-solid surface tension.

4.6 Conclusions

In summary, we have reported an approach to induce the entropy-driven formation and stabilization of fivefold and icosahedral twin crystals in melt. The free-energy calculation showed that the thermodynamic stability of the cubic diamond and hexagonal diamond can be controlled by the proper design of vertex and edge truncation of the hard tetrahedron. When the free-energy

difference between the two diamond crystals is negligible, the formation of stacking faults or twin boundaries is easily induced in melt. This enables the entropy-driven formation of fivefold and icosahedral twin clusters in melt via seed-assisted growth, and the error-and-repair mechanism was observed during the formation of the multiply twinning. We confirmed that a high fluid-solid interfacial tension is required to stabilize the icosahedral cluster through comparison with the hard sphere system.

In conclusion, our findings provide a quantitative understanding of the formation and stabilization of fivefold and icosahedral twin crystals in melt through entropy control. Importantly, we showed that the twinning behavior and the interfacial energy can be entropically controlled by particle shape. This implies that more diverse and complex entropic phase behaviors can be discovered through the thermodynamic stability control of the fluid-crystal or the crystal-crystal interfaces.

4.7 Method and supporting information

4.7.1 Particle geometry

The truncated tetrahedron (TT) shape is a member of the spherical triangle invariant 323 family, with truncation parameters (a, b, c) according to previous convention (*53*). In this study, $b = 1.0$ for all systems, and a and b vary in a range: $1.14 \leq a \leq 1.30$ and $2.10 \leq c \leq 2.30$.

4.7.2 Identification of staggered and eclipsed pairs

Each TT in a crystal phase has four nearest neighbors that form face-to-face contacts. We identified the type of each face-to-face contact (staggered or eclipsed), following a similar identification protocol described in Ref. (*92*). Briefly, for each face of the particle, we assigned three vectors that are defined from the face center to each tip of the face (Figure 4.5a). Then, for

each pair contact (i, j), we can define three vectors for i particle and three other vectors for j particle (Figure 4.5b and 4.5c). We calculated every pair combination between the i and j vectors, which is nine, and found the minimum angle (θ). All the θ from every pair contact accumulated from the entire system give an angle distribution. For instance, Figure 4.5d was calculated from hexagonal diamond crystal. The angle distribution shows a clear bimodal distribution, and the height of the peak at θ_1 and θ_2 are around a 1:3 ratio, indicating that the angles around θ_1 and θ_2 come from eclipsed pairs and staggered pairs, respectively. This allows us to decide the range of the pair contact angle to identify the eclipsed pair ($\theta_1 \pm 10°$) and the staggered pair ($\theta_2 \pm 10°$).

4.7.3 Monte Carlo simulation

Simulations were performed by the hard particle Monte Carlo (HPMC) implemented in the HOOMD-blue simulation package (*54*, *93*), which is available at https://github.com/glotzerlab/hoomd-blue. The system size for the self-assembly simulations (Figure 4.2) is $N = 2{,}000$, and all simulations were performed under periodic box condition. The translational and rotational movement of particles is decided to have around a 20% acceptance ratio for every simulation. The self-assembly simulations were initialized from a dilute isotropic fluid $\phi = Nv_0/V < 0.01$ and compressed until the desired thermodynamic condition (ϕ or reduced pressure $P^* = Pv_0/k_BT$) was reached. Here, v_0 and V are the volume of a particle and the volume of the simulation box, respectively. The unit length of simulation is defined by σ, and the particle volume of a TT is set to $v_0 = 1.0\sigma^3$ regardless of its truncation amount. After initialization, each run was continued in the isochoric ensemble (NVT) at constant packing density ϕ until equilibration was reached. For the most systems, crystallization occurs within 1.0×10^8 Monte Carlo sweeps.

4.7.4 Seed-assisted growth MC simulation

We conducted seed-assisted growth simulation at a fluid-solid coexistence state ($\phi = 0.575$ for $a = 1.18, c = 2.16$ system). For the growth of the fivefold twin cluster, we used a spherical cluster of a cubic diamond crystal ($N = 500$) as a seed, and the total system size including the seed is $N = 20{,}000$. At a very dilute condition ($\phi = 0.01$), we placed the seed at the center of simulation box and placed other particles around the seed without overlap. Holding the seed particles (no movement), we compressed the simulation box to the target packing fraction. Once it reaches the target packing fraction, we released the seed particles and equilibrated the system. For the growth of the icosahedral twin cluster, we used a fivefold twin cluster ($N \sim 1{,}000$) as a seed, and the total system size including the seed is $N = 20{,}000$. For compression and equilibration protocol is the same as the former case.

4.7.5 Local density calculation

Seed-assisted growth simulations for fivefold and icosahedral twin clusters were performed in a coexistence phase between a fluid phase and a crystal phase, where the solid clusters are fully surrounded by fluid. To distinguish the two phases, we calculated the local packing density ϕ_{loc}, which is defined as the packing density around a particle within a certain radius, r_{cut}. The r_{cut} in this study was determined by 3 times longer than the distance from a particle center to its nearest neighbor particle center, in order to properly average the local environment. TTs are identified as belonging to the fluid region if $\phi_{\mathrm{loc}} \leq 0.60$ and to the crystal region if $\phi_{\mathrm{loc}} > 0.60$. We utilized the freud (*94*) python library for this calculation.

4.7.6 Stability check of an icosahedral twin cluster of hard spheres

We performed MC simulation of hard sphere system to check the stability of icosahedral twin cluster of hard spheres at a fluid-solid coexistence phase ($\phi = 0.515$). We first constructed

an icosahedral cluster ($N = 5971$) with an ideal structure, following the method described in Ref. (*77*). At a very dilute condition ($\phi = 0.01$), we placed the ico-twin cluster at the center of simulation box and placed other hard spheres around the cluster without overlap. Holding the ico-twin particles (no movement), we compressed the simulation box to the target packing fraction. Once it reaches the target packing fraction, we released the ico-twin cluster particles and equilibrated the system. To check if the ico-twin cluster is stable in the coexistence phase, we monitored the bond-order diagram and the morphology of the ico-twin cluster during the simulation. We observed that the icosahedral structure is rapidly destabilized within 5×10^5 MC sweeps (Figure 4.4a).

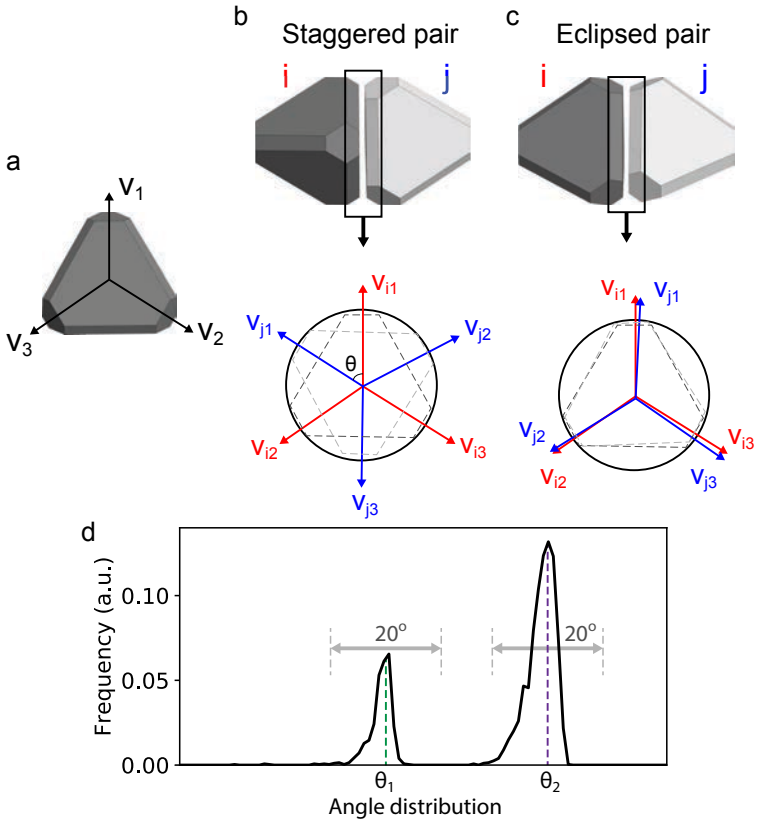


Figure 4.5: Pair contact type identification. **a**, Three vectors are assigned to each face of particle. **b,c**, For each pair contact of i and j TTs (top), we found a minimum angle (θ) between the three vectors of i and the three vectors of j (bottom). **d**, Distribution plot of the θ for a hexagonal diamond crystal.

CHAPTER 5

Clathrate Colloidal Crystals with DNA

This chapter is reproduced from Ref. (*59*), a publication in *Science*, authored in 2017 by Haixin Lin[†], Sangmin Lee[†], Lin Sun, Matthew Spellings, Michael Engel, Sharon C. Glotzer[*] and Chad A. Mirkin[*] ([†]equal contribution, [*]corresponding authors). My contributions to this work included structure identification of the crystal structures found in experiments, simulation model development of DNA-programmable colloids with polyhedral shape and performing all self-assembly simulation and their analysis.

5.1 Introduction

DNA-programmable assembly has emerged as a powerful method for generating colloidal crystals with exquisite control over crystal symmetry, lattice parameter, and in certain cases macroscopic habit (*12*, *95–97*). Indeed, more than 500 different crystals spanning 36 different symmetries have been made via this approach and the complementary contact model and associated design rules that govern DNA-programmable colloidal crystal formation (*13*, *15*, *98–100*). The vast majority of such structures studied to date consist of spherical nanoparticle building blocks, isotropically functionalized with DNA, termed programmable atom equivalents (PAEs) (*101*). With the advent of methods for preparing anisotropic PAEs, DNA-modified structures can be made where the shape of the nanoparticle core directs the DNA bonding elements in an anisotropic manner (*99*, *102–105*). This is making it possible to synthesize colloidal examples of

many solid-state structures that naturally occur, as well as several others that have no known mineral equivalent (*4*, *99*, *102*, *106*). Clathrates, structures consisting of polyhedral cages with large pores that can be used for host-guest chemistry, represent a challenging target for colloidal assembly (*36*, *107–109*). Because of the symmetry and complexity of clathrates, it is not clear how such architectures can be made from spherical building blocks, but they may be attainable with anisotropic ones. Herein, we explore, via experiment and simulation, how the symmetry of nanoscale gold bipyramids can be used in conjunction with programmable assembly to generate several clathrate architectures.

5.2 Result and discussion

Clathrates occur as hydrates of a variety of host molecules such as methane and as open four-fold coordinated crystals in the carbon group (*110–113*). The cages are formed from molecular or atomic nodes that adopt discrete bond angles between 100° and 125°. With the development of synthetic methods for forming monodisperse samples of anisotropic particles (*114*, *115*), shapes can be explored that impose directionality on oligonucleotide bonding motifs (*99*, *102*). In this regard, oblate trigonal bipyramids (TBPs) with {110} facets, which direct bonding elements with ~110° angles between them, are ideal structures to explore nanoparticle-based colloidal clathrate formation (Figure 5.1A and 5.1B). To prepare building blocks suitable for assembly, we synthesized 250-nm long-edge-length TBPs via reported procedures (*116*) and functionalized them with 28 base hexylthiol-modified DNA (Figure 5.1C, red; see Table 5.2 for sequences of all DNA strands. After purification by centrifugation and washing with phosphate-buffered saline, these particles were hybridized with a variety of different linker strands (Figure 5.1C, blue, 23 to 228 bases) that vary in length but are all terminated with a self-complementary GCGC sticky end. Prior to their binding to the particles, the long linkers were introduced to

partially complementary strands (Figure 5.1C, green) to form a primarily duplex region except for the four base self-complementary termini, 10 A bases near the particle that were maintained as single-stranded DNA for flexibility purposes, and additional individual bases (black) at designated locations that served as flexors facilitating assembly (*13*).

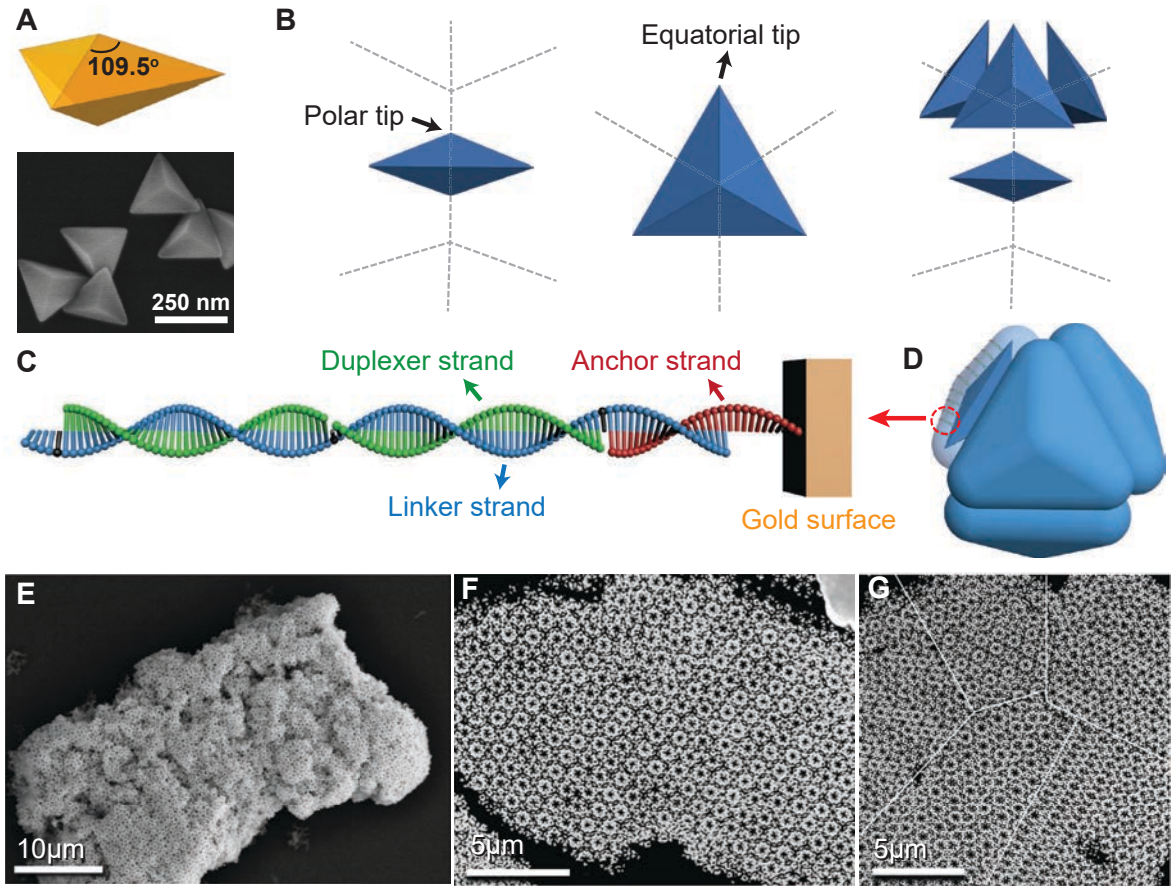


Figure 5.1: Self-assembly into clathrate colloidal crystals. (A) Geometry and scanning electron microscopy (SEM) image of gold TBPs with {110} facets and 109.5º large-edge angle. (B) When placing TBPs at the middle of network edges and rotating them appropriately, all triangular TBP facets align nearly parallel. In the resulting tetrahedral network, polar tips join at network nodes and equatorial tips join in the center of network cages. (C) Illustration of the DNA linker design. The linker binds to the surface of Au bipyramid through a 28-base hexylthiol-anchor strand (red), which can recognize a linker strand (blue). Duplexer strands hybridize the linker strand to form double-strand segments, except for specific single-base sites (black) and the 4-bases self-complementary sticky end. DNA length can be tuned by varying the number of duplexed block segments. (D) The TBP tetramer with its DNA shell has rounded edges and vertices.

(E) Whole-view SEM image of the self-assembled TBP superlattice. (F and G) TEM image of the superlattice zoomed in on single-domain and multidomain regions from a sectioned sample.

We hypothesized that using long DNA strands that have enough length and flexibility would be important to reduce the strain in the material imposed upon assembly. In fact, if we tried to physically assemble TBPs of this size by sticking faces together, structures with relatively large gaps of up to 20° will form, creating the need for long, flexible DNA bonds, which simultaneously also round the outer DNA shell (Figure 5.1D). Indeed, when the particles were allowed to assemble by maintaining the temperature of the solution below the DNA melting temperature, long DNA linkers enabled the formation of high-quality crystalline structures (Figure 5.1E). Closer inspection of TEM images suggested that the assemblies were clathrates (Figure 5.1F). In addition to large-area single-domain crystals, we observed multi-domain architectures that contained a mixture of at least two different but related materials (Figure 5.1G and Figure 5.12).

To investigate the effect of the DNA linkers, we tuned their length by changing the number of 40-base duplexed block segments (Table 5.3). For the shortest DNA linker without block segments, TBPs only formed tetrahedral units and higher-level structures were not found (Figure 5.2A). Some local order was observed when we used DNA linkers with one block segment, but substantial distortion was still apparent (Figure 5.2B). Multishell order resembling clusters but with defects and cracks occurred when using DNA with two and three block segments (Figure 5.2, C and 5.2D). Clearly identifiable clathrates appeared when using DNA with four block segments, but the domain size was usually small (Figure 5.2E). The highest-quality crystals formed when DNA with five block segments was used, and single-crystalline clathrate domains as large as tens of micrometers were found in such structures (Figure 5.2F). As expected for clathrates, cavities

were distributed throughout the sample (Figure 5.2G and 5.2H), which were generally guest-free in all of the samples.

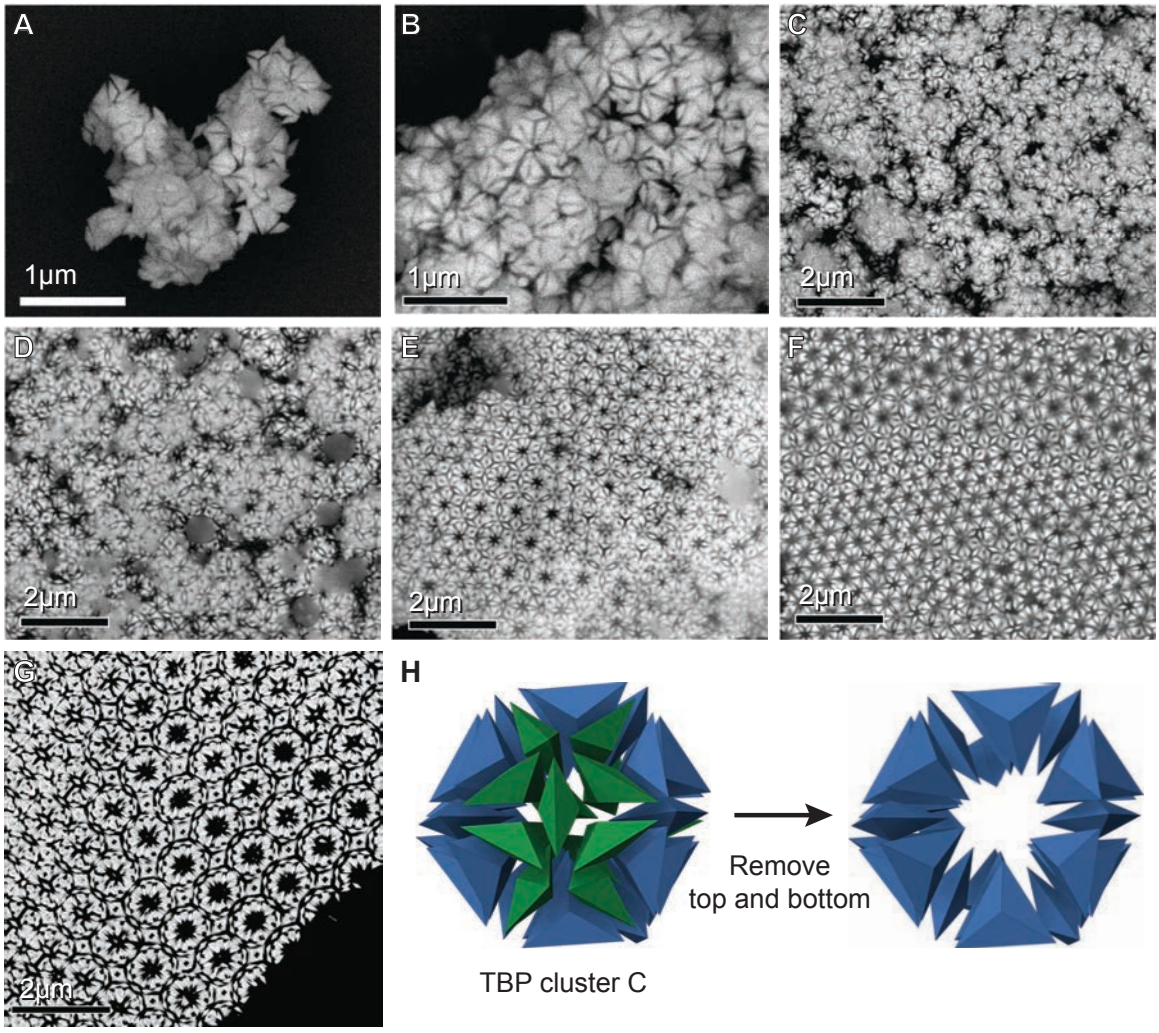


Figure 5.2: Effect of DNA length on TBP assembly and presence of cavities. SEM images of samples using TBP particles with DNA bonding elements that contain (A) zero, (B) one, (C) two, (D) three, (E) four, (F) five block segments paired with duplexer strands. An improvement of the assembly quality with the increase of the block segment number is apparent. (G) Cavities are observed in the middle of the clusters representing clathrate cages. Shown in this TEM image is a thin sectioned sample that contains the middle portion of clusters showing the empty spaces (center of the image) and off-center sectioned clusters including the upper part of the clusters (top left of the image, see also Figure 5.11h). The particles in this sample have DNA ligands with five block segments. (H) Illustration of a single cluster (identified as cluster C, Figure 5.6D) before and after removing the top and bottom TBPs.

To resolve the complex structure of the clathrates, we first reproduced the assemblies thermodynamically using a discrete element method with implicit solvent (*117*). We constructed a minimal model whereby a single PAE is represented as a rigid TBP core surrounded by a DNA ligand shell (Figure 5.3, 5.4, 5.6A and *Section 5.4* for details). The interaction between PAEs is described by an effective pair potential that depends only on the separation distance measured from patches on the TBP core (Figure 5.4 and Table 5.1). In this simplified representation the increased rounding of the PAE observed experimentally with increased DNA length is automatically included. We chose the shell thickness, determined by DNA length, to match that used in the experiments.

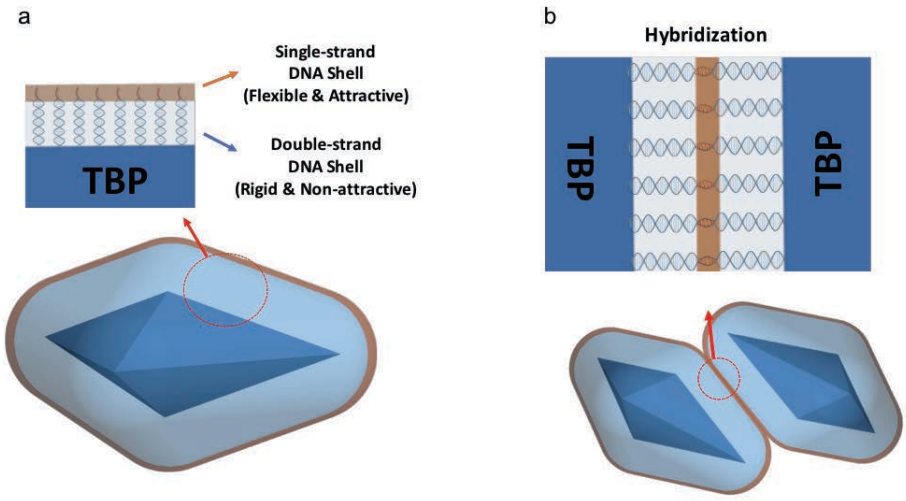


Figure 5.3: Schematic diagram of the minimal model used in simulation. (a) The TBP core is surrounded by a DNA shell composed of a rigid and non-attractive ds-DNA part (light blue) and a flexible and attractive ss-DNA part (orange). (b) The overlap of the ss-DNA shell represents DNA hybridization.

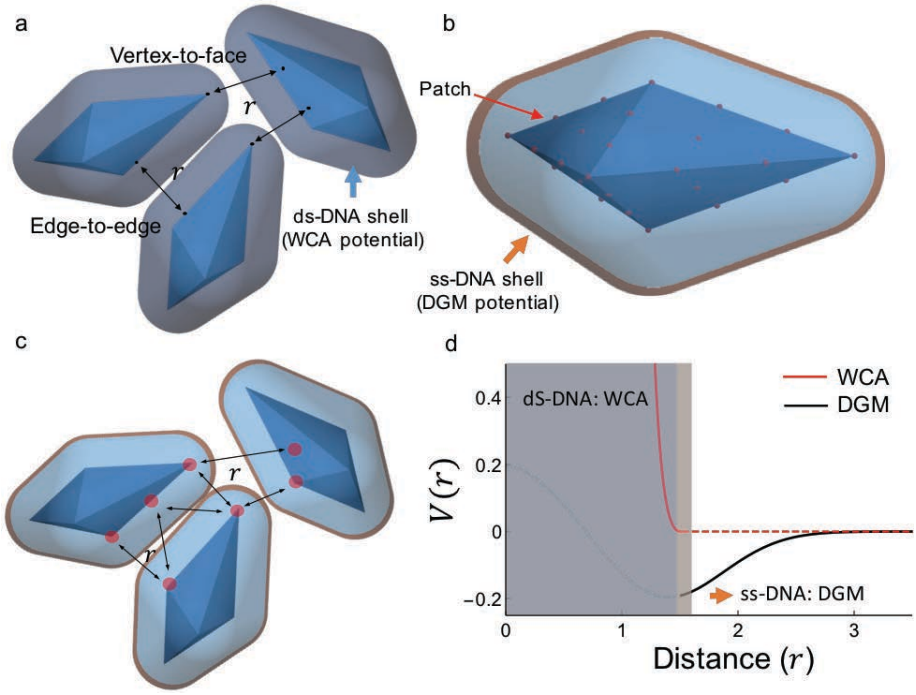


Figure 5.4: Schematic representation of interactions in the minimal model. (a) The non-attractive and rigid ds-DNA shell (light blue shell) is modeled in DEM with the WCA pair potential. (b) Attractive force from DNA hybridization of ss-DNA are included with patches distributed along the edges and faces of TBP cores. Coordinates of the patches are listed in Table 5.3. (c) Each patch interacts with all other patches, resulting in a uniformly attractive ss-DNA shell (orange shell) around the ds-DNA shell. (d) The WCA dominates in the blue region (ds-DNA part). The orange region contains the maximum attraction at a distance where DNA hybridization occurs.

Table 5.1: Parameters for the inter-PAE potentials.

DNA	TBP core edge length	WCA		DGM			Annealing temperature	Observed crystal phase
		ε_{WCA}	σ_{WCA}	ε_{DGM}	σ_{DGM}	r_0		
$n = 3$‡	3.0	0.20	1.08	0.20	0.50	1.18	2.80 – 2.50	Cla I / Cla II / Cla IV
$n = 4$‡	3.0	0.20	1.36	0.20	0.50	1.46	2.80 – 2.50	Cla II / Cla IV
$n = 5$	3.0	0.20	1.62	0.20	0.50	1.72	2.80 – 2.50	Cla II / Cla IV
$n = 8$†	3.0	0.20	2.46	0.20	0.50	2.56	2.70 – 2.40	Cla II

‡ Crystal phases are partially observed.

† $n = 8$ DNA linkers are tested only in simulation.

At high temperature, the model PAEs remained in a disordered fluid phase, but they aggregated at low temperature because of DNA hybridization and eventually ordered into crystals given sufficient time (Figure 5.5) (*118*, *119*). A strong tendency for alignment and face-to-face contact to maximize hybridization was observed, in agreement with experiment. This tendency brought the two polar tips of the TBPs together as well as the three equatorial tips of the TBPs (Figure 5.1B). We identified aspects of crystallographic order that developed spontaneously by visualizing the PAEs (Figure 5.6B and 5.6C). A geometric analysis revealed that the assembled structures contain clusters resembling stellated polyhedra (Figure 5.6D). Within a cluster, one of the three equatorial tips of each TBP pointed inward toward the cluster center. The remaining two equatorial tips joined to form the outer (stellated) points of the cluster, and the polar tips formed the inner points of the cluster surface (Figure 5.6E).

We found four cluster types, all stellated polyhedra, in the simulation data. Thirty PAEs formed a great dodecahedron (a nonconvex polyhedron, cluster A), and 36, 39, and 42 PAEs arrange into the lower-symmetry configurations clusters B, C, and D, respectively (Figure 5.6D). By relying only on the inner points corresponding to polar tips on the cluster surface, clusters A, B, C, and D were mapped onto a pentagonal dodecahedron (cage A, 5^{12}), tetrakaidecahedron (cage B, $5^{12}6^2$), pentakaidecahedron (cage C, $5^{12}6^3$), and hexakaidecahedron (cage D, $5^{12}6^4$), respectively (Figure 5.6E and Figure 5.8). These polyhedra are well known as clathrate cages.

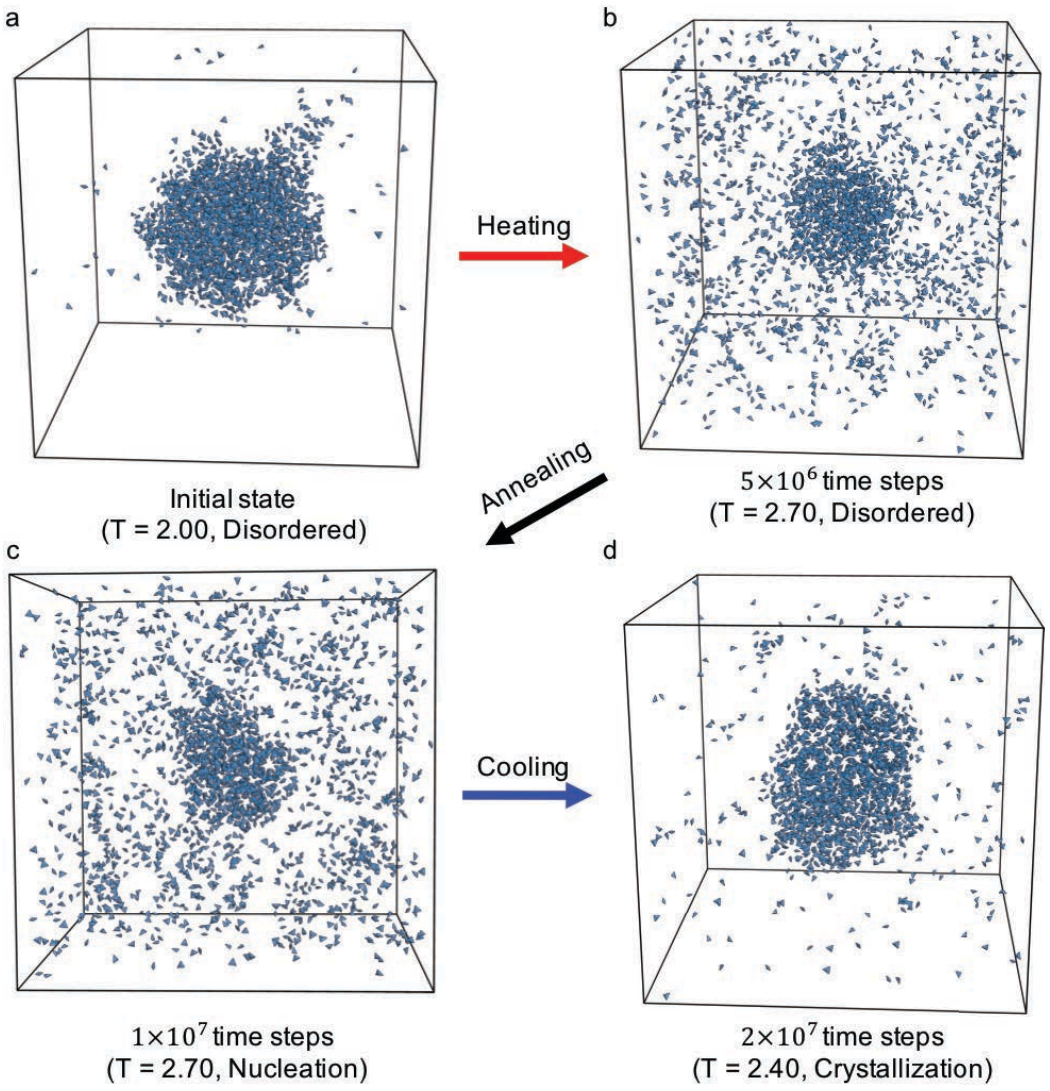


Figure 5.5: Simulation snapshots with D8-DNA (103.2 nm length) ligands for each stage of clathrate crystal growth. PAEs are visualized without DNA shell (only TBP cores) to allow identifying local order more clearly. (a) Shortly after initialization from a disordered spherical cluster. (b) Heating up the cluster close to melting to relax particles. (c) Keeping the high temperature for a while until nucleation occurs. (d) Cooling down slowly to grow the crystal.

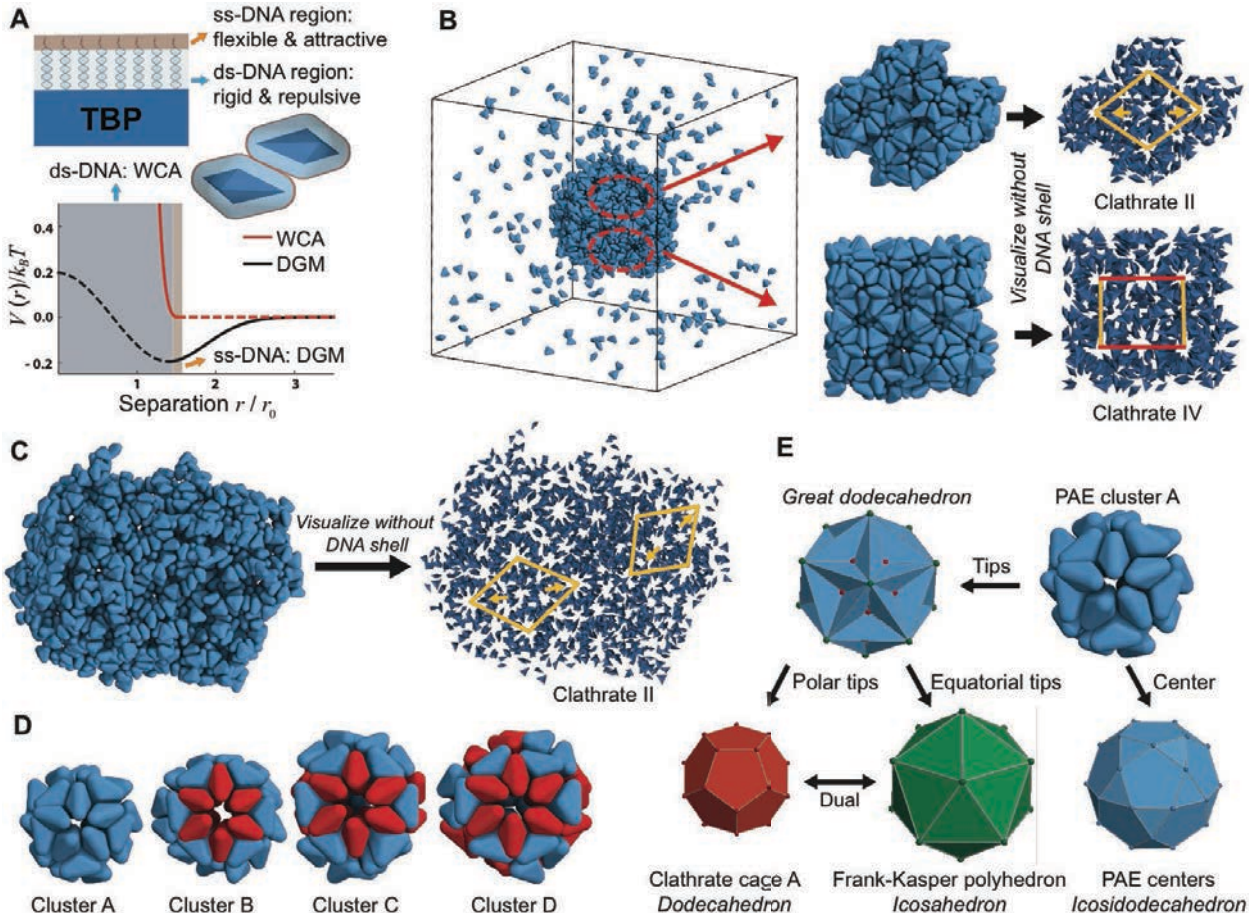


Figure 5.6: Model and simulation of DNA-tethered nanocrystals. (A) The nanocrystal core (triangular bipyramid, TBP) is surrounded by a wide shell of double-stranded DNA (ds-DNA) that terminates in a narrow shell of single-stranded DNA (ss-DNA). The interaction of PAEs is captured by an effective pair potential consisting of a Weeks-Chandler-Anderson (WCA) repulsion upon shell overlap plus a double-Gaussian model (DGM) attraction representing DNA hybridization of the ss-DNA. (B) Nanocrystals with DNA ligands containing five block segments (68.7 nm length) cluster together in a simulation snapshot and spontaneously order. Without the DNA shell local motifs of clathrates II and IV structure are identified. (C) Simulations using DNA ligands containing eight block segments (103.2 nm length) show exclusively clathrate II. (D) Clathrates are built by four types of clusters. Particles in sixfold rings are colored in red. (E) Relation of geometric frameworks derived from PAE cluster A. The tips of the cluster form a great dodecahedron. The cluster can be mapped onto the (pentagonal) dodecahedron by connecting polar tips. It can also be mapped onto the icosahedron, which is the dual of the dodecahedron and the Frank-Kasper polyhedron with coordination number 12 by connecting equatorial tips. Connecting PAE centers defines an icosidodecahedron. The same principles can be applied to other PAE cluster types (Figure 5.8).

The four cluster types can be arranged into three crystals, clathrates I, II, and IV, following literature classification (Figure 5.7) (*118*). We briefly discuss each of the three clathrate types by comparing TEM images and our structure model. Among hundreds of images, we selected for discussion those that are aligned with their main crystallographic axes. Clathrate I (also known as CS-I or sI) is a cubic phase that is uniquely characterized by square tiles when projected along a fourfold axis with hexagonal rings at their vertices (Figure 5.7A-5.7C). Clathrate IV (HS-I_Q) is a hexagonal phase that corresponds to triangle tiles when projected along a sixfold axis with hexagonal rings at their vertices (Figure 5.7D-5.7F). Finally, clathrate II (CS-II or sII) is another cubic phase that has a fundamental rhombohedral building block. When viewed along a twofold axis, pentagonal rings arrange into rhombs (Figure 5.7G-5.7I). Additional evidence is provided in *Section 5.4*, where we characterized images not aligned with main crystallographic axes (Figures 5.9 to 5.12). Using this classification, we identified the order observed in our simulations (Figure 5.6B) as a mix of clathrates II and IV. For thicker DNA shells, exclusively clathrate II is observed (Figure 5.6C).

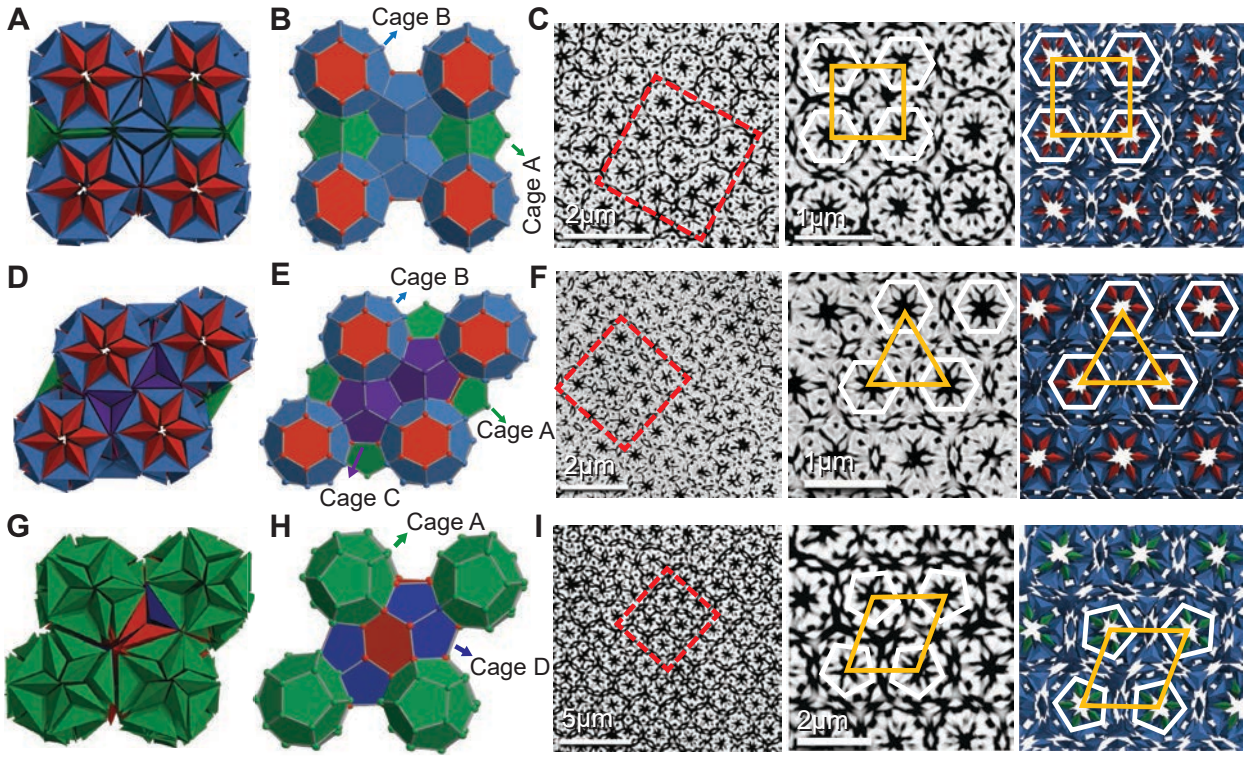


Figure 5.7: Identification of the three basic clathrate crystal structures. In the experimental data, we observe crystals analogous to (A-C) clathrate I oriented along [100], (D-F) clathrate IV oriented along [0001], and (G-I) clathrate II oriented along [110]. Each row shows the construction of a unit cell. (A, D, and G) Non-rounded TBPs highlight the local geometry. (B, E, and H) Connecting TBP polar tips reveals the clathrate cage representation. (C, F, and I) Comparison of electron microscopy images (left), zoom-ins of the red areas (middle), and TBP cores in the structure model (right). Pentagonal rings and hexagonal rings are indicated as white overlays. Characteristic structural features seen in projection along high-symmetry axis are outlined as orange overlays.

We hypothesize that the relative stability of the clathrate crystals can be estimated from geometric principles and the complementary contact model (*13*). Clathrate I has the most similar clusters (cluster composition AB_3) but also the lowest frequency of clusters with icosahedral symmetry (cluster A). Overall, it requires the least deformation of the DNA shell. Thus, we expect clathrate I to be the dominant phase for thin DNA shells. Clathrate IV has a larger range of cluster sizes ($A_3B_2C_2$) and provides TBPs more wiggle room, which is an advantage for DNA shells of intermediate thickness. Clathrate II has the largest range of cluster sizes (A_2D) including the most

open cluster (cluster D). It requires the strongest deformation of the DNA shell. Our simulation results are consistent with the hypothesis that clathrate II is the dominant phase for thick, flexible DNA shells.

5.3 Conclusions

Depending upon the application, the clathrate colloidal crystals we observed may be described equally well using the language of different material geometries (Figure 5.6E). For example, the equatorial tips define a tetrahedral (Frank-Kasper (*120*)) network, and the TBP centroids are located in a fashion analogous to the oxygen positions in the recently discovered chibaite mineral network (*119*). The cavities in the center of the clusters might have applications for host-guest recognition applications (e.g. proteins or virus) at the mesoscale. In addition, they may allow access to multiple properties either synchronously or asynchronously, yielding materials behavior not yet achievable in simpler colloidal crystals. Finally, other complex crystals predicted by simulation (*4*) but not yet realized should now be possible through the use of the complementary contact model and judicious combination of particle shape and DNA linker.

5.4 Method and supporting information

5.4.1 Discrete Element Method Molecular Dynamics

Molecular dynamics simulations with the discrete element method (*117*) (DEM) are performed to simulate the self-assembly behavior of DNA-modified TBPs (or programmable atom equivalents, PAEs) (*101*). The DEM method for simulations of nanoscale anisotropic particles is implemented in the HOOMD-blue simulation package (*9*, *55*), which is available as open source and may be downloaded at https://github.com/glotzerlab/hoomd-blue. HOOMD-blue runs on central processing units (CPUs) as well as on graphics processing units (GPUs). For our

simulations, GPU runs achieve a speed-up of up to 75 times compared to single CPU runs. Simulations are further accelerated by performing parallel runs using MPI domain decomposition on multiple NVIDIA Tesla K20X GPUs.

A minimal model allows simulating thousands of PAEs within a reasonable timescale. The Weeks-Chandler-Andersen (WCA) (*121*) potential and the double Gaussian model (DGM) (*122*) reproduce the influence of the DNA shell surrounding a TBP core. The WCA potential represents the structural property of the double-strand DNA shell, and the DGM models the self-complementary DNA hybridization of single-strand DNA shell.

We perform DEM simulations with between $N = 2065$ and $N = 2274$ TBP particles, which is large enough to form a sufficient number of clathrate unit cell and small enough to still allow crystallization within a reasonable timescale. Typical simulation times are 290 hours on 8 GPUs for 2×10^7 DEM time steps. We used periodic boundary conditions and simulated for a total of up to 2.5×10^7 time steps in individual runs. All simulations are performed in the NVT ensemble at packing fraction 3%, which is low enough to minimize the influence of boundary effects and high enough to avoid an unnecessarily large vapor phase around the particles. The temperature is varied across a region including the melting temperature of the clathrate colloidal crystals to find a temperature value that maximizes assembly propensity. TBP particles spontaneously cluster below the melting temperature without external pressure because the DGM potential is attractive. This allows us to neglect pressure effects and thus to focus solely on the role of DNA hybridization and temperature.

Each simulation follows four steps that are chosen to closely emulate the experimental assembly protocol: (i) Disordered particle configurations within a spherical cluster are initialized at a low enough temperature $T = 2.0$ (dimensionless units) to keep the cluster stable. (ii) The system is heated up instantly to $T = 2.8$, which is close to but below the melting temperature $T_{\mathrm{melt}} = 3.0$ and high enough for the cluster to relax particles efficiently. (iii) The temperature is kept constant until nucleation occurs. (iv) The nucleated TBP cluster is slowly cooled down to $T = 2.4$ in stages of $\Delta T = 0.1$ over 3×10^6 time steps for each stage to grow the crystal and gradually improve structural order.

5.4.2 Minimal Model for DNA-Tethered Nanocrystals

Particles (PAEs) were modeled by a hard TBP core with the vertices listed in Table 5.3, which has the three-fold symmetry axis aligned along the *z*-coordinate axis. [We note the possibility of an alternative, integer representation of the TBP coordinates that is obtained by aligning the two polar tips along a three-fold axis, $v_{\mathrm{p1}} = (1,1,1)$ and $v_{\mathrm{p2}} = -(1,1,1)$, and the three equatorial tips perpendicular to this axis, $v_{\mathrm{e1}} = (2,2,-4)$, $v_{\mathrm{e2}} = (2,-4,2)$, $v_{\mathrm{e3}} = (-4,2,2)$.] The self-assembly behavior of TBP PAEs is simulated with a minimal model. The model assumes three structural properties of DNA-TBPs: (i) the double-strand DNA (ds-DNA) shell is rigid and non-attractive; (ii) the single-strand DNA (ss-DNA) shell is flexible and short-range attractive; (iii) the DNA shell causes an effective rounding of the hard TBP core. Schematic diagrams of the structure model are shown in Figure 5.3.

The non-attractive and rigid ds-DNA shell is modeled using contact interactions applied to vertex-to-face and edge-to-edge pairs of TBP cores at nearest contacts (Figure 5.4a) as standard

for the DEM framework applied to polyhedral particles. For each contact, a steep and purely repulsive WCA potential (*121*) is applied to prevent overlap of the ds-DNA shells,

$$V_{\mathrm{WCA}}(r) = \begin{cases} 4\varepsilon_{\mathrm{WCA}}\left[\left(\frac{\sigma_{\mathrm{WCA}}}{r}\right)^{12} - \left(\frac{\sigma_{\mathrm{WCA}}}{r}\right)^{6}\right] - 4\varepsilon_{\mathrm{WCA}}\left[\left(\frac{\sigma_{\mathrm{WCA}}}{r_{\mathrm{cut}}}\right)^{12} - \left(\frac{\sigma_{\mathrm{WCA}}}{r_{\mathrm{cut}}}\right)^{6}\right], & r < r_{\mathrm{cut}} \\ 0 \quad , & r \geq r_{\mathrm{cut}} \end{cases}$$

where r is the contact distance, $\varepsilon_{\mathrm{WCA}}$ is the interaction energy between contact points (*117*), and $r_{\mathrm{cut}} = 2^{1/6}\sigma_{\mathrm{WCA}}$ (blue region and red line in Figure 5.4d). The rounding radius $\sigma_{\mathrm{WCA}}/2$ is set to be the length of ds-DNA, which is parameterized from the experiment. The ds-DNA shell has a rounding effect on the TBP shape because the potential applied for the contact interaction is isotropic.

The attractive force from DNA hybridization of ss-DNA is expressed using a number of attractive patches on TBP cores (Figure 5.4b and 5.4c), which interact with the DGM potential (*122*)

$$V_{\mathrm{DGM}}(r) = \varepsilon_{\mathrm{DGM}} \exp\left[-\frac{1}{2}\left(\frac{r}{\sigma_{\mathrm{DGM}}}\right)^{2}\right] - \varepsilon_{\mathrm{DGM}} \exp\left[-\frac{1}{2}\left(\frac{r - r_0}{\sigma_{\mathrm{DGM}}}\right)^{2}\right]$$

where r is the distance between patch centers, r_0 denotes the distance where DNA hybridization occurs, $\varepsilon_{\mathrm{DGM}}$ is the interaction energy between patches, and σ_{DGM} is the width of the attractive well (Figure 5.4d). Unlike the contact interaction in DEM, the patches between TBP cores interact with one another regardless of their position (Figure 5.4c), resulting in a uniformly attractive ss-DNA shell around the ds-DNA shell. The combined effect of the repulsive and attractive interactions is achieved by summation (Figure 5.4d). The ds-DNA shell, where the WCA potential

dominates, has a steep and repulsive interaction, while the ss-DNA shell shows a slower varying and attractive interaction from the DGM potential. Parameter values are listed in Table 5.1.

5.4.3 Local Order: Clusters and Cages

Following the simulation stages described in *Section 5.4.1* (initialization – heating – annealing – cooling), the first feature we could observe in the simulation results was a strong tendency to maximize face-to-face contacts between PAEs. This tendency results in the spontaneous formation of four types of cluster (Figure 5.8a). Cluster A, which is the most common cluster, is composed of thirty PAEs forming a five-fold stellate features, while others (B, C, and D) have two, three and four additional six-fold stellated features, respectively. Each cluster can be mapped on a polyhedron by connecting PAE centers. Each such polyhedron is characterized by the number of triangles, pentagons, and hexagons. For instance, a polyhedron denoted by $3^n5^m6^l$ indicates the structure with *n* triangles, *m* pentagons and *l* hexagons. The cluster A, B, C and D correspond to the icosidodecahedron $3^{20}5^{12}$, and the polyhedra $3^{24}5^{12}6^2$, $3^{26}5^{12}6^3$, and $3^{28}5^{12}6^4$, respectively (Figure 5.8b), ordered in sequence of increasing cluster volume.

The four PAE clusters can also be mapped onto four clathrate cages by connecting the locations of the shared polar tips (Figure 5.8c) resulting in the pentagonal dodecahedron (5^{12}), tetrakaidecahedron ($5^{12}6^2$), pentakaidecahedron ($5^{12}6^3$), and hexakaidecahedron ($5^{12}6^4$), respectively. These clathrate cages combine to form various clathrate crystal structures depending on their combination (*118*). Furthermore, PAE clusters can be mapped onto Frank-Kasper (FK) polyhedra (*123*) by connecting the locations of their shared equatorial tips (Figure 5.8d). FK polyhedra are relevant for tetrahedral close packing and can be labeled by coordination number

(CN). PAE cluster (from A to D) then map onto FK CN-12, CN-13, CN-14, and CN-15 polyhedra, respectively. FK polyhedra and clathrate cages are dual to another.

5.4.4 Clathrate Crystal Structures

Although PAE clusters are locally observed at a very early stage in our simulation before the annealing step (Figure 5.5b), most of the clusters are imperfect and arranged in a disorderly fashion. We therefore keep the temperature close to but below the melting temperature for an extended period of time to give the clusters enough time to develop crystallographic order. After 10 million time steps, nucleation is observed (Figure 5.5c), and we slowly cool down the system (Figure 5.5d).

The crystal structure formed by the PAEs can be identified more clearly by connecting polar tips, because it integrates multiple PAEs sharing the polar tips into a single particle and each TBP cluster is uniquely related to a clathrate cage. This reveals that the PAEs with n = 5 DNA (68.7 nm) form a mixed clathrate structure with type II and IV structure in agreement with experimental observation, as indicated by the rectangular and rhombic tiling of five-fold rings we observe. The n = 8 DNA (103.2nm) system, which is tested only in simulation, forms exclusively clathrate II with a rhombic tiling of five-fold rings.

Table 5.2: DNA sequences used for the linker design. The number of block segments n controls the total length (number of bases) of the DNA linker.

Component	Base sequence	Number of bases
anchor Strand	$^{5'}$AAG AAT TTA TAA GCA GAA-A_{10}-$C_3SH^{3'}$	28
linker $n = 0$	$^{5'}$TTC TGC TTA TAA ATT CTT-A-GCGC$^{3'}$	23
linker $n \geq 1$	$^{5'}$TTC TGC TTA TAA ATT CTT-A-(block sequence)$_n$-GCGC$^{3'}$	23 + 41 n
block sequence	$^{5'}$-TTT TTT TTT TTT ACT GAG CAG CAC TGA TTT TTT TTT TTT T-A-$^{3'}$	41

duplexer strand	$^{5'}$AAA AAA AAA AAA ATC AGT GCT GCT CAG TAA AAA AAA AAA A$^{3'}$	40

Table 5.3: Estimated DNA shell thickness for a given number of block segments.

Number of block segments	Total number of bases (base pairs)	Estimated thickness of DNA shell
0	33 (18)	11.2 nm
1	74 (58)	22.7 nm
2	115 (98)	34.2 nm
3	156 (138)	45.7 nm
4	197 (178)	57.2 nm
5	238 (218)	68.7 nm

Table 5.4: Coordinates of TBP core vertices and patch positions used for DEM simulations.

Vertices of the rigid core (TBP)											
	x	***y***	***z***		***x***	***y***	***z***		***x***	***y***	***z***
v_{e1}	1.5	0.866	0	v_{e2}	−1.5	0.866	0	v_{e3}	0	−1.732	0
v_{p1}	0	0	0.612	v_{p2}	0	0	−0.612				
Patch positions on the TBP											
1	0	0.577	0.204	**11**	−1	0.577	0.204	**21**	1	0.577	0.204
2	−0.5	−0.289	0.204	**12**	−0.5	0.289	0.408	**22**	0.5	0.289	0.408
3	0.5	−0.289	0.204	**13**	−1	0.577	−0.204	**23**	1	0	0
4	0	0.577	−0.204	**14**	−0.5	0.289	−0.408	**24**	0.5	−0.866	0
5	−0.5	−0.289	−0.204	**15**	0	−1.155	−0.204	**25**	1.5	0.866	0
6	0.5	−0.289	−0.204	**16**	0	−0.577	−0.408	**26**	−1.5	0.866	0
7	0.5	0.866	0	**17**	0	−1.155	0.204	**27**	0	−1.732	0
8	−0.5	0.866	0	**18**	0	−0.577	0.408	**28**	0	0	0.612
9	−1	0	0	**19**	1	0.577	−0.204	**29**	0	0	−0.612
10	−0.5	−0.866	0	**20**	0.5	0.289	−0.408				

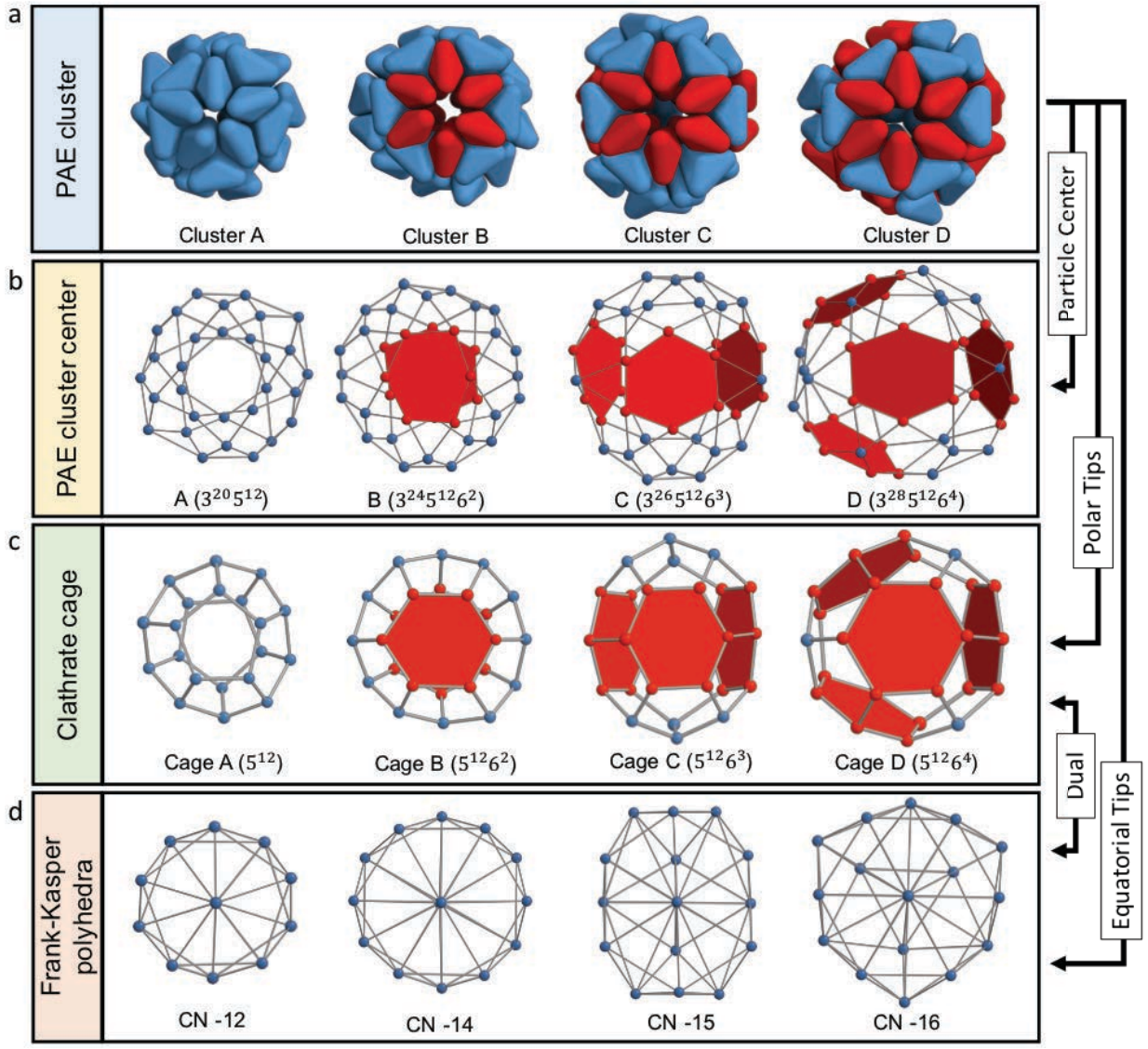


Figure 5.8: Relationship of geometric frameworks derived from PAE clusters. (a) Four PAE clusters spontaneously assemble in simulation, cluster A (30 PAEs), B (36 PAEs), C (39 PAEs), and D (42 PAEs). All cluster contain 12 five-fold stellated features (blue), and cluster B, C and D have additional 2, 3, and 4 six-fold stellated features (red), respectively. (b) Connecting PAE centers gives four polyhedra. (c) Clathrate cages result by connecting polar tips of PAEs. (d) Connecting equatorial tips defines four Frank-Kasper coordination (CN) polyhedra.

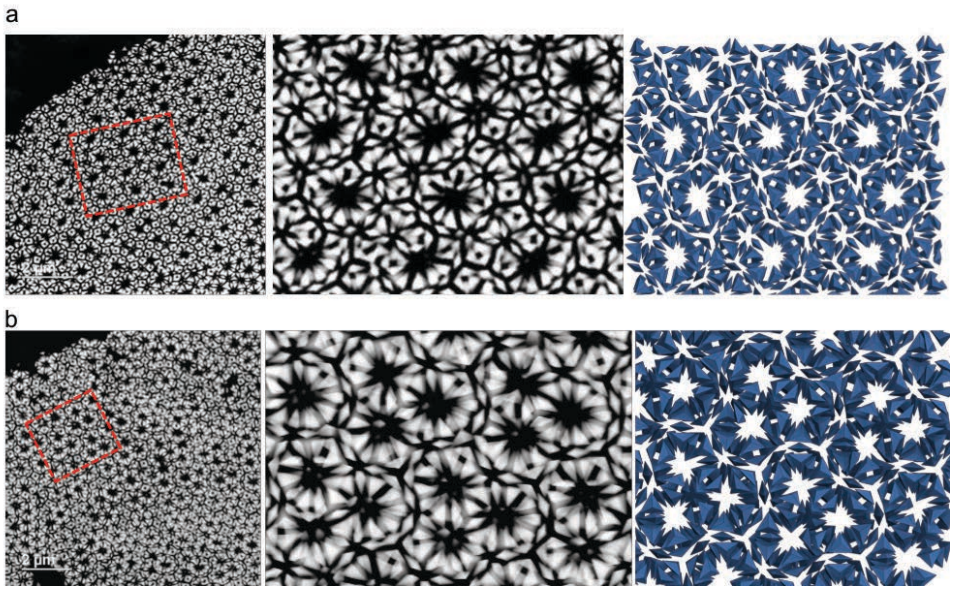


Figure 5.9: Experimental examples for the clathrate I structure. (a, b) Electron microscopy images of clathrate I captured from a direction not aligned with a main crystallographic axes (left) and zoom-ins (middle). Extracted images from the constructed structure model of clathrate I for comparison (right).

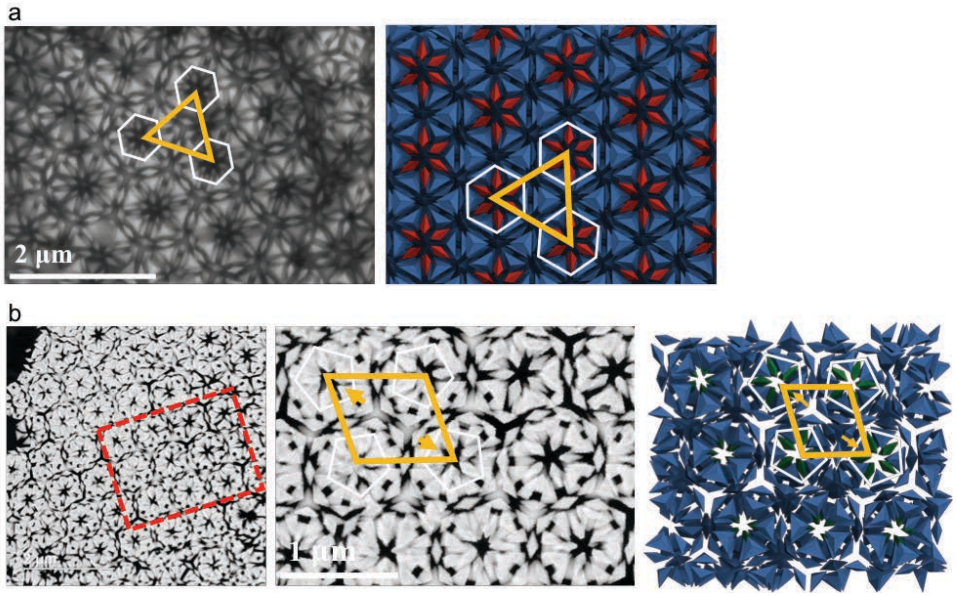


Figure 5.10: Experimental examples for the clathrate II structure. (a, b) Electron microscopy images of clathrate II projected along three-fold symmetry and two-fold symmetry axes (left). Extracted images from the constructed structure model of clathrate II for comparison (right).

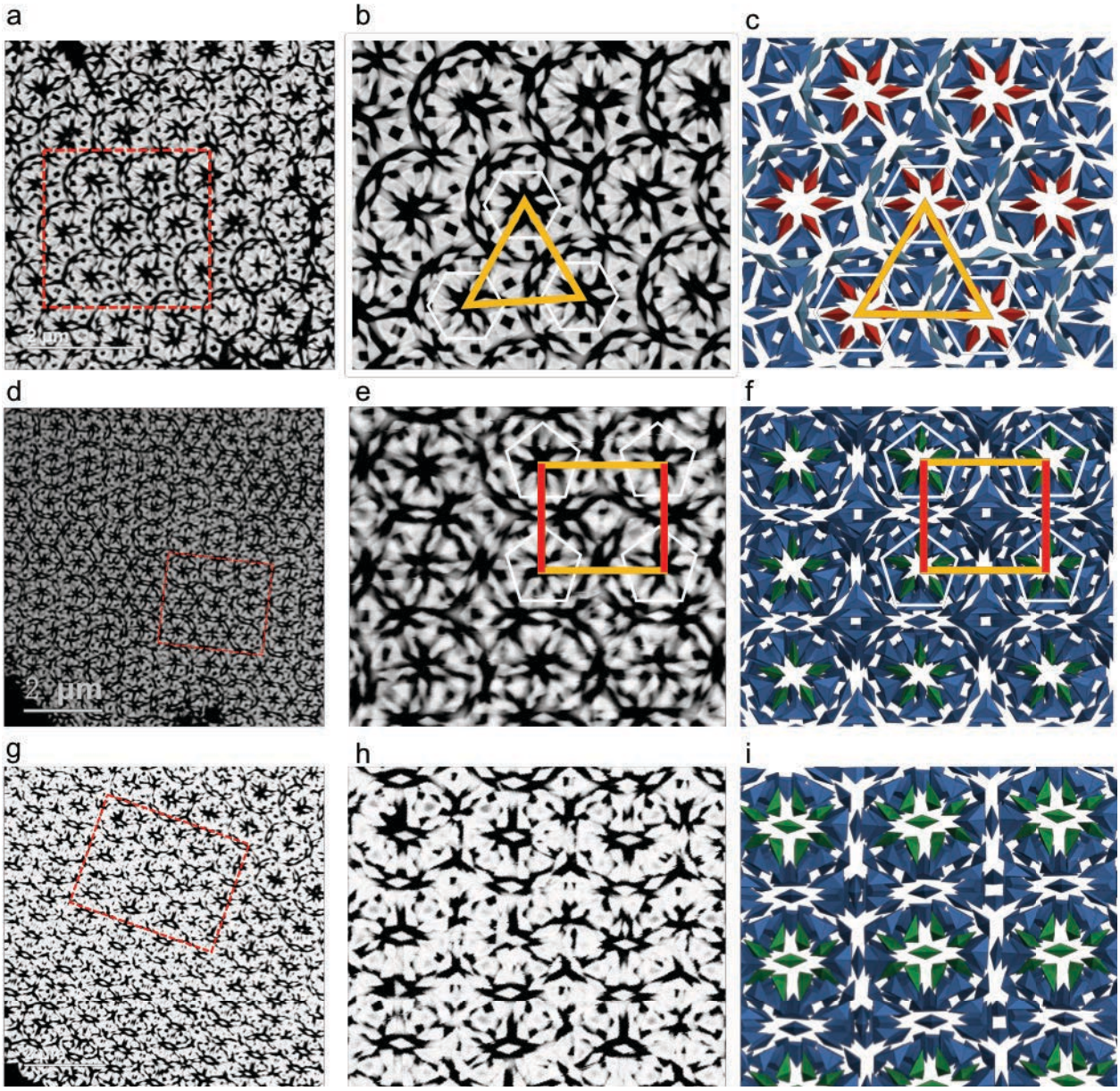


Figure 5.11: Experimental examples for the clathrate IV structure. Electron microscopy images of clathrate IV projected along (a) three-fold, (d) two-fold, and (g) a minor crystallographic axes. (b,e,h) Zoom-ins and (c,f,i) the extracted images from the constructed model of clathrate IV for comparison.

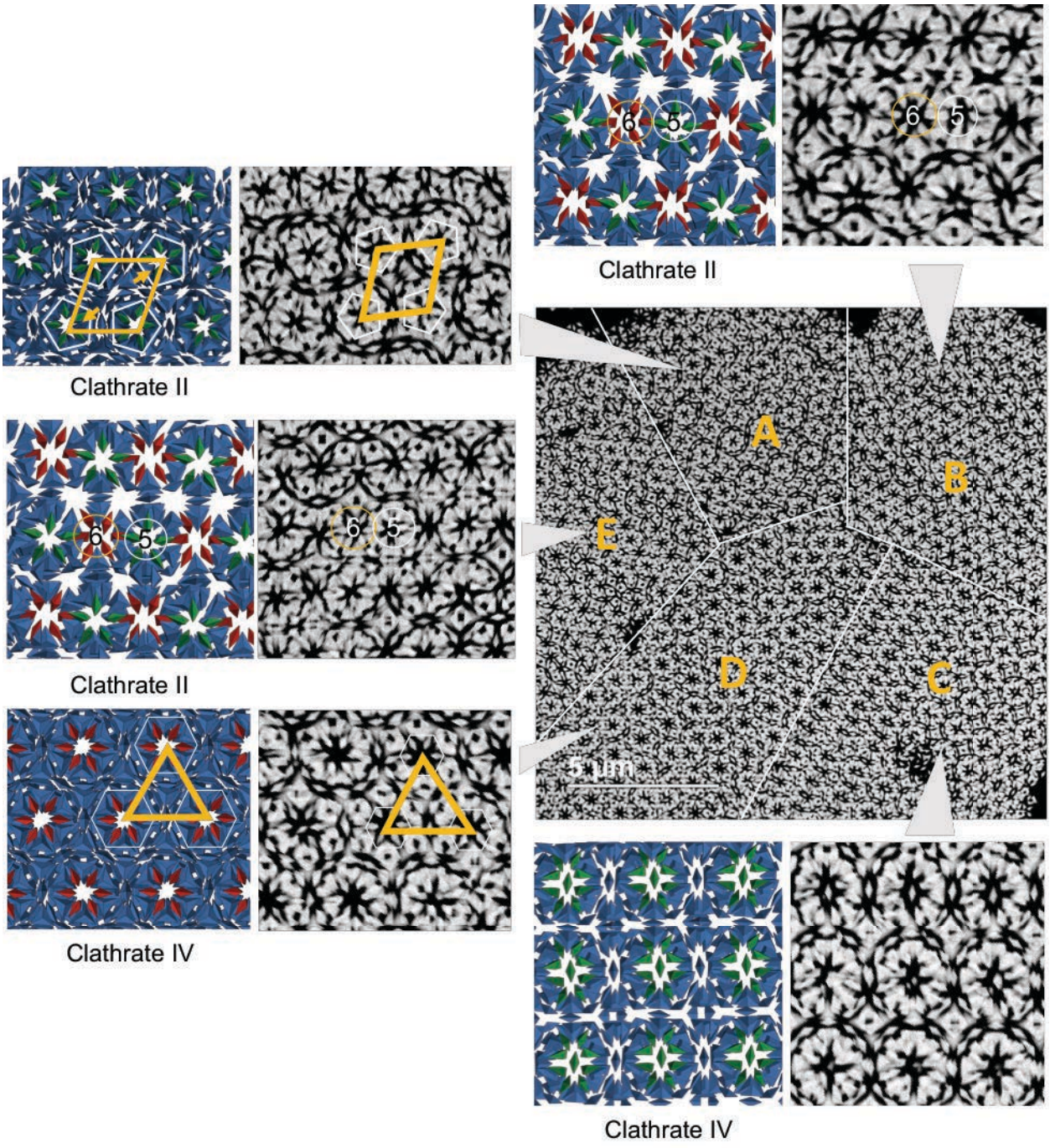


Figure 5.12: Electron microscopy image of a superlattice with multiple coexisting domains. The sample and the image in the middle, right is identical with Figure 5.1g in the main text. Zoom-ins of each domain (right side for each domain) are compared with images extracted from the structure model (left side). Clathrate II (A, B, E) and clathrate IV (C, D) coexist in different domains in this sample.

CHAPTER 6

Valency Control in DNA-Programmable Colloidal Crystals through Electron Equivalents

This chapter is reproduced from a publication in preparation, authored by Shunzhi Wang[†], Sangmin Lee[†], Jingshan S. Du[†], Benjamin E. Partridge, Ho Fung Cheng, Wenjie Zhou, Vinayak P. Dravid, Byeongdu Lee[*], Sharon C. Glotzer[*] and Chad A. Mirkin[*] ([†]equal contribution, *corresponding authors). My contribution to this work included MD simulation model development of DNA-programmable electron equivalent (EE) colloids, analysis of spatial distributions and dynamic characters of EEs, identifying triple double-gyroid structure, and confirming dynamic redistribution of EEs by temperature in computer simulations.

6.1 Introduction

The concept of valence electrons is central to chemical bonding theory, which defines how atoms can combine to form well-defined polyhedral geometries in molecules and crystalline solids (*124*, *125*). Despite their prevalence in nature, low symmetry structures are not commonly found in colloidal crystals, regardless of the fact that particles are often analogized with atoms in such structures (*61*, *126*, *127*). The absence of such low symmetry structures is because colloidal particles typically assemble through surface ligand-mediated interactions that favor the formation of closely-packed and highly symmetric phases, whereas in atomic or ionic crystals, various

coordination geometries exist due to electronic interactions and atomic valency. Although progress has been made in using anisotropic particles and polyhedral templates to direct particle "bonding" interactions and access low symmetry structures (*12*, *59*, *105*, *128–132*), these approaches are still limited with respect to the number and types of building blocks that can be synthesized and, as a result, the types of crystals that can be accessed. In addition, generating low-symmetry phases from isotropic building blocks is particularly challenging. A strategy to program the formation of colloidal crystals from particles with the equivalent of valence electrons that define possible coordination geometries would enable access to broad unexplored regions of phase space.

Colloidal crystal engineering with DNA has emerged as a powerful approach to control nanoparticle superlattice symmetry, crystal habit, and lattice parameters, regardless of particle composition (*6*, *95*, *133–135*). With this methodology, DNA-modified nanoparticles act as programmable atom equivalents (PAEs) that arrange themselves in three-dimensional space via a well-defined set of rules described by the complementary contact model (CCM) (*13*). Recently, Girard *et al*. (*136*) made the remarkable finding that when reduced in size and DNA grafting density, small PAEs can behave as classical electron equivalents (EEs), allowing them to roam through and stabilize lattices defined by larger PAEs (*136*, *137*). However, it remains unclear whether these EEs can be programmed to dictate the symmetry of PAE lattices. If possible, new phases would become accessible and designable from isotropic building blocks by taking advantage of such particle valency.

Herein we describe an approach to reducing lattice symmetry in colloidal crystals by taking advantage of anisotropic distributions of mobile "valence" EEs around PAEs concomitant with crystal lattice formation (Figure 6.1A). In particular, we hypothesized that by controlling the relative densities of DNA grafted to the particle surfaces, we could establish a type of valency

defined by EE distribution around PAEs (Figure 6.1B) and that the symmetry breaking would be determined by: 1) the stoichiometric ratio of PAEs to EEs in a given lattice; 2) the local geometric configurations of the interstitial EEs, which are defined by the polyhedral arrangement of PAEs around them (Figure 6.1A); and 3) the delocalization and diffusion of EEs within the lattice (Figure 6.1C). Through a combination of small-angle X-ray scattering (SAXS), electron microscopy, and molecular dynamics (MD) simulations, we show that the assembly of nine equilibrium phases is driven by the maximization of DNA hybridization. The flexible DNA shells and delocalization of EEs are responsible for stabilizing distorted EE local configurations and nonequivalent sites in certain lattices. Lastly, we report two enantiotropic phase transformations that rely on the redistribution of EEs upon thermal treatment, characterized by *in situ* SAXS experiments and well reproduced by MD simulations. These findings are summarized by five observations that reveal the role of EEs in enabling complex colloidal crystals and will further serve as design principles to engineer different phases.

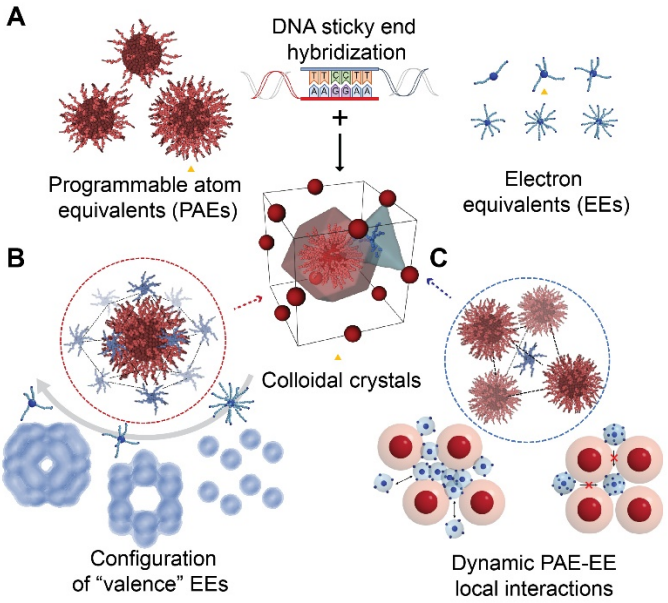


Figure 6.1: Assemblies of PAE-EE colloidal crystals and simulation models. (**A**) Colloidal crystals synthesized from co-assembling PAEs and complementary EEs with different numbers of DNA linkers.

Wigner-Seitz cell representation of the PAE sublattice is highlighted in red, and the local EE configuration is highlighted in blue. (**B**) A zoomed-in view of local PAE coordination environment surrounded by EEs (left) and examples of anisotropic EE spatial distributions around a single PAE. (**C**) A zoomed-in view of local geometric configuration of an EE site, where the arrangement of PAEs and relative EE diffusion is controlled by the strength of PAE-EE interaction.

6.2 Results and discussion

Observation 1: The number of valence EEs per PAE inversely correlates with the EE to PAE DNA sticky end ratio

In a typical experiment, PAEs and EEs were synthesized by modifying 10-nm and 1.4-nm AuNPs with two types of single-stranded, propylthiolated DNA (ssDNA) (Table 6.3), yielding particles with an average of 160 and 15 DNA ligands, respectively (*136*). These ssDNA-modified PAEs and EEs were then duplexed with specific numbers of linker strands complementary to the ssDNA sequence. These linkers contain exposed 6-base sticky end overhangs. The sticky ends on PAEs and EEs are complementary to each other and allow for DNA bond formation via hybridization. Upon mixing in their duplexed form in 0.5 M NaCl solution, PAEs and EEs rapidly assembled into aggregates, which were heated above their melting temperature and then slow-cooled to form crystalline products (*138*).

By varying the number of input linker strands (ranging from 4 to 20 linkers per EE and 30 to 200 linkers per PAE) and EE-PAE stoichiometry, an experimental phase diagram was established where nine distinct crystalline phases emerge from this deceptively simple system, as characterized by SAXS experiments (Figure 6.2A to 6.2I). Due to their small scattering cross section, 1.4-nm AuNPs contribute negligibly to SAXS scattering. Therefore, these structures are defined based upon their PAE arrangement (Table 6.1), where sharp scattering peaks and high

angle resolution suggest large, highly crystalline domains. To further confirm the crystal structures, all samples were immobilized in the solid state *via* silica embedding (*139*), followed by resin encapsulation and ultramicrotome sectioning into thin slices. These cross-sectional samples were imaged, and the crystalline structures were confirmed through scanning transmission electron microscopy (STEM) (Figure 6.2). The formation of specific phases is deliberately controlled by designed and sequence-specific DNA-DNA interactions, which is distinct from classical binary nanoparticle superlattices (*61*). Four of the nine phases [body-centered tetragonal (BCT-1), BCT-2, triple double-gyroid, and high-pressure gallium (HP-Ga)] have not been reported before from isotropic particles (*140*). Among them, BCT-2 and HP-Ga formed when 10-nm AuNPs were duplexed with two types of linkers, one with a sticky end and a second type without a sticky end, to modulate the number of sticky ends on the PAE while maintaining the steric bulk of the PAE DNA shell. For all other phases, a single type of linker, which contains a sticky end, was duplexed onto the PAE.

Table 6.1: Structural information of nine identified PAE-EE assemblies in the fully localized state.

Symmetry	**Pearson symbol**	**Space group**	**Lattice constant (nm)**	**PAE Wyckoff positions**	**EE Wyckoff positions based on W-S analysis**	**Equilibrium PAE:EE stoichiometry***
FCC	cF4	$Fm\bar{3}m$	a = 57	4a	32f	AB_8
BCC	cI2	$Im\bar{3}m$	a = 42	2a	12d	AB_6
BCT-2	tI2	*I4/mmm*	a = 46, c = 42	2a	4d, 8j	AB_6
FK-A15	cP8	$Pm\bar{3}n$	a = 70	2a, 6c	6d, 16i, 24k	A_8B_{46}
BCT-1	tI2	*I4/mmm*	a = 34, c = 63	2a	4d, 4e	AB_4
SH	hP1	*P6/mmm*	a = 31, c = 28	1a	2d	AB_2
Gyroid	cI48	$Ia\bar{3}d$	a = 108	48g	16a, 16b, 48f	$A_{48}B_{80}$
HP-Ga	cI12	$I\bar{4}3d$	a = 72	12a	16c	$A_{12}B_{16}$
SC	cP1	$Pm\bar{3}m$	a = 27	1a	1b	AB

* Assuming each Wyckoff position is occupied by a single EE.

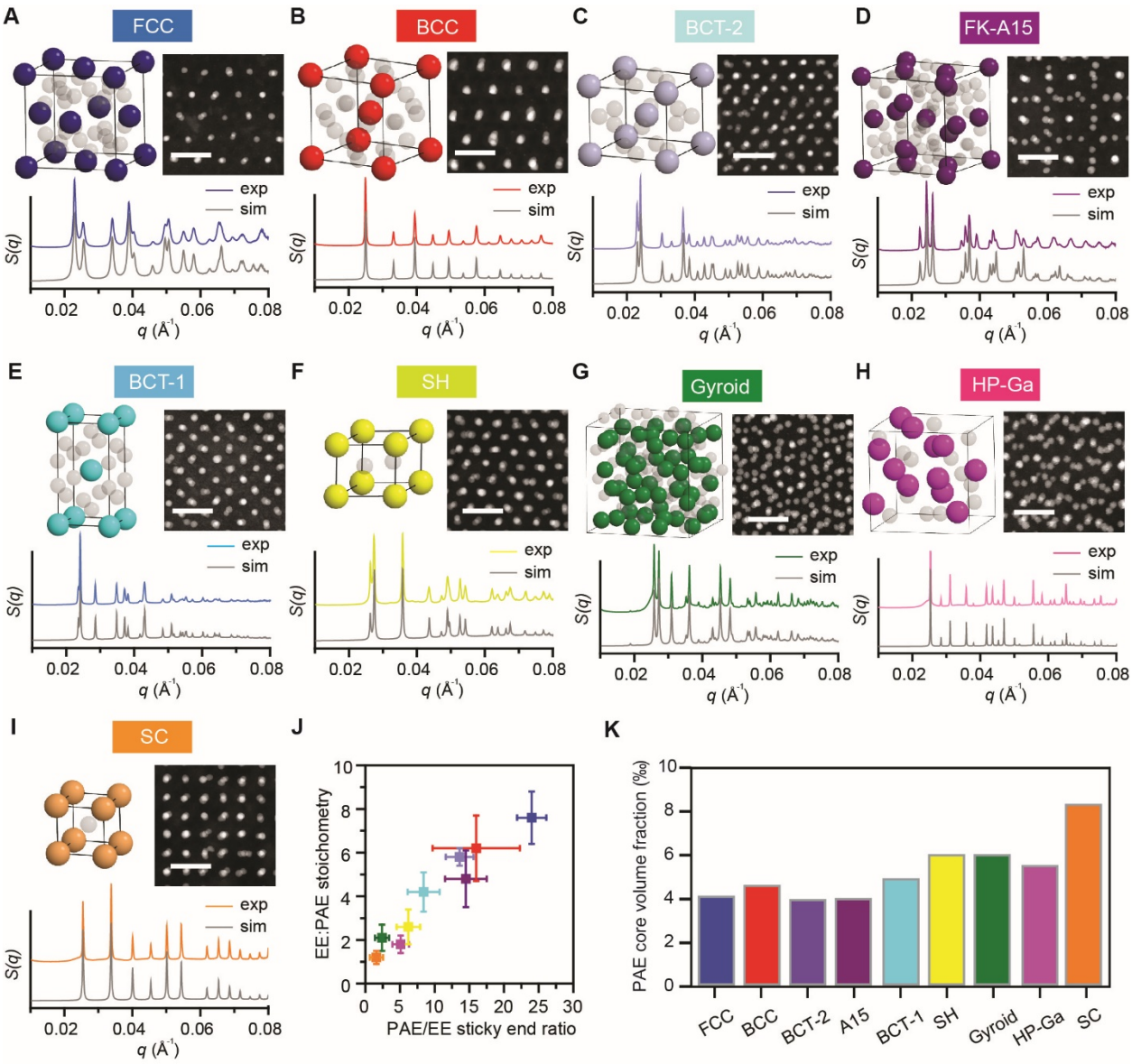


Figure 6.2: Formation of nine distinct PAE-EE assemblies as controlled by DNA-based interaction. (**A-I**) Top row of each panel: unit cell model (PAEs in color, EEs in light grey) and representative cross-sectional low-angle annular dark-field (LAADF) image for each crystal. Bottom row of the panel: structure factor *S*(*q*) obtained from synchrotron-based SAXS experiments (color) and simulated SAXS patterns based on PAE crystals alone (grey). Scale bars are 50 nm for all images. (**J**) A positive correlation is observed between PAE-EE stoichiometry and their surface sticky end density ratio (based on input linkers). (**K**) Packing fraction of PAE AuNP core for each phase following the order of EE:PAE stoichiometry ratio.

To quantify the EE-PAE stoichiometric ratio for each phase, we dissolved pure-phase crystals synthesized with Cy5 dye-labeled EEs in water and then measured the maximum

absorbance of Cy5 at 647 nm and the localized surface plasmon peak of 10-nm Au-PAEs at 520 nm using UV-vis spectroscopy. The input EE-PAE DNA linker and, therefore, sticky end ratio directly correlates with the EE-PAE stoichiometry observed in crystalline phases, providing a means to deliberately access distinct structures (Figure 6.2J). Across all phases, longer PAE-PAE distances as well as lower PAE volume fractions (Figure 6.2K) are observed for phases generated from lower EE-PAE sticky end ratios, providing more spaces to host EEs, which is consistent with an increase in the EE-PAE stoichiometry. Such a correlation can be explained by the maximization of DNA sticky end hybridization, where unpaired PAE-EE sticky ends are unfavorable.

Observation 2: Within a crystalline lattice, EE spatial distribution around a single PAE forms a polyhedral valence shell defined by the surface of a PAE Wigner-Seitz cell, with maximum EE density at its vertices.

Having established that the average number of valence EEs is structure-dependent and can be controlled by the surface density of their DNA sticky ends (Figure 6.2J), we next explored the EE spatial distribution for each phase, which defines the geometric shape of the EE valence shell. We hypothesized that EEs reside – in clusters or singly – at the Wigner-Seitz cell (W-S cell, also known as the Voronoi cell) vertices of the PAE lattices (Figure 6.1A and Table 6.1) because: 1) the W-S cell surface represents the interface at which EEs are equidistant between two or more neighboring PAEs; 2) at W-S cell vertices, one EE can interact with the most PAEs simultaneously to maximize DNA hybridization; and 3) W-S cell vertices define the largest distance between PAEs and EEs, minimizing steric repulsion between unpaired DNA strands and increasing EE configurational entropy. In this vein, the PAE coordination environment is defined by EEs at the W-S cell vertices, and the local geometry of an EE site is conversely defined by the polyhedral arrangement of its nearest PAE neighbors.

To test this W-S cell hypothesis, MD simulation models were built to study the colloidal assembly processes and equilibrium crystal structures *in silico* (Figure 6.3A to 6.3C, and *Section 6.4*) using the HOOMD-Blue particle simulation toolkit (*93*). In this model, PAEs and EEs are represented by rigid spherical cores (diameter: 10 and 1.4 nm) surrounded by soft DNA ligand shells (thickness: ~9 nm), where size parameters were selected to match experimental values (*59*). The finite number of sticky ends on EEs are modeled as spatially uniform anchoring points that can bind to isotropic DNA shells on the PAEs (Figure 6.3A and *Section 6.4*) (*141*). To account for the compressibility of the DNA shells, an overlap index (ω_{PAE} and ω_{EE}) was assigned, which is correlated to the DNA grafting density on PAEs and EEs (Figure 6.3B and 6.3C and *Section 6.4.2*). Specifically, a low DNA grafting density allows for significant spatial overlap with other particles and is represented by a high ω_{EE} value. In contrast, a high DNA grafting density corresponds to EEs that behave like hard spheres with little distortion or overlap (low ω_{EE}). Modeled PAEs and EEs were initialized in a disordered fluid phase with a low pre-assigned PAE packing density ($\phi_{PAE} = \mathrm{N_{PAE}} \times v_{\mathrm{PAE}} \,/\, V_{\mathrm{box}} = 0.05$) and then equilibrated at a temperature (T^*) lower than the melting temperature of the system (T_m^*, with $T^*/T_m^* \sim 0.95$) to induce crystallization (Table 6.2). By simulating crystal formation with different input parameters (ω_{EE}, ω_{PAE}, and the number of anchoring points on EEs), all nine different crystal structures were observed as stable products, as confirmed by diffraction patterns, bond order diagrams, and direct visualization (Figures 6.9 to 6.12). The observed equilibrium phase behavior can be rationalized by calculating the relative stability of different crystals based on their overall packing density, where a higher overall packing density maximizes DNA bond formation (*Section 6.4.3*). This model expands the CCM to include structures such as these by considering the overlap of like particles (via the overlap indices ω_{EE} and ω_{PAE}) in addition to complementary particles. Stacking these packing density plots for each

crystal lattice symmetry at a given set of overlap indices generates a landscape, in which the topmost structure has the highest packing density (lowest potential energy) and is predicted to be the equilibrium phase (Figure 6.8).

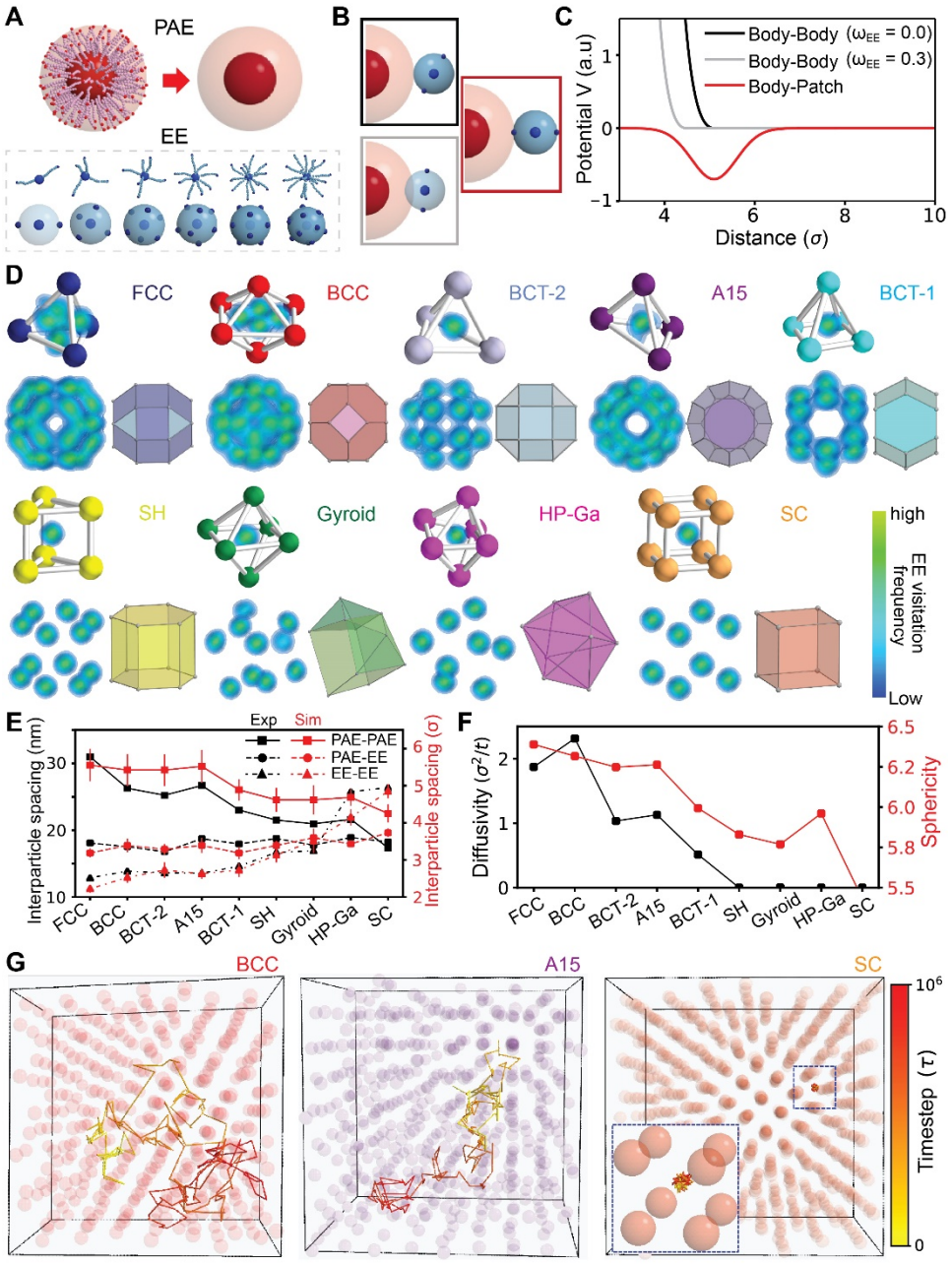


Figure 6.3: MD simulations of the structural configurations and dynamics in equilibrium phases. **(A)** Double-stranded DNA around PAEs and EEs are modeled by isotropic shells surrounding the particle cores. The finite number of sticky ends on EEs was modeled as spatially uniform anchoring points that can bind

onto the isotropic DNA shells on the PAEs, representing DNA hybridization of ssDNA sticky ends. (**B, C**) In three cases modeled in (B), different pair potential functions are assigned for PAE-EE pairs without patch contact nor shell overlap (black), without patch contact but with shell overlap (grey), and with patch contact (red). (**D)** The isosurface plots of EE visitation frequency in a PAE local environment (upper left insets), all EEs surrounding one PAE at the center (lower left insets) and local PAE coordination polyhedra (lower right insets) for the nine different crystals obtained from MD simulation for 10^7 timesteps. (**E**) Interparticle spacing (surface-to-surface) for each phase obtained from experiments (black) and MD simulations (red). Experimental PAE-EE and EE-EE spacings are calculated based on the constructed W-S models. Error bars indicate the standard deviations. (**F**) A positive correlation is observed between EE diffusivity and the sphericity of local PAE coordination polyhedra. (**G**) Trajectories of a single EE in the BCC (left), FK-A15 (center), and SC (right) phases for 10^6 MD timesteps. The trajectory in SC is confined within a single unit cell (inset).

The sequence of symmetries realized experimentally when the EE-PAE sticky end ratio increases is reproduced well in MD simulations by simultaneously increasing the number of anchoring points on EEs and decreasing the magnitude of ω_{EE}. Isosurfaces of EE visitation frequency (Figure 6.3D and Section 6.4) reveal that the distribution of EEs around a PAE has multiple well-defined local maxima, with a geometry that evolves from a 24-coordinated rhombicuboctahedron for the face-centered cubic (FCC) structure to an 8-coordinated cube for the simple cubic (SC) structure (Figure 6.3D). The positions of these local maxima either agree with the PAE W-S cell vertex positions (Table 6.1), or in the cases of FCC and triple double-gyroid phase, correspond to groupings of several vertices in close proximity. Importantly, the simulated spacings between PAE-PAE and PAE-EE pairs (Au surface-to-surface) are consistent with SAXS measurements and the W-S cell analysis (Figure 6.3E). The spacings between PAE cores and neighboring EE sites (derived from W-S cells) are nearly constant for all phases, and they match the length of a pair of DNA linkers hybridized through sticky ends.

Observation 3: The valence shell geometry is determined by the extent of EE delocalization, and defines the degree of metallic or covalent character.

As the number of sticky ends per EE decreases, more "valence" EEs are recruited within a crystal, and the EEs become more delocalized. This transition is evidenced by a decrease in the average spacing between neighboring EE sites (Figure 6.3E) and a more isotropic distribution of EEs around a PAE as seen from isosurfaces of EE visitation frequency. We quantified the sphericity (*142*) ($\psi = \pi^2(6V_p)^{\frac{2}{3}}/A_p$, where V_p and A_p are the W-S cell volume and surface area, respectively) of EEs distributed around a central PAE, which represents the isotropy of the local EE spatial distribution. Structures with higher sphericity (isotropy) better resemble an ideal free-electron Fermi surface. Phases that form at lower EE-PAE sticky end ratios show higher sphericity in MD simulations (depicted by fewer DNA anchor points on EEs) (Figure 6.3F).

These phases also show enhancement of EE diffusion and delocalization (*136*). To study the dynamics of EEs in different crystal structures, we tracked the trajectories of EEs as a function of time in MD simulations of crystals under equilibrium conditions (Figure 6.3G). The diffusivity of EEs ($D = \sigma^2/\tau$, where σ and τ are simulation displacement and timestep, respectively) for each phase was further quantified using the slope of the mean-squared displacement (MSD) of EEs as a function of time (Figure 6.15). EEs within the crystals of the BCT-1, Frank-Kasper (FK)-A15, BCT-2, body-centered cubic (BCC), and FCC phases (low EE-PAE sticky end ratio) are clearly diffusive, with higher diffusivity observed for lower EE DNA grafting density (Figure 6.3G, left and middle panels).

Accordingly, low diffusivity correlates with low sphericity and high EE DNA grafting density. For the SC, simple hexagonal (SH), HP-Ga, and triple double-gyroid phases, which all have a high EE-PAE sticky end ratio, we observed only local thermal vibrations of EEs and no

long-range diffusion (Figure 6.3G, right panel). For these phases, we confirmed that single EEs reside at EE sites (Figure 6.16).

At intermediate sphericity and diffusivity, EEs are partially delocalized, forming clusters at EE sites. The clustering of EEs was determined in simulation by the time-averaged distribution of the number of EEs at each EE site (Figure 6.16). For these phases, multiple EEs can reside at a single EE site because a significant spatial overlap between EEs is allowed (Figure 6.16). These phases – BCT-2, FK-A15 and BCT-1 – represent first examples of colloidal crystals with strongly delocalized EEs.

Based on these results, the nine colloidal crystal PAE-EE phases can be categorized into three groups: 1) "Metallic" phases – BCC and FCC, which exhibit high EE diffusivity and near-spherical local PAE coordination geometries. In this group, a single EE does not have enough sticky ends to stabilize a tetrahedral interstitial site (Figure 6.3D), which instead must be occupied by multiple EEs to form a collective bond, significantly increasing local EE disorder and macroscopic diffusion (*136*). 2) "Tetrahedral" phases – FK-A15, BCT-1 and BCT-2, which exhibit near-tetrahedral EE local configurations but with appreciable EE diffusivity. In this group, EEs are mostly localized in clusters at the vertices of the W-S cells, and can diffuse across multiple unit cells to form strongly delocalized, collective bonds, 3) "Covalent" phases – triple double-gyroid, HP-Ga, SH and SC, which show high EE coordination numbers (≥6), anisotropic PAE coordination geometries, and only localized EE thermal motion around specific sites, due to strong PAE-EE bonding interactions (Figure 6.3D). In the covalent group, single EEs are largely constrained at local sites that bind multiple neighboring PAEs, showing strong covalent-like bonding characteristics; these structures from a building block standpoint are not fully differentiated from conventional PAE-PAE binary crystals, highlighting the PAE:EE duality (*136*).

Taken together, the structural transition from "metallic" to "covalent" phases reveals a negative correlation between EE dynamics and the PAE-EE interaction strength, suggesting a decrease in colloidal crystal metallicity along with the phase transition (*136*). Importantly, although PAE-EE interactions can be continuously tuned by controlling DNA sticky ends, the structural outcome is a discrete set of lattice symmetries.

Observation 4: Flexible DNA ligands and partial EE delocalization are consistent with distorted EE coordination environments and splitting of nonequivalent EE sites.

The formation of low symmetry phases requires the local coordination environment to be distorted and, in certain cases, split into multiple nonequivalent EE sites. These subtle differences in the EE local configurations are evidenced by local maxima in the EE density determined by STEM analysis and the EE visitation frequency derived from MD simulation (Figure 6.4). We first consider the three "tetrahedral" phases – BCT-1, BCT-2 and the FK-A15 phase – as they all have similar tetrahedral EE sites (i.e., each EE site is surrounded by four PAEs) but distinct PAE arrangements. In all three phases, there may be one or multiple EEs at an EE site. In BCT-1, there is one W-S cell: a rhombo-hexagonal dodecahedron where one PAE is surrounded by 18 EE sites (Figure 6.4A). Two types of nonequivalent EE sites exist: a distorted tetrahedral 4d site and an irregular square pyramidal 4e site. In contrast, EEs within the BCT-2 lattice reside at distorted, tetrahedral 4d sites and octahedral 8j sites (Figure 6.4B). PAEs are surrounded by 24 EE sites, rather than 18 as for BCT-1, leading to distinct scattering and lattice parameters despite having body-centered tetragonal symmetry.

In the third case, the FK-A15 unit cell has two types of W-S cells for PAEs: two PAEs reside in 20-coordinated pyritohedra of EE sites, and six PAEs are surrounded by 24 EE sites that

define a truncated hexagonal trapezohedron, also known as the Weaire-Phelan structure (Figure 6.4C). The absence of (110) and (220) scattering peaks indicates that the A15 ($Pm\bar{3}n$) crystal realized here is a single-component system from the PAE perspective. There are three types of EE sites (6d, 16i and 24k sites) within the A15 structure. All three types of EE sites comprise the 24-coordinated W-S cell, while only two (16i and 24k) comprise the 20-coordinated W-S cell, and all are tetrahedrally coordinated.

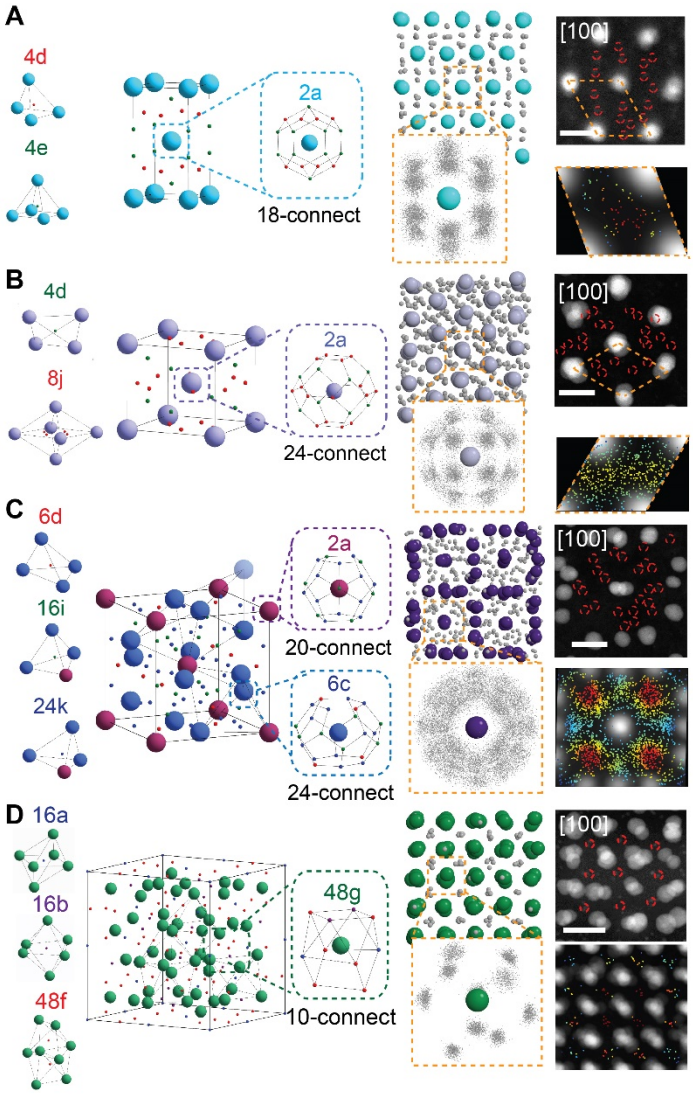


Figure 6.4: Local structural analysis based on STEM and MD simulations. Structural characterization of four representative crystals: BCT-1 (A), BCT-2 (B), FK-A15 (C), and gyroid (D). For each panel:

illustration of a single unit cell composed of PAE and EE sublattices with their corresponding local polyhedral coordination environments and Wyckoff positions (left); MD simulation snapshots of PAE-EE assemblies and EE spatial distribution around a single PAE (middle); representative LAADF image of a single unit cell (EEs are highlighted by red dashed circles), and overlay of averaged-intensity LAADF images of identified EE spatial distribution along the [100] direction (right).

Additionally, we observed the DNA directed-assembly of PAEs into an intriguing triple double-gyroid structure in the cubic $Ia\bar{3}d$ space group (lattice parameter a = 108 nm). This is a rare space group that is found in ternary garnet-type inorganic minerals or in soft matter assemblies as single or double gyroids (*143–145*). Due to the relatively high number of EE sticky ends, EE localization is more pronounced compared to that of the BCT-1, BCT-2, and FK-A15 structures, and only one EE is observed at each EE site. In this triple double-gyroid structure, there is a single coordination environment (W-S cell) of PAEs (48g), and it is an irregular 11-sided polyhedron with 10 vertices (Figure 6.4D). PAEs are surrounded by three inequivalent octahedrally coordinated EE sites (16a, 16b, and 48f), resulting in a unique distorted octahedral close-packing of PAEs that has not been reported before in the colloidal domain. The 48 PAEs of the unit cell can be divided into 3 subgroups (Figure 6.5A), each containing 16 PAEs. Connecting neighboring PAEs within any 1 of these 3 subgroups constructs a skeleton representation of a double gyroid network (Figure 6.5B) (*146*). Consequently, the PAEs of the entire gyroid structure, including all three subgroups, sit at the nodes of the six intertwined gyroid networks (Figure 6.5C and 6.5D). Additionally, connecting EEs on all 16b EE sites gives an additional double-gyroid structure with a slight distortion (Figure 6.13). Together, the PAEs and 16b EEs comprise an 8-srs network structure (*147*), which exists in biological photonic crystals and often exhibits chiroptical properties (*148*). This quadruple double-gyroid structure of PAEs and EEs is among the most topologically complex structures reported to date (*149*) and has not been reported in nanoparticle

assemblies. These four examples illustrate that complex low symmetry phases arise from the distortion of EE local configurations, as confirmed by STEM experiments and MD simulations.

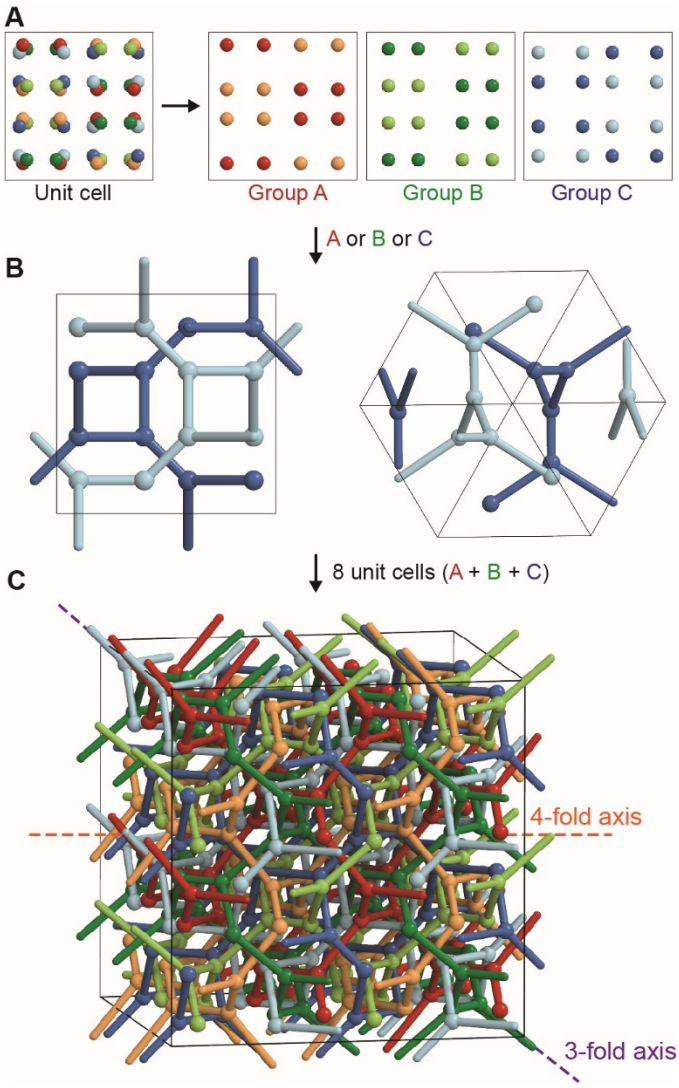


Figure 6.5: Colloidal gyroid crystal structure. (**A**) The 48 PAEs in a unit cell of the triple double-gyroid crystal can be divided into three subgroups, where each subgroup contains 16 PAEs. (**B**) The PAEs in each subgroup are superimposable onto the vertices of skeletal graphs of a double gyroid structure. One of the three double gyroid structures viewed along 4-fold axis (left) and 3-fold axis (right). (**C**) The PAEs in the whole gyroid crystal can be superposed onto the vertices of six intertwined gyroid networks.

Observation 5: Dynamic redistribution of EEs enables reversible *in situ* phase transformations between different crystal structures.

The nine phases described above were all obtained at ambient temperature from different structural configurations of PAEs and EEs. We hypothesized that since EE mobility increases with temperature (*136*) and different equilibrium phases may be favored, PAE lattice symmetry changes could be triggered with a deliberate change in temperature (Figure 6.6A). Importantly, *in situ* variable temperature SAXS experiments reveal that FCC and BCT-2 structures exhibit enantiotropic phase transformations (fully reversible, solid-state transformations without an intermediate liquid phase). Both FCC and BCT-2 transition to the BCC structure upon heating (Figure 6.6B and 6.6C). In the first case, a pure FCC sample annealed by the standard slow-cooling method was heated on a thermal stage at a constant rate of 0.25 °C/min and monitored by SAXS. Distinct from conventional PAE-PAE assemblies that directly melt into an isotropic liquid state, a set of sharp BCC scattering peaks gradually emerge near 41 ºC accompanied by a reduction in the intensity of FCC features (Figure 6.6B). The FCC to BCC transition is complete at 44 ºC, and the BCC structure persists up to 46 ºC before completely melting. The observed BCC phase exhibits high crystallinity as evidenced by the sharp scattering peaks that are well resolved even at high angles. The absence of additional scattering peaks other than those of the BCC and FCC structures suggests that the transformation occurs via a solid-state nucleation-growth pathway rather than the theoretical Bain's pathway (*150*), which involves an intermediate BCT structure (*151*). Indeed, MD simulation shows that nucleation of the BCC phase preferentially occurs through the slip of (111) facets near domain boundaries and subsequent growth into FCC crystallites (Figure 6.6D, 6.6E) rather than through any intermediate phase. In this process, EEs must dynamically reconfigure their locations to induce and accommodate the structural transformation, supported by

the EE density change between the crystalline phase and the surrounding solution, as observed in the MD simulation (Figure 6.17). The redistribution of EEs is further confirmed by accumulated EE positions during the phase transition (Figure 6.6F). The *in situ* change in EE distribution (i.e., a change in PAE symmetry breaking in response to temperature) provides a novel route to rare colloidal crystal-to-crystal transitions (*152*, *153*).

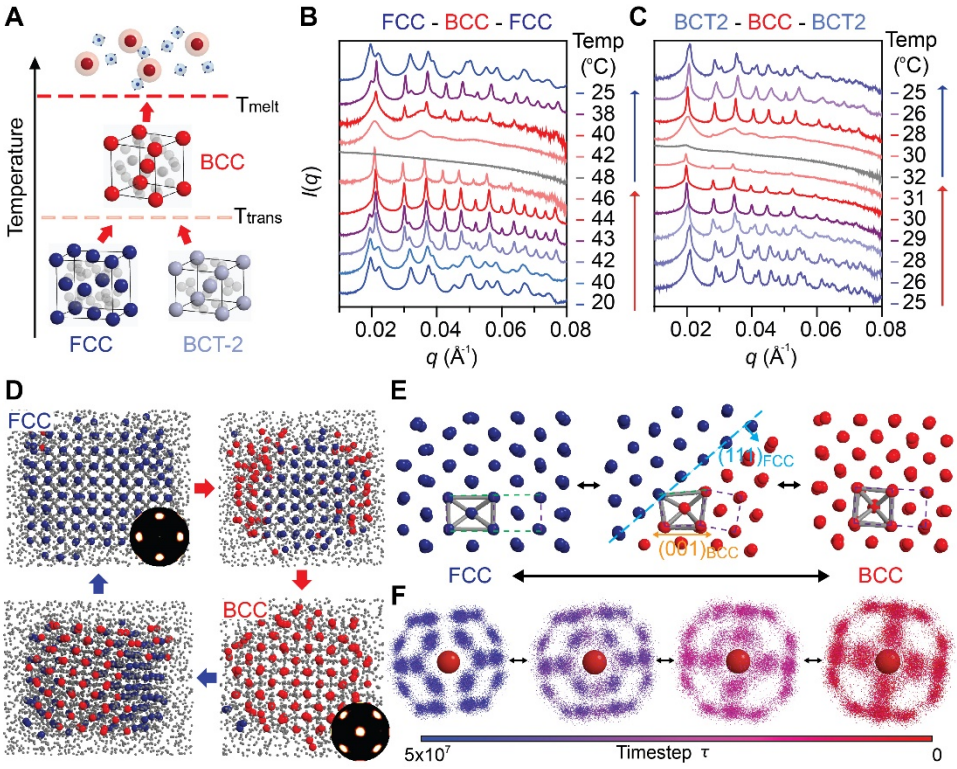


Figure 6.6: Enantiotropic crystal-crystal phase transitions induced by redistribution of EEs. (**A**) Schematic representations of crystal-crystal phase transformation at elevated temperature before complete melting into liquid phase. (**B** and **C**) SAXS patterns for FCC and BCT-2 structures heated and cooled *in situ* resulting in phase transitions. (**D**) MD simulation snapshots of the reversible FCC-BCC phase transition by system temperature change (red arrow = heating, blue arrow = cooling). PAEs are colored by the bond-order parameter (Q_4, *Section 6.4*), where red and blue PAEs correspond to BCC and FCC, respectively. (**E**) Local PAE coordination environment highlighted by a single unit cell (green or purple) during the phase transition in MD simulation. (**F**) The spatial distribution of EEs over time for the FCC-BCC phase transition.

6.3 Conclusions

In summary, we have reported an approach to induce colloidal crystal symmetry breaking via the anisotropic distribution of EEs in colloidal crystals of PAEs engineered with DNA. In this PAE-EE system, the strength of local DNA bonding interactions can be regulated by two means: temperature and the EE-PAE sticky end ratio. Weak interactions occur at high temperature or low EE DNA loading (e.g., 4 sticky ends/particle), where PAEs densely modified with DNA interact isotropically with many EEs, analogous to the spherical free-electron Fermi surface. Strong local interaction occurs at low temperature or high EE DNA loading, where the clustering of DNA bonds and stronger steric repulsion between EEs prevents their free diffusion and leads to symmetry breaking. Intermediate between these extremes, partial delocalization of EEs leads to symmetry breaking through localized, collective bonds involving multiple EEs. We note that the observed symmetry breaking in the PAE-EE system *via* EE redistribution is reminiscent of phase transitions in metals via charge exchange. For instance, at ambient conditions, a handful of metallic elements such as Li (*154*) and Mn (*155*) adopt close-packed BCC or FCC structures as a result of nondirectional metallic cohesion involving itinerant electrons. At elevated temperature or pressure, these metals undergo a phase transformation to multiple low symmetry structures, due to significant frustration of the free-electron Fermi surface topology by strong local interactions (*156*). Furthermore, certain mechanistic features are shared between this proposed mechanism and mesophase formation involving soft matter, which is typically driven by noncovalent interactions between spherical micelles formed by lipids (*157*), molecular (*130*) and nanoparticle (*158*) surfactants, dendrimers (*159*), and block copolymers (*64*). In these systems, local asymmetries originate from molecular flexibility, building block size dispersity, micellar size and shape

polydispersity, and inter-micelle mass transfer, causing relaxation of cellular domains to unequal volumes with minimized contact surface and overall increased packing density (*152*, *160*).

In conclusion, our strategy for symmetry breaking relies on the establishment of the new concept of "valence electrons" in the PAEs and EEs that comprise colloidal crystals. Importantly, the nine crystals reported herein are only a small fraction of the possible structures. Based on this approach, we expect more low symmetry and complex colloidal architectures to be discovered, including other Frank-Kasper phases, Laves phases, and quasicrystalline structures.

6.4 Method and supporting information

Simulations of this study were performed using molecular dynamics (MD) implemented in the HOOMD-blue simulation package (https://github.com/glotzerlab/hoomd-blue) (*93*, *161*, *162*). HOOMD-blue can be performed either on the central processing units (CPUs) or on the graphics processing units (GPUs), but it is highly optimized on GPUs. As such, most simulations of the present work were performed on NVIDIA Tesla P100 GPUs (*163*).

6.4.1 Simulation setup

We developed a simplified model of PAEs and EEs to simulate a large number of particles on a reasonable timescale. A PAE with a large number of DNA ligands is modeled as a sphere that has an isotropic interaction with EEs and other PAEs. The EE model contains two parts: attractive patches and a spherical body on which the patches reside on. The patches are uniformly positioned around the EEs where the sum of the distances between them is maximized. The purely repulsive Weeks-Chandler-Andersen (WCA) pair potential (*121*),

$$V(r) = \begin{cases} 4\varepsilon_{\mathrm{WCA}}\left[\left(\frac{\sigma_{\mathrm{WCA}}}{r}\right)^{12} - \left(\frac{\sigma_{\mathrm{WCA}}}{r}\right)^{6}\right] - 4\varepsilon_{\mathrm{WCA}}\left[\left(\frac{\sigma_{\mathrm{WCA}}}{r_{\mathrm{cut}}}\right)^{12} - \left(\frac{\sigma_{\mathrm{WCA}}}{r_{\mathrm{cut}}}\right)^{6}\right], & r < r_{\mathrm{cut}} \\ 0, & r \geq r_{\mathrm{cut}} \end{cases} \quad \text{(Eq. 6.1)}$$

is applied to the PAE–PAE, PAE–EE body, and EE body–EE body interactions (Figure 6.3B and 6.3C), where r is the center-to-center separation distance between two particles. The σ_{WCA} between *i* and *j* particles is determined by the radius of each particle and the overlap parameters (ω_{ij}), $\sigma_{\mathrm{WCA\_}ij} = (1 - \omega_{ij}) \times (r_i + r_j)$. The $r_{\mathrm{cut}} = \sigma_{\mathrm{WCA}} \times 2^{1/6}$ and the $\varepsilon_{\mathrm{WCA}}$ are 1.0 for all systems. The shifted Gaussian model is applied to the PAE–EE patch interaction:

$$V(r) = \varepsilon_{Gauss} \exp\left[-\frac{1}{2}\left(\frac{r-r_0}{\sigma_{Gauss}}\right)^2\right] \quad \text{(Eq. 6.2)}$$

which mimics the attraction caused by DNA hybridization. The shifted Gaussian pair potential is truncated for separation beyond $V(r) < 10^{-2}$. We controlled the number of patches ($4 \leq N_{patch} \leq 12$) on EEs as a simulation parameter to mimic the estimated number of ligands on EEs in the experiments.

6.4.2 Overlap parameters

In an atomistic model of the PAE (*16*, *164*), the ligand layers would be somewhat compressible and overlap. In our patch particle model, to control the amount of overlap between particles, we introduced an overlap parameter (ω) in the simulation. The amount of overlap of a particle is defined by $\omega = l/r$, which is a ratio between the thickness of the overlap region that is allowed to overlay with other particles (l) and the radius of the particle (r) (Figure 6.7A). As such, $\omega = 0$ defines that overlap is not allowed on the particle, and $\omega = 1$ defines that the particle can be fully overlapped with other particles. In principle, the overlap parameter between two particles (ω_{ij}) should be solely determined by ω_i and ω_j, but we introduced a constant coefficient c_{ij} for allowing us to slightly adjust the overlap specific types of pairs (Figure 6.7B). As such, the general formula of ω_{ij} is defined as:

$$\omega_{ij} = \frac{l_i + l_j}{r_i + r_j} - c_{ij}\frac{(r_i - l_i) + (r_j - l_j)}{r_i + r_j} = \frac{\omega_i r_i + \omega_j r_j - c_{ij}[(1-\omega_i)r_i + (1-\omega_j)r_j]}{r_i + r_j} \quad \text{(Eq. 6.3)}$$

For $c_{ij} = 0$, the ω_{ij} is determined by ω_i and ω_j only. We kept $c_{\mathrm{PAE-EE}} = 0.05$ and $c_{\mathrm{EE-EE}} = 0$ as constants for all calculations and simulations in this work. In addition, the value of $c_{\mathrm{PAE-PAE}}$ is changed when we modulate the rigidity of PAE shells due to the dummy linkers. The values of the simulation parameters used in this work are listed in Table 6.2.

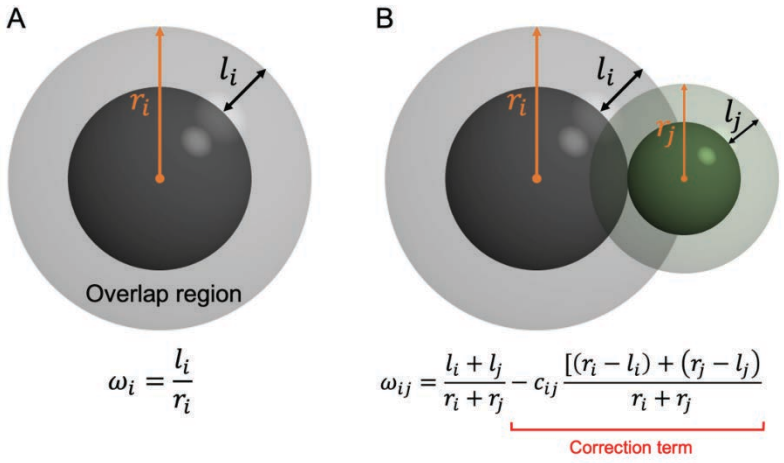


Figure 6.7: Schematic diagram showing the definition of the overlap parameter. (A) r_i is the radius of a particle that includes the nanoparticle core and DNA ligand shell. l_i is the thickness of the shell (grey) that is allowed to overlap with other particles. (B) The amount of overlap between the two particles (ω_{ij}) can be tuned by the correction term, which is determined by a constant coefficient of c_{ij}. If $c_{ij} = 0$, the overlap parameter of pair *ij* is purely determined by ω_i and ω_j. The diagram in (B) shows a maximum overlap between *i* and *j* particles when $c_{ij} = 0$.

Table 6.2: Simulation parameters where the nine phases are stable. For all system, $r_{\mathrm{PAE}} = 2.70\sigma$, $r_{\mathrm{EE}} = 1.84\sigma$, $\phi_{PAE} = 0.05$, $\varepsilon_{Gauss} = 0.70$, $\sigma_{Gauss} = 0.50$, $r_0 = r_{\mathrm{PAE}} \times 2^{1/6}\,\sigma$, $c_{EE-EE} = 0$, and $c_{PAE-EE} = 0.05$.

Phase	N_{Patch}	N_{PAE}: N_{EE}	T_m^*	ω_{EE}	ω_{PAE}	$c_{PAE-PAE}$
FCC	4	1:8	~0.30	0.35	0.00	0.00
BCC	4	1:6	~0.30	0.35	0.00	0.00

BCT-2	4	1:4	~0.32	0.30	0.00	-0.05
A15	4	1:6	~0.26	0.33	0.00	0.00
BCT-1	5	1:4	~0.44	0.30	0.10	0.00
SH	8	1:2	~0.42	0.10	0.00	0.00
Gyroid	10	1:2	~0.38	0.00	0.00	0.00
HP-Ga	6	1:2	~0.37	0.10	0.05	-0.05
SC	12	1:1	~0.60	0.00	0.10	0.00

6.4.3 Overlapping packing model

An overlapping packing model (OPM) compares the maximum packing density of various crystals where overlap between particles is allowed. The overall packing density ϕ of the PAE-EE binary system is defined as $\phi = (N_{\mathrm{PAE}} \times v_{\mathrm{PAE}} + N_{\mathrm{EE}} \times v_{\mathrm{EE}})/V_{\mathrm{Box}}$, where N, v, and V_{Box} are the number of particles, volume of a particle, and total volume of a system, respectively. In this study, we compared the maximum packing density of nine PAE-EE crystals depending on the PAE overlap (ω_{PAE}) and the EE overlap (ω_{EE}).

Steps for calculating the maximum packing density of an input lattice for a given set of overlap parameters are:

1. Assign an input lattice with a low-enough packing density such that no overlap exists between any particle. In this study, the initial packing density was set to $\phi = 0.1$.
2. Assign the radius of each type of particle (r_{PAE} and r_{EE}) and the overlap parameters (ω_{PAE}, ω_{EE}, and c).
3. Slightly compress the crystal with a proper scale. We compressed the system 0.5% in the length scale in each step.
4. Calculate the PAE-PAE, EE-EE, and PAE-EE distances that are averaged by the number of the nearest neighbor.
5. Check if:

- $d_{\mathrm{PAE-PAE}} < (1 - \omega_{\mathrm{PAE-PAE}})(r_{\mathrm{PAE}} + r_{\mathrm{PAE}})$
- $d_{\mathrm{EE-EE}} < (1 - \omega_{\mathrm{EE-EE}})(r_{\mathrm{EE}} + r_{\mathrm{EE}})$
- $d_{\mathrm{PAE-EE}} < (1 - \omega_{\mathrm{PAE-EE}})(r_{\mathrm{PAE}} + r_{\mathrm{EE}})$

6. If all three conditions are false, go back to step 3. Otherwise, go to the next step.
7. Calculate the packing density of the current state. This is the maximum packing density of the input lattice for a given set of overlap parameters.

For all nine crystals found in this study, we carried out the OPM calculation for $0 \leq \omega_{\mathrm{PAE}} \leq 0.30$, and $0 \leq \omega_{\mathrm{EE}} \leq 0.55$ with 0.01 intervals, $c_{\mathrm{PAE-EE}} = 0.05$, $c_{\mathrm{PAE-PAE}} = 0$, and $c_{\mathrm{EE-EE}} = 0$. Figure 6.8A – 6.8C shows three example results (BCC, BCT-2, and gyroid) of the OPM calculation where ω_{PAE} and ω_{EE} vary. By stacking all 3D plots for the nine crystal structures (Figure 6.8D), we can identify the structure that has the maximum packing density for a given set of overlap parameters (Figure 6.8E; see also Figure 6.8F for a cross section at a fixed ω_{PAE} value). Note that ϕ can be larger than 1.0 because there exists a significant amount of overlap between particles at the maximum packing state.

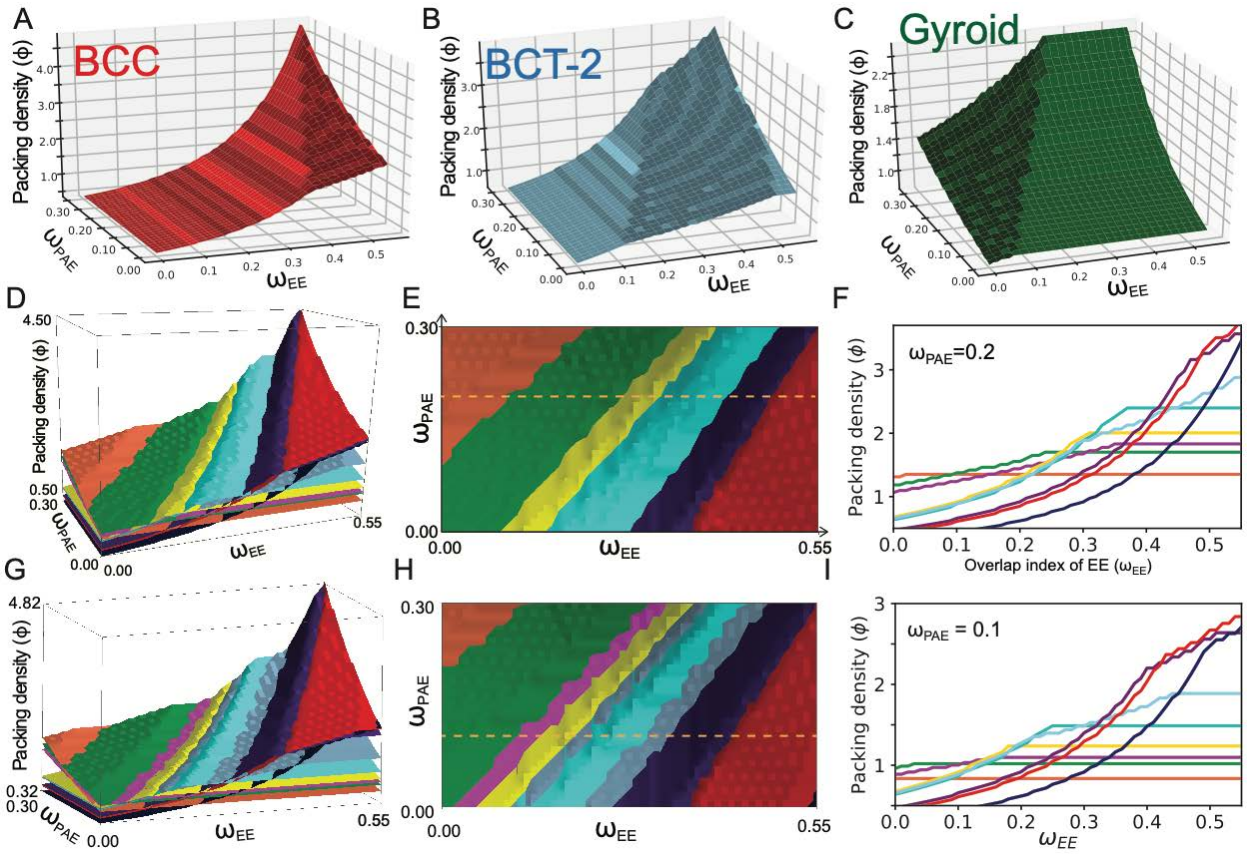


Figure 6.8: Overlap packing model (OPM) calculation results. (A–C) Maximum packing fraction calculated by OPM for (A) BCC, (B) BCT-2, and (C) gyroid crystals in the range of $0 \leq \omega_{\mathrm{PAE}} \leq 0.30$ and $0 \leq \omega_{\mathrm{EE}} \leq 0.55$. $c_{\mathrm{PAE-EE}} = 0.05$, $c_{\mathrm{PAE-PAE}} = 0$, and $c_{\mathrm{EE-EE}} = 0$. (D) OPM calculations for all nine lattices at $c_{\mathrm{PAE-EE}} = 0.05$, $c_{\mathrm{PAE-PAE}} = 0$, and $c_{\mathrm{EE-EE}} = 0$. (E) Top view of the 3D plot in (D), showing the maximum packing structures. (F) Plot of the maximum packing fraction of the nine lattices as a function of ω_{EE} when ω_{PAE} is a constant ($\omega_{\mathrm{EE}} = 0.2$, yellow dashed line in (E)). (G) OPM calculations for all nine lattices at $c_{\mathrm{PAE-PAE}} = -0.05$. HP-Ga (magenta) and BCT-2 (light blue) phases appear as the maximum packing structure for a certain range of ω_{PAE} and ω_{EE}, which was not observed when $c_{\mathrm{PAE-PAE}} = 0$ (Figure 6.3D-6.3E). (H) Top view of the 3D plot in (G). (I) Plot of the maximum packing fraction of the nine lattices as a function of ω_{EE} when ω_{PAE} is a constant ($\omega_{\mathrm{EE}} = 0.1$, yellow dashed line in (H)). The color scheme is the same as the colors used in the main text.

6.4.4 Crystallization simulations

For the assembly simulations, the number of PAEs is fixed ($N_{\mathrm{PAE}} = 500$), and the number of EEs (N_{EE}) varies from 500 to 8000. This system size is large enough to contain a few hundred unit cells of the crystal structures in this study. We kept the packing fraction of PAEs constant

($\phi_{PAE} = N_{\mathrm{PAE}} \times v_{\mathrm{PAE}}/V_{\mathrm{Box}} = 0.05$), where v_{PAE} is the volume of a PAE. For each MD timestep, $\delta\tau = 0.005$ is applied for all systems. Brownian dynamics are applied, and all simulations are conducted under periodic boundary conditions.

As an initial setup, we place the PAEs and the EEs in a large simulation box (ϕ_{PAE}~0.01) with no overlap between the particles and assign a random orientation to the EEs. Then, we compress the system to a target density ($\phi_{PAE} = 0.05$) and equilibrate the system at a constant temperature (T^*) by checking the potential energy of the system every $10^3\tau$. If the T^* is lower than T_m^*, a supercooled liquid cluster, where the PAEs and the EEs are mixed, is formed. Once the system overcomes the nucleation barrier within the simulation time, spontaneous crystallization occurs, as identified by a drop in the potential energy of the system, the diffraction patterns, and the bond-order diagrams. Typically, nucleation of the crystals was observed around T^*/T_m^*~0.9 within 10^8 MD timesteps (τ), which takes around 48–72 GPU h.

6.4.5 Phase identification from simulation results

In order to identify the crystal structures obtained in the MD simulations, we used diffraction patterns of PAE crystals, bond orientational order diagrams (BOD) of PAEs, and direct visualization of simulation snapshots (Figure 6.9 to 6.12). The diffraction patterns were calculated by FFTs of the PAEs configurations in real space. The BOD shows a stereographic projection of the accumulated orientation of the nearest-neighbor bonds along a crystal symmetry axis. In order to get references for the structural information, we constructed the ideal structures of the nine PAE lattices observed in experiments and obtained the BODs and the diffraction patterns from a main zone axis (Figure 6.9). We used them as references to identify the crystal structures obtained from MD simulations.

For the FCC-BCC transition simulation (Figure 6.6D-6.6E and E11), we used the rotationally invariant Steinhardt bond order parameter (Q_l) (*94*, *165*) that associates a set of spherical harmonics with every bond between particle i and its neighbors j. We denote

$$Q_{lm}(\vec{r}) = Y_{lm}(\theta(\vec{r}), \phi(\vec{r})) \quad \text{(Eq. 6.4)}$$

where $\vec{r}$ is the vector of the bond that connects from i to j, and $Y_{lm}(\theta, \phi)$ are spherical harmonics. To determine the local environment, we averaged over all nearest neighbors

$$\bar{Q}_{lm} = \langle Q_{lm}(\vec{r}) \rangle \, . \quad \text{(Eq. 6.5)}$$

To remove local orientational order, one should use the rotationally invariant combination

$$Q_l = \left[\frac{4\pi}{2l+1} \sum_{m=-l}^{l} |\bar{Q}_{lm}|^2 \right]^{1/2} \quad \text{(Eq. 6.6)}$$

We used Q_4 to distinguish between BCC ($Q_4 < 0.1$) and FCC ($Q_4 \geq 0.1$) phases (*35*). Only the bonds between PAEs are considered when the Q_4 is calculated.

6.4.6 PAE-EE phase diagram simulation

We first performed MD simulations by changing the number of patches on the EE, $4 \leq N_{\text{patch}} \leq 12$, and the amount of overlap of the EEs, $0 < \omega_{\text{EE}} < 0.5$, while keeping $\omega_{\text{PAE}} = 0$ and $c_{\text{PAE-PAE}} = 0$. We observed spontaneous formation of gyroid, SH, BCT-1, A15, and BCC lattices as we increase the ω_{EE} and decrease the number of EE patches (Figure 6.10 to 6.12). This observation agrees well with the experiments where gyroid and SH were found with high EE grafting density, that BCT-1 and A15 were found with intermediate grafting density, and that BCC was found with high grafting density. When we increase ω_{PAE} by 10% in the high grafting EE density region ($0 < \omega_{\text{EE}} < 0.1$), we observed the formation of the SC phase. In the low grafting

EE density region ($0.35 < \omega_{EE} < 0.50$), when we increased the EE concentration, we observed a phase transformation from BCC to FCC. Note that FCC was assembled only when the EE density is high enough (PAE:EE > 1:25 in this study). In the range of $0.05 < \omega_{EE} < 0.15$ and $0.30 < \omega_{EE} < 0.40$, when we decreased $c_{PAE-PAE}$ 5%, which effectively increased the rigidity of the PAE shell, we observed that HP-Ga and BCT-2 phases were well stabilized, consistent with experiments.

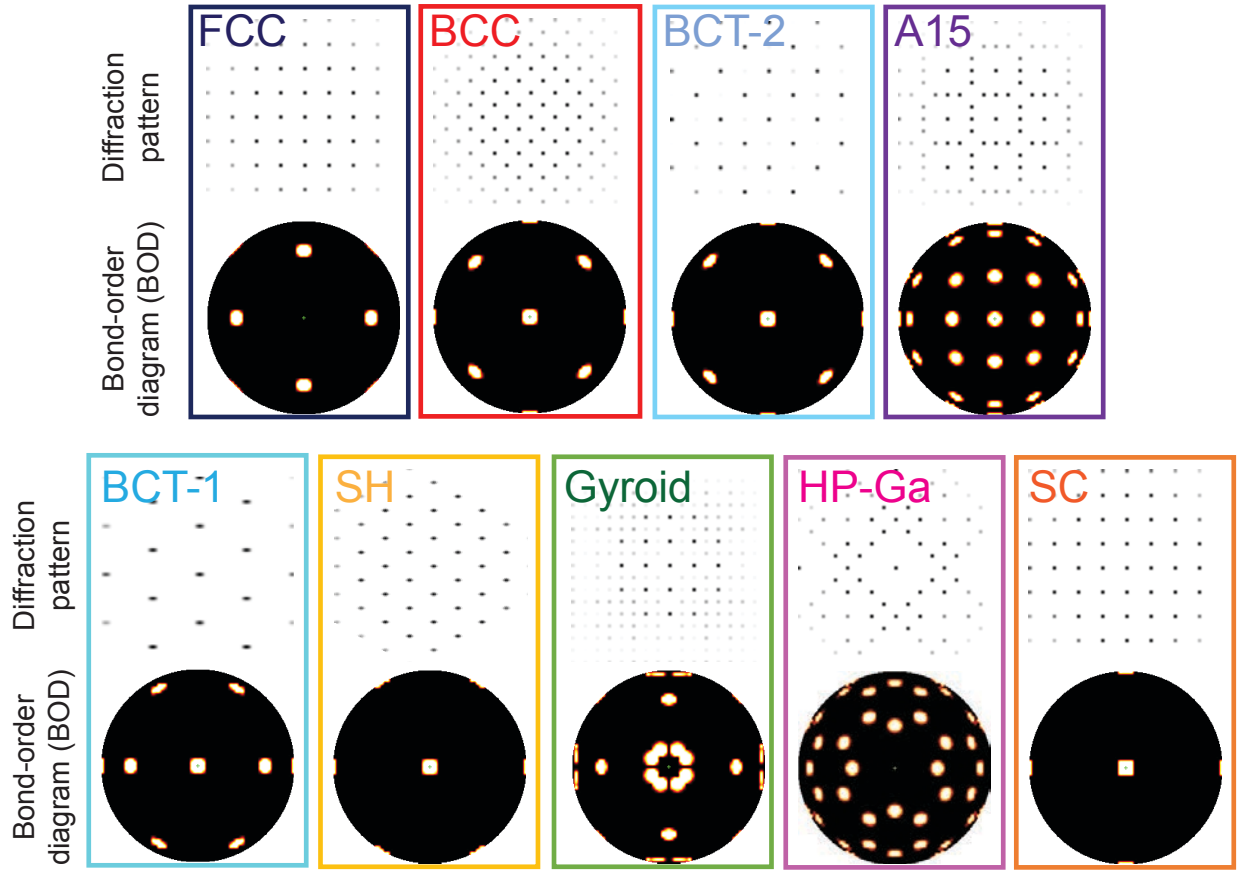


Figure 6.9: Diffraction patterns and bond-order diagrams (BOD) of the nine PAE lattices from a main zone axis. These crystal structures are constructed in ideal positions to use their diffraction patterns and BODs as a reference to identify the structures found in the MD assembly simulation.

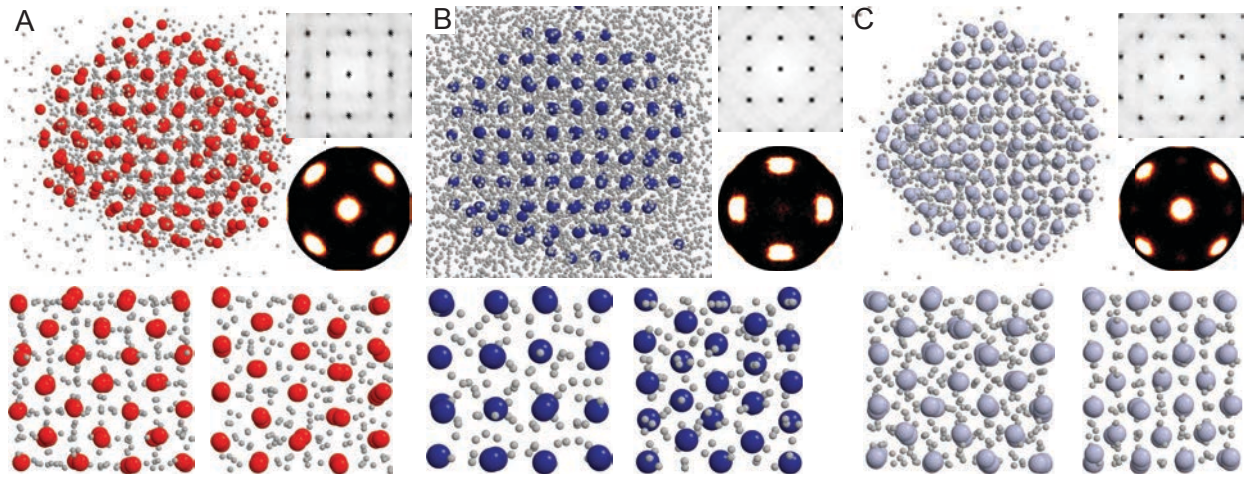


Figure 6.10: Assembly results of BCC, FCC and BCT-2 from the MD simulation. Each lattice formation is identified by direct visualization of snapshots, diffraction patterns, and bond-order diagrams. (top left) Simulation snapshots of (A) BCC, (B) FCC, and (C) BCT-2 lattices along the four-fold symmetry axis. (top right) Diffraction pattern (up) and bond-order diagram (down) along the four-fold symmetry axis. (bottom row) Zoomed-in view of each lattice from a different symmetry axis.

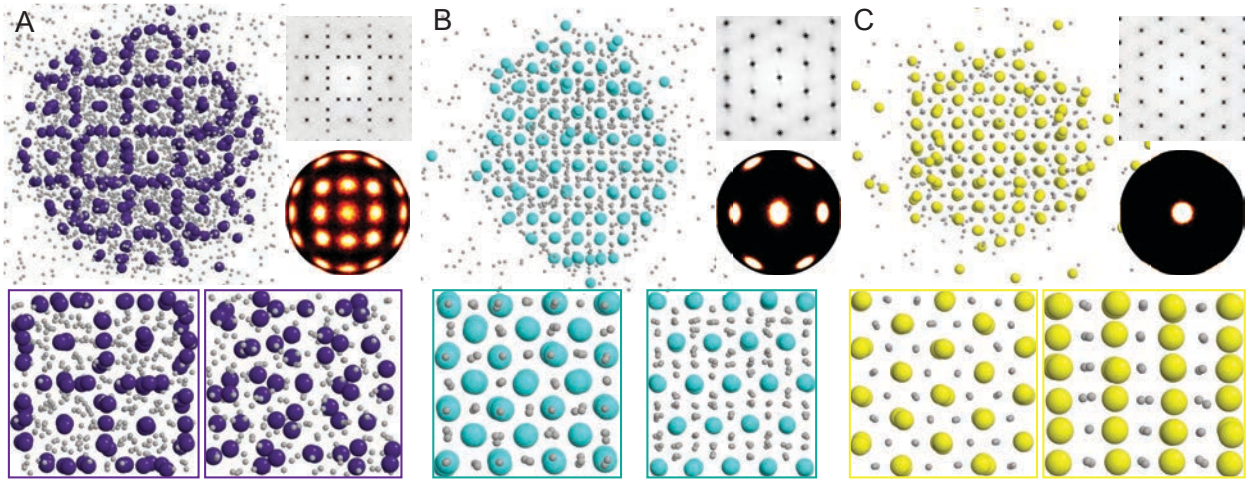


Figure 6.11: Assembly results of A15, BCT-1 and SH from the MD simulation. Each lattice formation is identified by direct visualization of snapshots, diffraction patterns, and bond-order diagrams. (top left) Simulation snapshots of (A) A15, (B) BCT-1, and (C) SH lattices along the four-fold, two-fold, and three-fold symmetry axes, respectively. (top right) Diffraction pattern (up) and bond-order diagram (down) along each symmetry axis. (bottom row) Zoomed-in view of each lattice from a different symmetry axis.

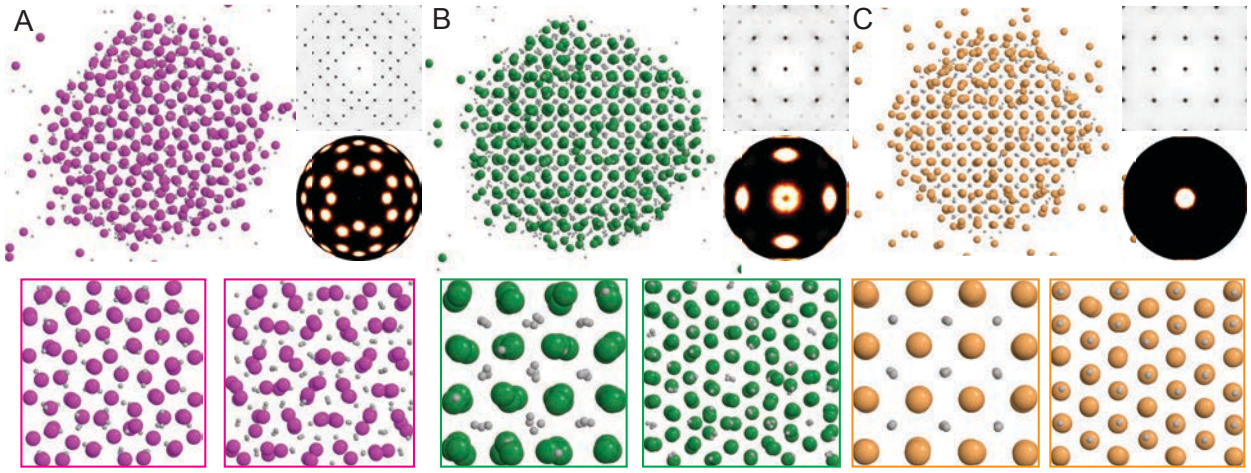


Figure 6.12: Assembly results of HP-Ga, gyroid and SC from the MD simulation. Each lattice formation is identified by direct visualization of snapshots, diffraction patterns, and bond-order diagrams. (top left) Simulation snapshots of (A) HP-Ga, (B) gyroid, and (C) SC lattices along the four-fold symmetry axis. (top right) Diffraction pattern (up) and bond-order diagram (down) along the four-fold symmetry axis. (bottom row) Zoomed-in view of each lattice from a different symmetry axis.

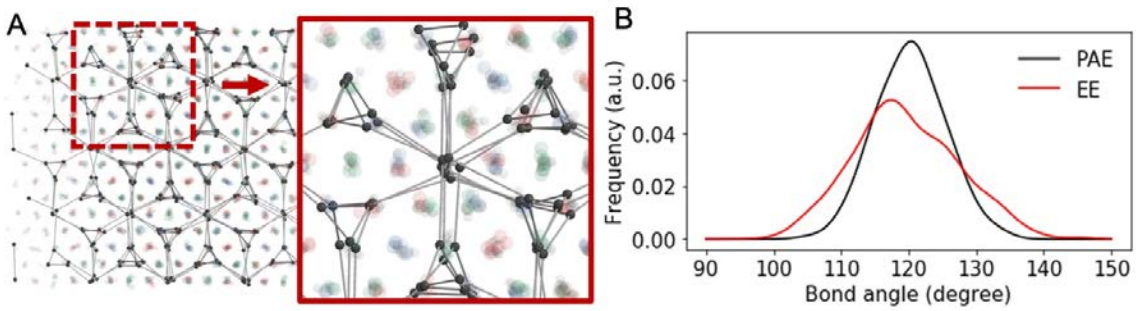


Figure 6.13: Double-gyroid structure constructed by EEs in 16b sites. (A) MD simulation snapshot of EEs in 16b sites (dark grey) that are connected by light grey sticks to represent the double-gyroid structure. (B) Bond-angle distributions calculated from the PAE gyroid (black) and the EE gyroid (EE). The bond angles were obtained from the bonds that connect three nearest particles of each particle in a gyroid network. The EE gyroid shows a broader distribution than that of the PAE gyroid.

6.4.7 Spatial distribution of EE in PAE lattice

The change of spatial distribution of EEs in a PAE lattice as a function of EE grafting density is important for understanding the PAE-EE systems, which have nine different crystal structures formed under different conditions. However, in experiments, it is difficult to specify the

location of EEs in a PAE lattice if EEs are highly mobile (e.g., for BCC or A15). Therefore, we investigated the spatial distribution of EEs in MD simulations from the accumulated positions of EEs during a given simulation time window. In order to remove the surface effect between the crystal and the surrounding gas phase, we constructed periodic bulk crystals of all nine phases. After thermalizing the constructed bulk crystals for $10^6\tau$ in $T^*/T_m^* \sim 0.95$, we obtained the positions of EEs around a PAE for every $10^3\tau$ during an entire time window of $10^7\tau$ and then visualized the accumulated EE positions (about a few thousand EE positions) around a PAE, which give us statistically relevant spots showing EE localization in each PAE lattice (Figure 6.14). If there are multiple types of PAEs in terms of their local environments, the EE distribution for each case was calculated separately. As expected, each PAE lattice has a different EE spatial distribution, and more importantly, the PAE lattices that are formed with low EE grafting densities (e.g., A15, FCC, and BCC) show a broader and more isotropic distribution of EEs. This implies that the EEs are highly dynamic and exchange their positions frequently.

Based on the accumulative EE positions around a PAE for each lattice, we calculated isosurface plots that visualize the surfaces with the same EE visitation frequencies (Figure 6.5A). To achieve this, we constructed 3D girds ($50 \times 50 \times 50$ cells, each cell has an edge length of 0.4σ) in the space around the PAEs and obtained a 3D histogram of the EE visitation frequency. From the histogram, we made a contour plot that connects grids with the same visitation frequencies. A Gaussian filter from SciPy was further used to smooth the plots.

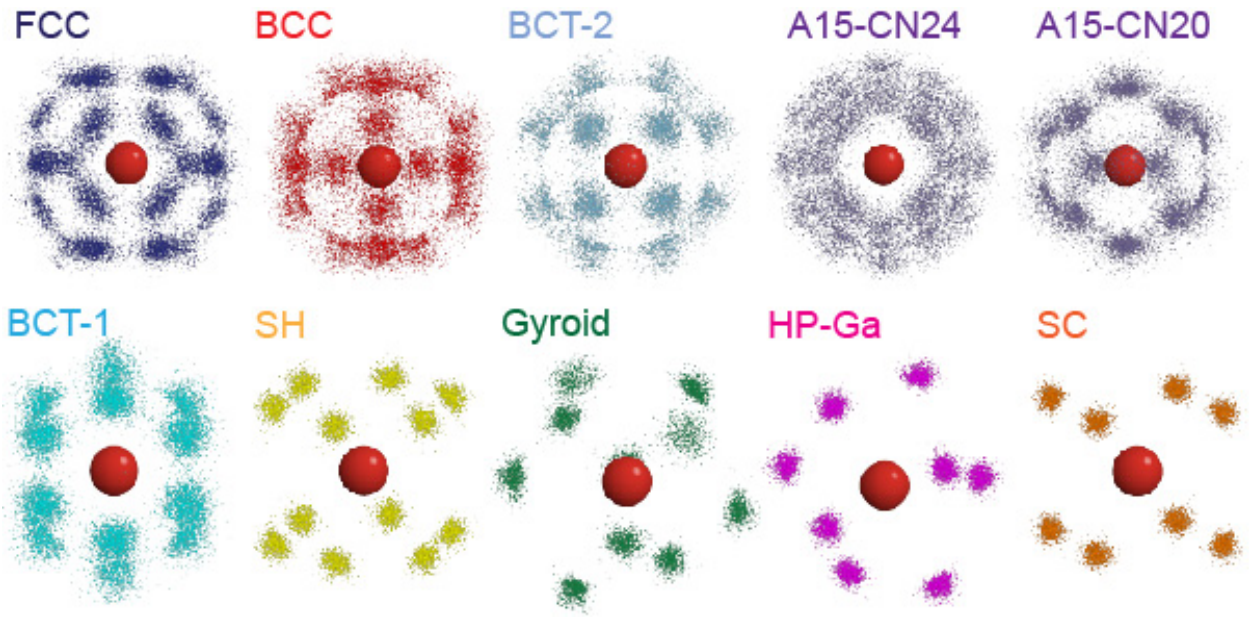


Figure 6.14: Accumulated position of EEs around a PAE for the nine different PAE lattices. The EE positions are collected from periodically constructed crystals at T^*/T_m^* ~0.95 every $10^3\tau$ during $10^7\tau$. The A15 phase has two different types of PAEs in terms of the local environment in the lattice (coordination number: CN = 20 or 24).

6.4.8 Dynamics of EE

The mean-squared displacement (MSD), which is defined as $\frac{1}{N}\sum_{i=1}^{N}\left|\boldsymbol{r}^i(\tau)-\boldsymbol{r}^i(0)\right|^2$, was calculated to measure the diffusivity of the EEs. We calculated the MSD of EEs in the nine periodic PAE-EE crystals in this work at a constant temperature (T^*/T_m^* ~0.95) for $10^7\tau$ (Figure 6.15). The packing density of PAEs (ϕ_{PAE}) in each system is determined by the local packing density of PAEs when the crystal cluster is fully surrounded by a gas phase at T^*/T_m^* ~0.95. We observed a linear increase of the MSD for five crystals (BCC, FCC, A15, BCT-2, and BCT-1), which indicates a diffusive motion of EEs within the PAE lattice (Figure 6.5D). On the other hand, the other four crystals (SC, Gyroid, HP-Ga, and SH) show negligible changes in the MSD during the simulation time window, suggesting that the EEs are highly localized in these PAE lattices.

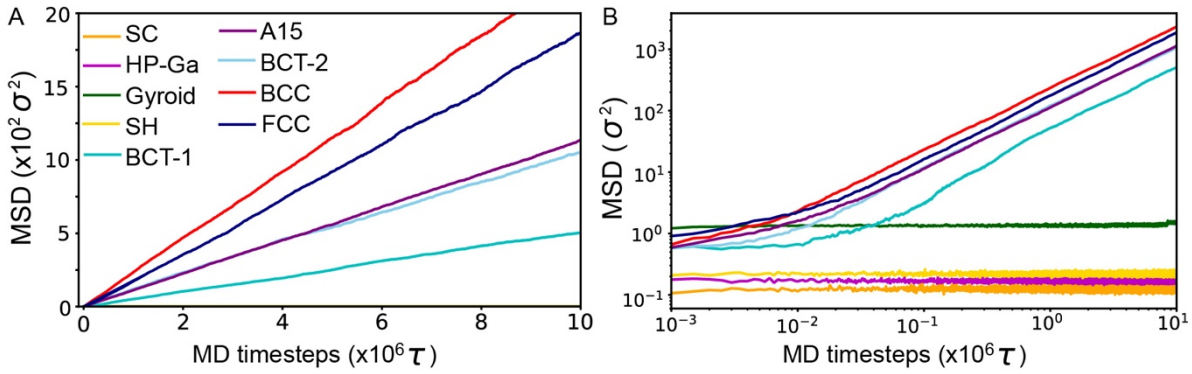


Figure 6.15: (A) Mean squared displacement (MSD) of the EEs in the nine PAE lattices. (B) The MSD plots in the log scale. The MSD was calculated from the periodically constructed crystals at T^*/T_m^* ~0.95 during $10^7\tau$. The ϕ_{PAE} of the periodic crystal for FCC, BCC, BCT-2, A15, BCT-1, SH, gyroid, HP-Ga, and SC are 0.33, 0.34, 0.40, 0.33, 0.43, 0.42, 0.38, 0.55, and 0.56, respectively. The ϕ_{PAE} of each system was determined by the local packing density of the PAEs when the crystal cluster is fully surrounded by the gas phase at T^*/T_m^* ~0.95. Other simulation parameters are given in Table 6.2.

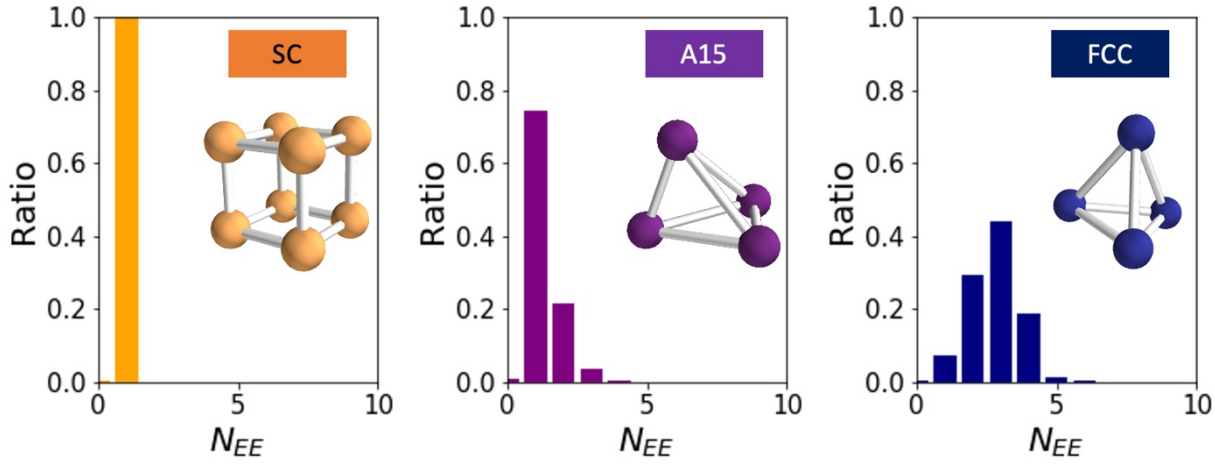
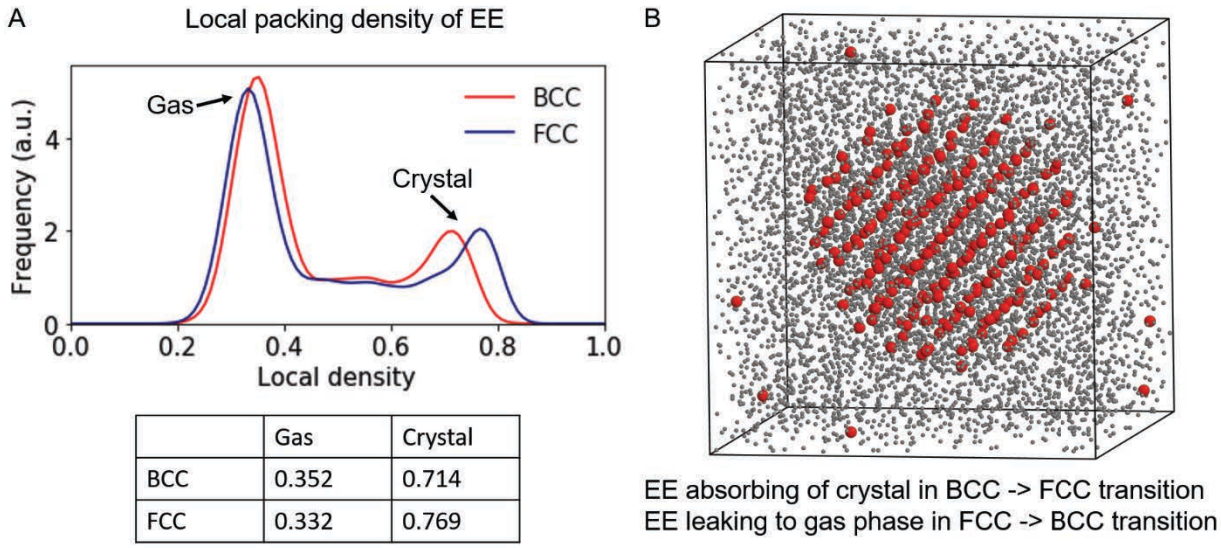


Figure 6.16: Clustering tendency at an EE site for three representative systems: SC for covalent phases (left), A15 for tetrahedral phases (middle), and FCC for metallic phases (right). This is measured by the distribution of the number of EEs at a EE site, which is calculated from every $10^3\tau$ for $10^7\tau$ in equilibrium (Table 6.2).



	Gas	Crystal
BCC	0.352	0.714
FCC	0.332	0.769

Figure 6.17: (A) MD simulation of the local packing density of EEs for BCC and FCC phases in the gaseous and crystalline phase, respectively. The local density of EEs does not account for the PAEs. (B) A simulation snapshot of a BCC cluster consisting of PAEs (red spheres) and EEs (grey spheres), which is surrounded by the gas phase of EEs, at T^*/T_m^* ~0.95. The EEs in the gas phase will be absorbed into the crystal cluster during the BCC to FCC phase transition.

Table 6.3: DNA sequences used in this study.

Strand	DNA sequence from 5' to 3'
Au NP-bound A	Thiol-$(Sp18)_2$-CATCCATCCTTATCAACT
Au NP-bound B	Thiol-$(Sp18)_2$-AACGACTCATACTCACCT
Dye-labeled strand 1	Thiol-(Sp18)-Cy5-(Sp18)-AACGACTCATACTCACCT
Dye-labeled strand 2	Thiol-(Sp18)-Cy3-(Sp18)-AACGACTCATACTCACCT
Nonself-comp linker A	TTCCTT-(Sp18)-AGTTGATAAGGATGGATG
Nonself-comp linker B	AAGGAA-(Sp18)-AGGTGAGTATGAGTCGTT
Nonself-comp linker A SE7	TTCCTTC-(Sp18)-AGTTGATAAGGATGGATG
Nonself-comp linker B SE7	GAAGGAA-(Sp18)-AGGTGAGTATGAGTCGTT
Dummy linker A'	$(Sp18)_8$-AGGTGAGTATGAGTCGTT
Self-comp linker A	CGCG-(Sp18)-AGTTGATAAGGATGGATG

All modified phosphoramidites were manufactured by Glen Research.

1. "**Thiol**" refers to 1-O-Dimethoxytrityl-hexyl-disulfide,1'-[(2-cyanoethyl)-(N,N-diisopropyl)]-phosphoramidite (Thiol-modifier C6 S-S).
2. "**Sp18**" refers to 18-O-Dimethoxytritylhexaethyleneglycol,1-[(2-cyanoethyl)-(N,N-diisopropyl)]-phosphoramidite (Spacer phosphoramidite 18).
3. "**Cy3**" refers to 1-[3-(4-monomethoxytrityloxy)propyl]-1'-[3-[(2-cyanoethyl)-(N,N-diisopropyl) phosphoramidityl]propyl]-3,3,3',3'-tetramethylindocarbocyanine chloride (Cyanine 3 Phosphoramidite).
4. "**Cy5**" refers to 1-[3-(4-monomethoxytrityloxy)propyl]-1'-[3-[(2-cyanoethyl)-(N,N-diisopropyl) phosphoramidityl]propyl]-3,3,3',3'-tetramethylindodicarbocyanine chloride (Cyanine 5 Phosphoramidite).

CHAPTER 7

Colloidal Fibers and Rings by Cooperative Assembly of Janus Particles

This chapter is reproduced from Ref. (*166*), a publication in *Nature Communications*, authored in 2019 by Joon Suk Oh, Sangmin Lee, Sharon C. Glotzer*, Gi-Ra Yi*, David J. Pine* (*corresponding authors). My contribution to this work included MD simulation model development of Janus colloids with varying patch size, reproducing assembly process of various superstructures, constructing a full phase diagram of Janus colloids, confirming cooperative assembly process of colloidal fibers and providing simulation evidence of the depletion effect on colloidal ring formation.

7.1 Introduction

The self-assembly of proteins into crystals or biologically active structures is a complex dynamical process that depends critically on both the shape and the chemical heterogeneity (patchiness) of the protein surface (*167*, *168*). Colloidal particles have served as model systems for exploring the role of shape in self-assembly (*4*, *38*, *169*), and notably for lock-and-key interactions using dimpled colloids (*170*).

The role of chemical heterogeneity may play an even greater role than shape in self-assembly, but there are few experiments on model systems where the strength and extent of the chemical heterogeneity are well controlled (*171*). One of the simplest systems exhibiting complex

structures due to chemical heterogeneity is the self-assembly of Janus colloids with one attractive patch, with the rest of the particle being repulsive. The different structures they form depend on the interaction strength ε of the attractive patch and the Janus balance, defined as fractional area χ of the attractive patch. Taken together, these play the same role as the hydrophilic-lipophilic balance (HLB) for surfactants. Although Janus particles with various patch ratios and interaction ranges have been explored in computer simulations (*172–178*), systematic experimental studies are rare, as it has proven difficult to adjust the patch ratio of Janus particles using available fabrication methods (*179*). Previous experimental work with hydrophobic patches on otherwise hydrophilic Janus particles suggests that kinetic trapping can select nonequilibrium structures over equilibrium ones, so that the ability of structures to anneal and equilibrate can play a role in structure formation.

Here we introduce a versatile and scalable method to synthesize Janus particles with well-controlled Janus balance χ. One face of the Janus particles is selectively functionalized with DNA having self-complementary sticky ends. We use DNA-mediated attraction rather than a hydrophobic attraction for three reasons: (1) the DNA-mediated interaction is thermally-reversible, (2) properly-functionalized DNA-coated colloids can anneal after they bind to each other (*180*), and (3) the strength of the DNA-mediated attraction can be continuously varied from essentially zero to about $10k_BT$ by varying the temperature over a range of about 10°C (*181*, *182*).

Because the self-assembly of micron-sized particles can be observed in real time under an optical microscope, which is not possible for proteins, we can follow their assembly. In addition, we conduct molecular dynamics (MD) simulations, which reproduce the formation process. Taken together, our experimental and computational observations reveal the dynamical processes that

lead to the formation of colloidal chains at different patch ratios, as well as a new dynamical process of surface-catalyzed ring formation of dimer chains (*183–185*).

7.2 Results

7.2.1 Fabrication of Janus particles

The Janus particles we make consist of a polystyrene (PS) patch and a cross-linked 3-(trimethoxysilyl)propyl methacrylate (TPM) patch (*186*). We introduce a fabrication method that combines incompatible elements, PS and TPM, in a cosolvent emulsion droplet and then uses phase separation to create the patches. The technique, which is scalable, is described in Figure 7.1a. The Janus balance χ can be readily controlled from 0.01 to 0.81 by varying the amount of TPM monomer introduced during synthesis relative to the size of the PS particle, as illustrated in Figure 7.1b. The PS surface of PS-TPM particles is selectively functionalized with a di-block copolymer brush using a swelling-deswelling process (*187*). The copolymer brush is terminated with DNA sticky ends by strain-promoted alkyne-azide cycloaddition (SPAAC) (*188*) (Figure 7.1c). The DNA density on the particles is 1 DNA strand per 10.2 nm^2, or 3.2 nm between polymer grafts. The DNA-coated

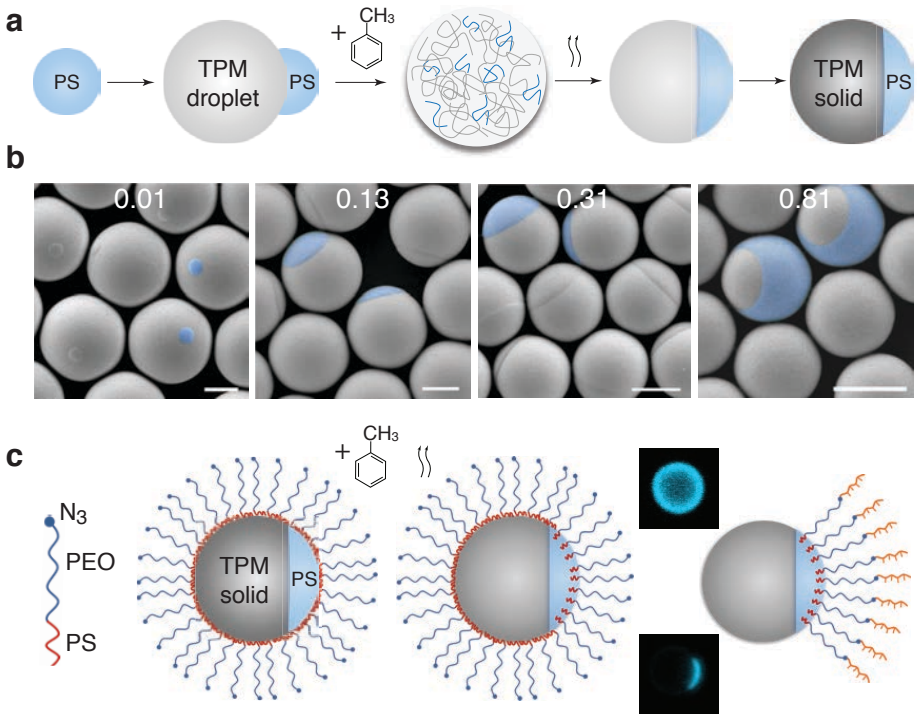


Figure 7.1: Janus particle fabrication. **a,** Left to right: A reactive silicone oil droplet, 3-(trimethoxysilyl)propyl methacrylate (TPM), is heterogeneously grown on the surface of a polystyrene (PS) microsphere by hydrolysis and condensation. Toluene is then added to swell the TPM-grown PS particles, forming fully liquid emulsion droplets in which PS and TPM oligomers are homogenously dissolved. Next, the PS and TPM oligomer droplets phase-separate as toluene gradually evaporates, resulting in a biphasic colloid composed of a PS cap and a TPM oligomer body. Finally, the TPM oligomer body is highly cross-linked by radical polymerization resulting in monodisperse Janus particles composed of solid PS and TPM patches (PS-TPM particles). **b,** SEM images show PS-TPM Janus particles with various patch ratios ranging from 0.01 to 0.81. False color aids identification of the PS patches. Scale bar = 1 μm. **c,** Amphiphilic block copolymers, consisting of a hydrophobic polystyrene (PS) block and a hydrophilic poly(ethylene oxide) (PEO) block with an azide functional end-group (PS-*b*-PEO-N_3), adsorbs to the entire surface of the PS-TPM particle (PS-*b*-PEO-N_3 are not drawn to scale: PS-*b*-PEO-N_3 block is approximately 10 nm thick; particles are approximately 1000 nm in diameter). Adding toluene leads to selective swelling of the PS patch, as the PS patch is not crosslinked, while the TPM is highly crosslinked. The PS block of the PS-*b*-PEO-N_3 to penetrates the swollen PS patch. After deswelling, the PS block is trapped inside the PS patch while the PEO block remains outside. DBCO-functionalized DNA (DBCO-DNA) is coupled to the azide functional group at the end of the PEO brush using SPAAC (*189*). The block copolymer coupled with DNA on the TPM surface is removed by washing with a surfactant solution (Triton X-100, 1-wt%). The

fluorescence images show the DNA-coated Janus particles before (top) and after (bottom) washing, confirming that only the PS patch is coated with DNA, which is fluorescently labeled.

7.2.2 Self-assembly of particles with small patch ratios

First, we present the results for the self-assembly of Janus particles with a patch ratio of $\chi = 0.13$. The bright-field and fluorescent images in Figure 7.2a show the cluster formation of dimers, trimers, and tetramers. Geometrically, the minimum patch ratio for a tetramer is $\chi = 0.092$. No pentamers or higher order clusters are observed, as these are geometrically prevented by the small size of the patches. To form these clusters, we heated the sample above the melting temperature T_m = 53°C. Cooling the sample to 52.0°C, just below T_m, allows the clusters to anneal and equilibrate. At 52.0°C, only singlets and dimers are observed. As the temperature is decreased, trimers form at 50.8°C and then increase in number at 49.6°C. Finally, at 48.4°C, the fraction of trimers decreases as the fraction of tetramers increases, as shown in Figure 7.2a.

The relative populations of singlets, dimers, trimers, and tetramers were measured as a function of temperature and are shown in Figure 7.2a and 7.2b. These are compared to the relative populations obtained from simulations (Figure 7.3, 7.4a, 7.4e and *Section 7.4*), which show the same general trend with the system tending towards trimers and tetramers at low temperature and towards singlets and doublets at higher temperature. The simulations show that to achieve the observed relative changes in the cluster fractions, the ratio $k_B\mathrm{T}/\varepsilon$ must change by more than a factor of 2, from about 0.38 to 0.85. In the experiments, the temperature changes by only a few degrees around a characteristic temperature of 325 K, which means that changes in the interaction strength $\varepsilon = \varepsilon(T)$ must account for the observed changes in cluster populations. In Figure 7.2c, we show histograms of the cluster fractions measured in experiments and simulation at four different temperatures. The values of $k_B\mathrm{T}/\varepsilon$ in simulations are chosen to best match the experimental data for the different temperatures. From this, we see that ε changes from $1.2k_BT$ at 51.2°C to $2.6k_BT$ at

48.4°C, a change that is consistent with the expected temperature dependence of the DNA-mediated attractive interaction(*96*, *181*).

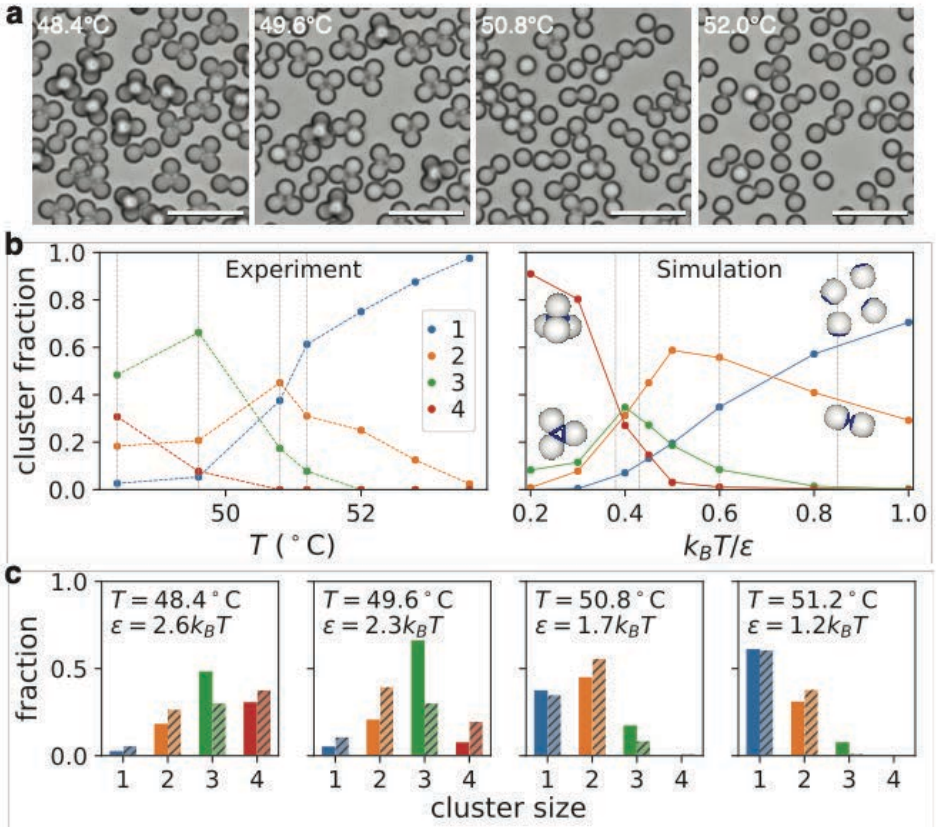


Figure 7.2: Assembly of Janus particles with patch ratio of $\chi = 0.13$. **a,** Bright-field images show self-assembled clusters of Janus particles at various temperatures. Particles disassemble at high temperature and self-assemble at low temperature at the DNA-coated surfaces, forming singlets and clusters of dimers, trimers, and tetramers. Scale bar = 10 μm. **b,** Distribution of singlets and clusters of 2, 3, and 4 particles at different temperatures from experiment (left) and computer simulation (right). Singlets and dimers are dominant at high temperature while trimers and tetramers are dominant at low temperature. Vertical gray lines in left plot correspond to temperatures indicated by vertical gray lines in right plot. **c,** Cluster fractions as a function of temperature from experiment (solid color bars) and computer simulation (hashed) where the depth ε of the attractive effective potential due to DNA binding is selected to best match the experimental measurements. Low temperature favors the formation of larger clusters, which are limited by geometry to four particles for the patch ratio χ=0.13 shown here. The interaction strength ε from simulations for each temperature from experiments is indicated in each plot.

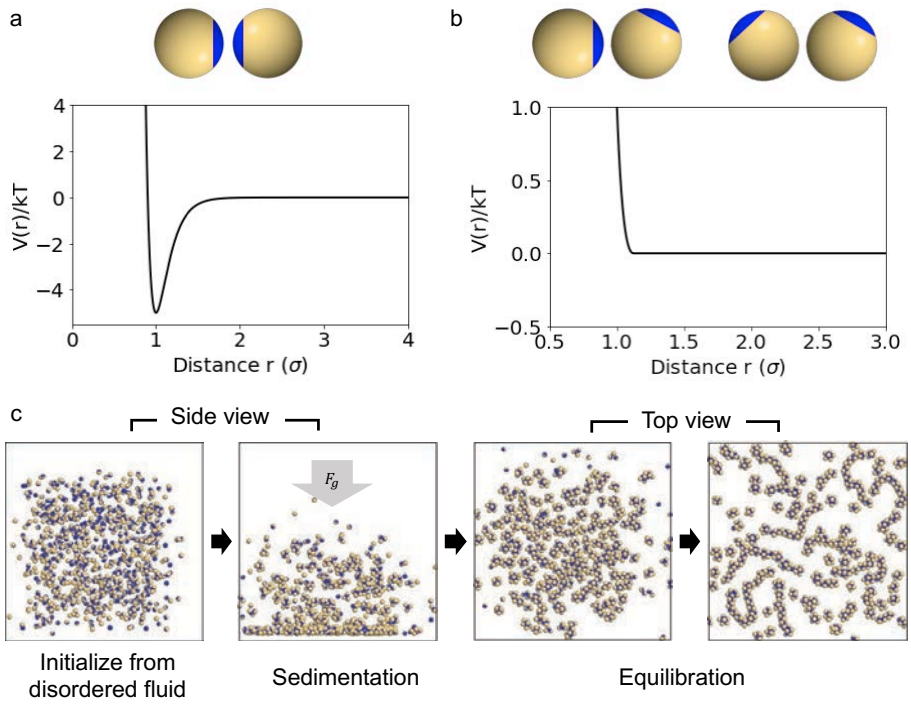


Figure 7.3: Simulation model and self-assembly protocol. (a) Plot for the Morse pair potential applied between patch and patch contact at $T^* = 1.0$. (b) Plot for the WCA pair potential applied between patch and non-patch (yellow) contact and between non-patch and non-patch contact at $T^* = 1.0$. On the top of the plots, examples of each contact are shown, where the blue part is DNA-coated patch region and the yellow part is non-patch region. (c) Simulation protocol for self-assembly of N=1,000 Janus particles with $\chi = 0.35$ at $T^* = 0.7$. From left to right, the snapshots are captured at $10^5\tau$, $5 \times 10^5\tau$, $10^6\tau$ and $10^7\tau$ simulation time steps.

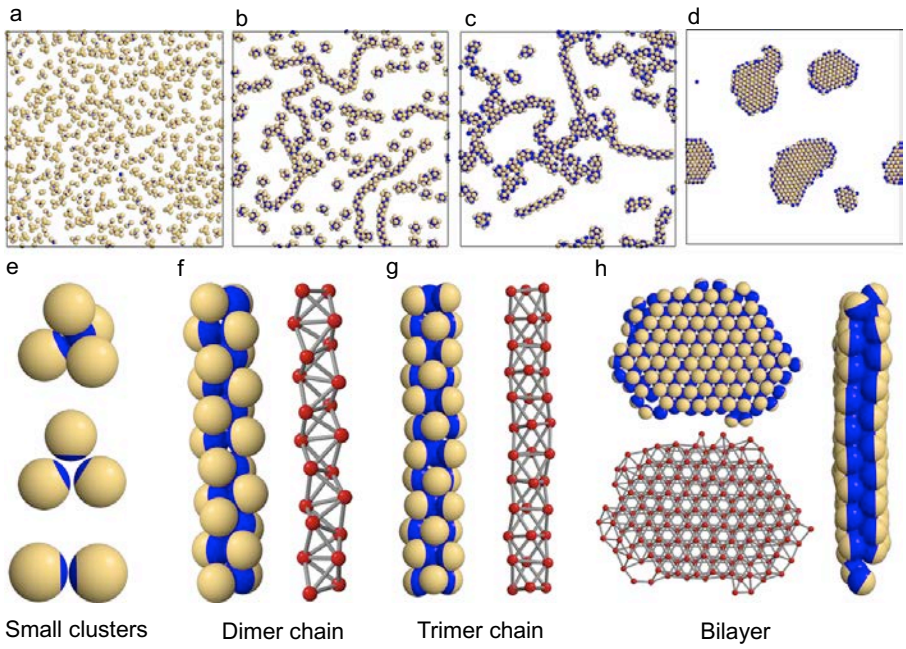


Figure 7.4: Self-assembly results of the MD simulation. (**a-d**) Simulation snapshots ($\tau = 10^8$) of N=1,000 Janus particles with **a**, $\chi = 0.13, \mathrm{T}^* = 0.40$, **b**, $\chi = 0.35, \mathrm{T}^* = 0.70$, **c**, $\chi = 0.475, \mathrm{T}^* = 0.90$ and **d**, $\chi = 0.6, \mathrm{T}^* = 1.25$. Depletion effect is not considered. (e) Small clusters found in $\chi = 0.13$ system. Tetramer (top), trimer (middle) and dimer (bottom). (f) Dimer chain shown in Janus particle representation (left) and ball-and-sticks representations (right) found in $\chi = 0.35$ system. The ball-and-sticks image shows centers of the Janus particles (red) and bonds (grey) with nearest neighbors. (g) Trimer chain shown in Janus particle representation (left) and ball-and-sticks representations (right) found in $\chi = 0.475$ system. (h) Top view (left top and bottom) and side view (right) of bilayer found in $\chi = 0.60$ system.

7.2.3 Cooperative chain formation

DNA-coated PS-TPM particles with larger patch ratios χ can form more than three bonds per particle, which leads to larger structures. Figure 7.5 shows the general trend of self-assembly for particles with patch ratios χ from 0.30 to 0.41. Above the melting temperature T_m = 50°C, particles are well dispersed and diffuse freely. When the temperature decreases to 49°C, the particles start to bind to each other forming clusters of two to seven particles. When the temperature further decreases to around 48°C, we observe mixtures of clusters and short chains.

At 47.2°C, long chain structures, some of them with branches, are formed as the number of clusters decreases. This picture is consistent with simulation results, which show the same trends as the strength of the attractive interaction $\varepsilon(T)$ increases (Figure 7.4b). The formation of chains occurs exclusively by cluster-cluster binding, most frequently by the addition of clusters of 4-7 particles to longer chains, but occasionally by longer chains fusing. Figure 7.5c show the sequential addition of small clusters to a growing chain, which is typical for the growth of dimer chains.

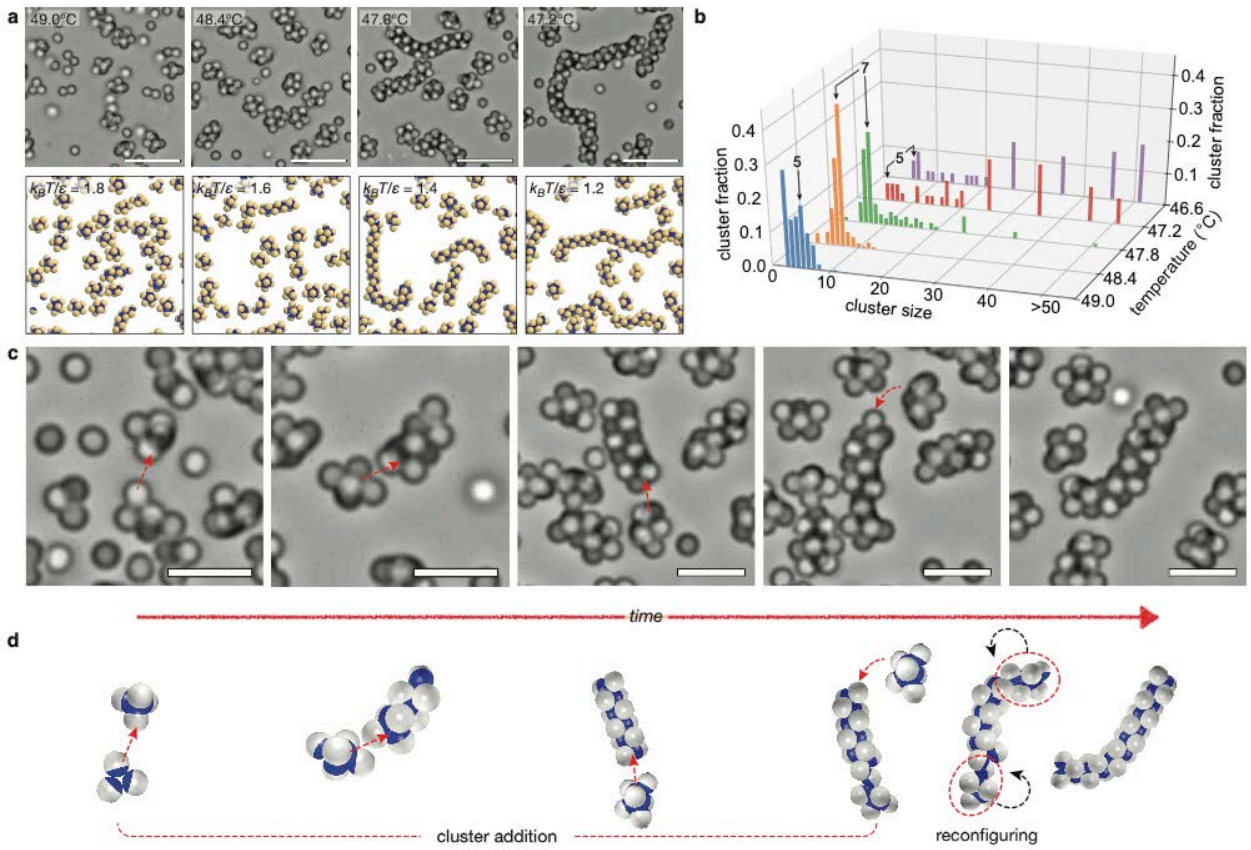


Figure 7.5: Self-assembly of Janus particles to clusters and chains for a patch ratio of 0.3. **a,** Bright-field images above and simulation renderings below show self-assembled structures of Janus particles at different temperatures. Particles self-assemble into small clusters at higher temperature and longer chain-like structures at lower temperature, where the DNA-mediated attraction between patches is greater. **b,** Fraction of clusters in clusters of different sizes as a function of temperature. At higher temperatures, small clusters with 8 or fewer particles dominate. As the temperature is lowered, chains develop and grow longer, mostly by the addition of small clusters but occasionally by fusion of long chains. For cluster sizes greater than

20, the histogram bars integrate over ten or more bins: 21-30, 31-40, 41-50, and >50. **c,** Snapshots and **d,** illustrations show the formation of chain structures by collective polymerization through cluster addition and reconfiguration.

This chain formation process is known as cooperative assembly (*190*), where a chain grows by the addition of clusters rather than by the addition of individual particles (*179*). A key characteristic of cooperative assembly is the coexistence of long chains in equilibrium with small clusters. Indeed, the equilibrium distribution of aggregate sizes, shown in Figure 7.5b, shows that small clusters coexist with chains. Upon closer examination, we see that at 49.0°C, there are only clusters of 8 particles or fewer, with a small excess of pentamers. When the temperature is lowered to 48.4°C, the distribution becomes strongly peaked around 7 particles, with a few chains appearing with a dozen particles or more. Upon lowering the temperature to 47.8°C, the number of small clusters (notably sizes of 6 and 7) decreases, and the number of chains grows. This trend continues as the temperature is decreased, with the growth of successively longer chains while a small but finite fraction of small clusters, with no fewer than 5 particles, always remains.

The microstructure of the chains that are formed from particles with $\chi = 0.30$ to 0.41 is one of alternating dimers, as illustrated in Figure 7.5d, 7.6a, and Figure 7.4f. Each dimer, which is oriented perpendicular to the contour of the chain, serves as a rotational axis for chain bending. Adjacent dimers along the chain have rotation axes oriented 90° with respect to each other, rendering dimer chains very flexible. Each particle in a dimer-chain binds to five other particles.

By contrast, when $\chi = 0.48$, chains consisting of an alternating stack of triangular trimers form, as shown in Figure 7.4c, 7.4g and 7.6b,. These chains are much stiffer than the dimer chains. Like the dimer chains, however, they form from small clusters and grow by cooperative

polymerization, but in this case by a two-stage process, which is very similar to the two-stage process described in simulations of one-patch colloids by Vissers *et al.* (*191*). The chains that initially form are disordered and flexible, and remain so for some time after formation, after which they transform into stiff ordered trimer chains. The transition from disordered to ordered chain is driven by the increased number of bonds that can be formed when the chain orders (*191*). The process can be observed in real time under a microscope and in simulation. Illustrative still frames are shown in Figure 7.6d.

Figure 7.6e show the equilibrium distribution of trimer chains measured at different temperatures. As for the case of dimer chains, long trimer chains coexist with a population of smaller clusters. In this case, however, the distribution of smaller clusters is broader and more disordered than it is for dimer chains. It is also striking how rapidly the population of long trimer chains increases as the temperature decreases. By decreasing the temperature from 48°C to 46°C, the system goes from having only small aggregates of 10 particles or fewer to a situation where roughly 40% of the clusters consist of 50 or more particles. This behavior is consistent with the phase diagram obtained from simulations and shown in Figure 7.6f, namely that small clusters form at the higher temperatures for $\chi < 0.55$ while various chains form at the lower temperatures (Figure 7.5a).

The structure of an alternating stack of triangles is consistent with our simulations and with those of Luitjen (*179*), but in contrast to the experimental observation of Boerdijk-Coxeter (BC) helices of Chen *et al.* for gold-capped hydrophobic-hydrophilic Janus particles with $\chi \sim 1/2$, which are less stable than trimer chains based on free energy calculations (*179*). Chen *et al.* suggested that the BC helices, which also grow by the addition of small clusters, are kinetically trapped and thus do not achieve the equilibrium trimer chain structure. Our chains, however, do reach the

predicted equilibrium structure. Recently, Fejer *et al.* pointed out that the thickness of the gold caps in the experiments of Chen *et al.* is greater at the center of the caps than at the edges, which leads to an orientation-dependent van der Waals interaction between the gold caps (*173*). The analysis of Fejer *et al.* suggests that orientation-dependent interaction between gold caps favors BC helices. Thus, the BC structures may be equilibrium structures after all rather than kinetically-trapped states.

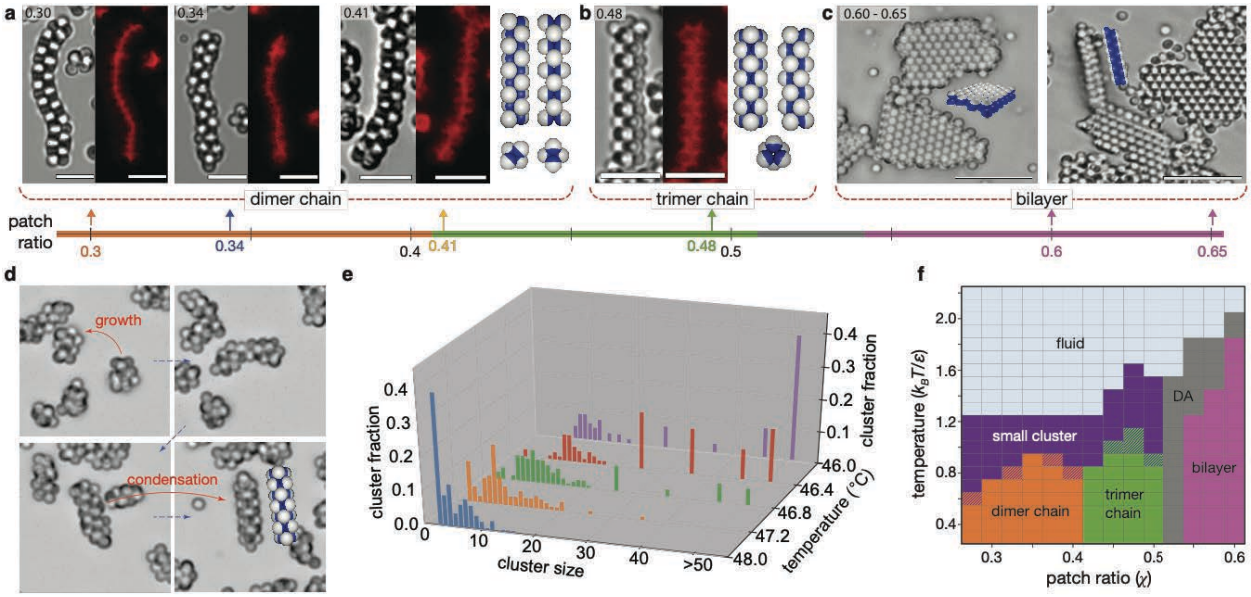


Figure 7.6: DNA-mediated self-assembly of Janus particles with patch ratios from 0.3 to 0.65. **a,** Bright-field and fluorescent images for self-assembled structures of DNA-coated Janus particles with patch ratios from 0.3 to 0.65. Particles with patch ratios from 0.3 to 0.41 self-assemble into flexible dimer chains. The colors on the patch-ratio axis follow those of the phase diagram shown in frame f. **b,** Particles with a patch ratio of 0.48 self-assemble into stiff densely-packed trimer chains. **c,** Particles with patch ratios of 0.6 and 0.65 self-assemble into bilayer structures. The right-most image shows a cross-section of a bilayer structure. **d,** Still frames showing the assembly of a trimer chain. Disordered chains form and grow by combining with clusters, and subsequently condense into trimer chains. **e,** Fraction of clusters that are found in clusters of different size. For cluster sizes greater than 20, the histogram bars integrate over ten or more bins: 21-30, 31-40, 41-50, >50. **f,** Phase diagram as a function of patch ratio and temperature, obtained by computer simulation (DA: Disordered aggregates).

7.2.4 Bilayers

When the patch ratios are increased to 0.60 and 0.65, the Janus particles self-assemble into bilayers, as shown in Figure 7.6c, which agrees with our simulations and those of others (Figure 7.4d and 7.4h) (*174*, *176*). Initially, particles form amorphous aggregates that gradually rearrange into hexagonally-packed bilayers upon annealing. Once they form bilayers, growth occurs at the edges of pre-existing bilayers. Top- and side-views of bilayers in Figure 7.6c confirm that the particles assemble into two layers of an ordered hexagonal array. Figure 7.6f shows the phase diagram determined from our simulations for Janus particles as a function of the dimensionless temperature k_BT/ε *vs.* patch ratio χ. The phase diagram is consistent with our experiments showing the same progression of structures occurring at patch ratios consistent with our experiments.

7.2.5 Surface-catalyzed formation of rings

One of our most striking observations is the formation of substrate-catalyzed rings of dimer chains when a weak attractive depletion interaction is introduced. It is well known that polymer micelles can induce a depletion attraction between particles (*192*, *193*) and between particles and substrates (*194*). We empirically selected 1.0% w/w of polymeric surfactants (F127 from BASF Corp.), above its critical micelle concentration (<< 0.1 % w/w around our annealing temperatures), as this introduces a weak depletion attraction that avoids depletion-mediated bulk crystallization and leaves the DNA-mediated attractive interaction as the primary force driving self-assembly. For a patch ratio of $\chi = 0.30$, dimer chains are observed, just as for the case without depletion. The most significant effect of adding a depletion interaction is that it pushes dimer chains against the walls of the sample container, as the depletion attraction between spheres and a flat substrate is twice as strong as between spheres (*194*). Geometrically, exactly half of the particles in a perfectly straight dimer chain can touch and adhere to a flat substrate via the depletion interaction (Figure

7.7e). Doing so maximizes the depletion attraction and lowers the free energy of the system. When a dimer chain bends, the geometrical constraints of a dimer chain dictate that the number of particles that can touch the substrate must decrease unless the radius of curvature of the chain is constant over the entire length of the chain, as illustrated in Figure 7.7e. Thus, a fluctuating dimer chain minimizes its free energy by adopting a circular C-shape of constant curvature when interacting with a flat substrate under the influence of the depletion interaction. This is exactly what we observe in both experiments and simulations, as illustrated in Figure 7.7a and 7.8. The strength of the depletion attraction is difficult to determine experimentally, but rings form in the simulations when the strength of the depletion attraction is 5% that of the DNA. The constraint of constant curvature leads to the formation of closed rings, also shown in Figure 7.7a, which are stable indefinitely. By contrast, rings are never observed in the absence of a depletion interaction. This reflects the entropic barrier to ring formation in the bulk suspension. However, restricting conformations to 2D structures with constant curvature reduces entropy to the point that the energetic advantage of forming rings overcomes the entropic barrier and thus catalyzes ring formation. Once the chain overcomes the barrier, rings persist even in the absence of the depletion effect.

We observe rings of varying size and number of particles. Geometrically, the smallest ring that can form without spheres overlapping consists of 32 particles, which is indeed what we observe. Larger rings may form, with the constraint that the total number of particles be divisible by four, as the dimers must come in pairs so the rings can close. We observe rings of 32, 36, 40, 44, and 48 particles in both experiment and simulation, which is consistent with the geometrical reasoning.

Rings are also observed for Janus particles with $\chi = 0.34$. Rings of 32 and 36 particles form without defects. However, defects generally appear in larger rings with 40, 44, and 48 particles. The defects are localized and consist of 6 particles in the chain adopting an octahedral conformation, as illustrated by the schematic of a ring with a single defect in Fig 5f. Such defects are stiff, which reduces the flexibility of the chain and thus reduces its conformational entropy. At the same time, each defect results in an additional interparticle bond, which lowers the energy of the chain, compensating for the entropic cost. The octahedral defects proliferate in the larger rings, which are more flexible and have more conformational entropy than the smaller rings. The defects are highlighted in the rings with 40, 44, and 48 particles in Figure 7.7f.

These octahedral defects can proliferate along a chain for patch ratios from 0.34 to 0.41 to the point that the chain consists of almost nothing but octahedral defects, effectively creating a new type of chain, which we call a compact chain (Figure 7.7b). We also observe the same depletion-induced transition in our simulations. For $\chi = 0.34$, dimer chains coexist with compact chains, often with a few dimer-chain defects, which make flexible kinks between the stiff compact chain segments (Figure 7.7a). Bright field and fluorescent images of compact chains for $\chi = 0.34$ and 0.41 are shown in Figure 7.7a and 7.7b along with illustrations of the compact chain. In fact, the situation is subtler than this. Just as dimer chains can exist with a few stiff octahedral defects, so too can compact chains exist with a few flexible dimer-chain defects. These dimer defects reveal themselves as kinks in otherwise stiff compact chains.

Figure 7.7g shows the phase diagram obtained from our simulations for the case where there is a weak depletion interaction and a flat surface. Once again, it is consistent with our experiments showing the same progression of phases with transitions between phases. Note that, in the region of the ring phase, our experiments show rings coexisting with dimer and compact

chains because not every chain can overcome the kinetic and energetic barrier to transform into the rings within the 24-hour duration of experiments.

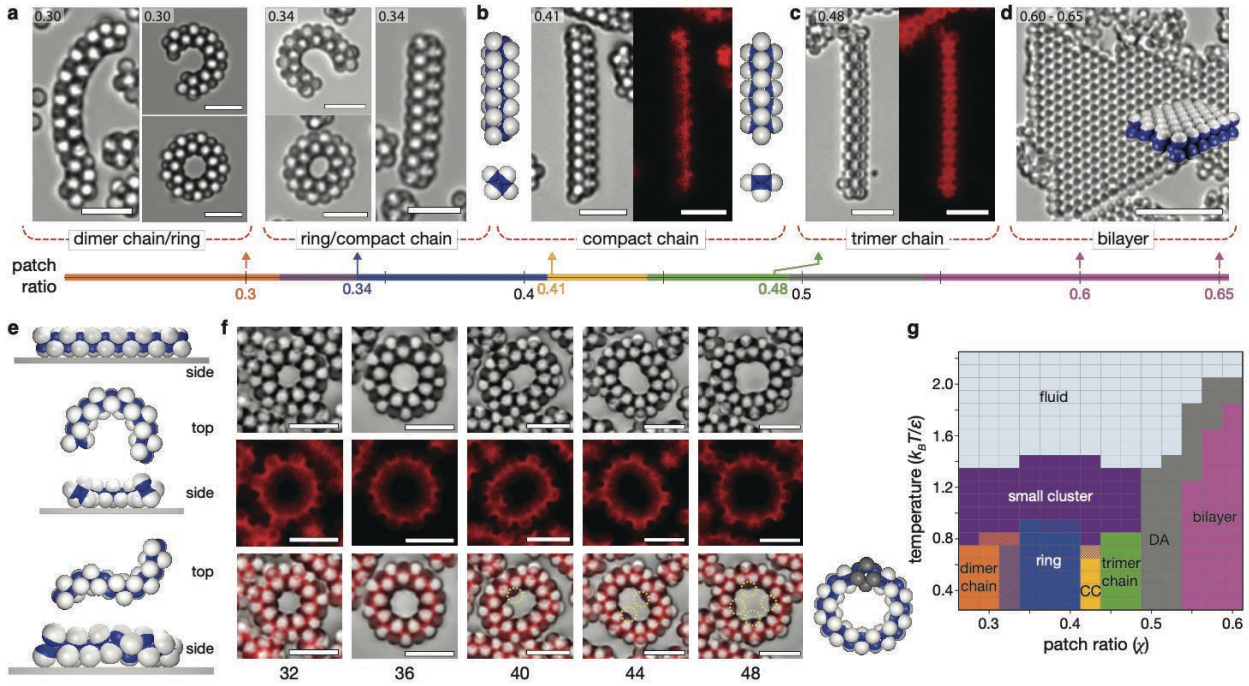


Figure 7.7: DNA-mediated self-assembly of Janus particles in the presence of depletion force. **a-c,** Optical microscope images for self-assembled structures of DNA-coated Janus particles with various patch ratios between χ=0.3 and 0.65 in the presence of 1% of F127, which forms micelles and introduces a weak attractive depletion interaction. **a,** When χ = 0.3, particles self-assemble into dimer chains as well as ring structures, which form only when there is a depletion interaction. **b,** For χ = 0.34, the same structures are observed, as well as densely packed chains which we call compact chains (CC). For χ = 0.41, most chains are compact chains. **c,** However, the structures of trimer chains (χ = 0.48) and **d,** bilayers (χ = 0.6 and 0.65) are unchanged. **e,** Half of the particles of a dimer chain can touch the substrate if the chain is straight (top) or has a constant radius of curvature (middle). However, when the dimer chain curvature is not constant, as for an S-shaped chain (bottom), one or more particles in the chain detach from a flat substrate. **f,** Bright field (top) and fluorescent (middle) optical microscope images of rings consisting of 32-48 particles. Yellow dashed circles in bright field/fluorescent overlay (bottom) show the locations octahedral defects that appear for χ = 0.34. Illustration on the right shows the structure of a single octahedral defect. **g**, Phase diagram of Janus particles obtained by simulation in the presence of a depletion interaction (ε_{dep} = 0.1).

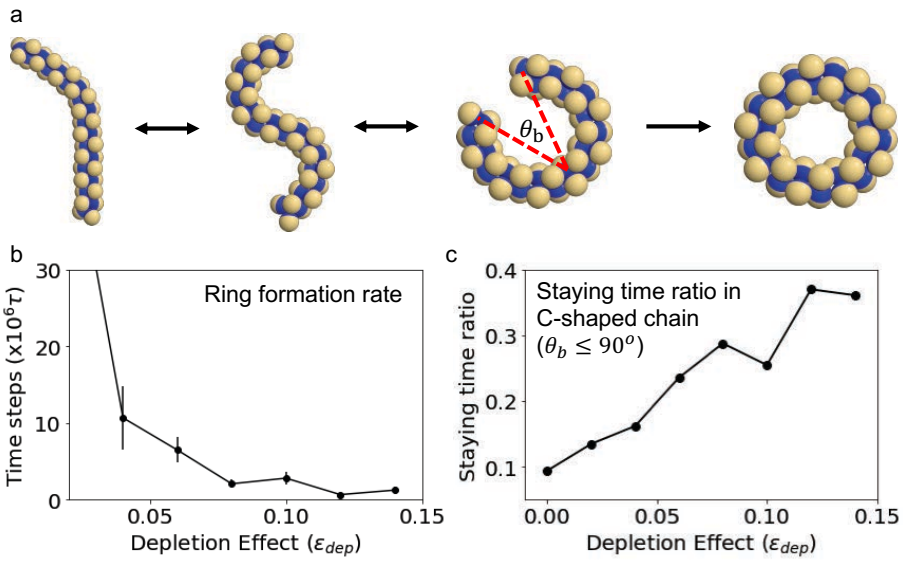


Figure 7.8: Ring formation in the MD simulation. (a) Tracking the formation process of ring in $\chi = 0.35$ system at $T^* = 0.55$ with $\varepsilon_{dep} = 0.1$ depletion effect, initialized from a flexible dimer chain. The dimer chain fluctuates between S-shaped and C-shaped chain and eventually forms a closed ring chain. (b) Ring formation rate depends on the strength of depletion effect (ε_{dep}), averaged by 10 different runs. As the depletion effect increases, the ring formation rate becomes faster. When there is no depletion effect, ring formation was not observed within the simulation time ($\tau = 10^8$). (c) The ratio of staying time in C-shaped chain before ring-formation, averaged by the 10 different runs. When the bent angle (θ_b) of the chain (third image in **a**) is less then 90°, we considered the chain is in C-shape. As depletion effect increases, the staying time ratio in the C-shape increases.

7.3 Discussion

Despite their simplicity, colloidal spheres with tunable Janus balance exhibit some of the most basic features in cooperative self-assembly, which is also thought to be active in biological molecules like proteins, although the evidence is indirect (*195*). From our observations, we can draw several lessons.

Small aggregates are self-limiting if the Janus balance is small but can serve to nucleate and grow fibers when the Janus balance reaches about 0.3. Thus, a small increase in the Janus balance can lead to the growth of long fibers.

Fiber growth generally occurs by cooperative polymerization, where fibers grow not by adding single particles but by adding small particle aggregates.

The chain structure is sensitive to the Janus balance, with different types of fibers forming when the Janus balance changes from about 0.4 to 0.47. Moreover, the dynamical pathways for fiber formation depend sensitively on the Janus balance. The slow two-step process by which trimer chains grow when the Janus balance is about 0.47 becomes a rapid, single-step process for dimer chains when the Janus balance falls to about 0.4.

We emphasize that all the structures formed by our DNA-coated Janus particles are equilibrium structures. This follows from the ability of the DNA-coatings we use to thermally anneal, as discussed and demonstrated in previous publications (*180*, *188*). This is in contrast to the nonequilibrium structures reported by Kang *et al.* (*196*, *197*) and possibly to those reported by Chen *et al*. (*179*)

Finally, we find that introducing a depletion interaction, which is the colloidal analogue of molecular solvation interactions, can have profound effects. First, it can alter the structure of the fibers that form, as illustrated by the formation of stiff compact chains instead of flexible dimer chains when $\chi \sim 1/3$. Second, it can catalyze entirely new and unanticipated structures. The dimer-chain rings we observe illustrate how interaction with a substrate can impose constraints on a polymer that favor certain configurations that otherwise would be very unlikely. This is reminiscent of ring-like amyloid protein aggregates, which can adopt different conformations when bound to a substrate (*167*), and may have high toxicity due to their topology (*198*). By using colloids, we can follow the dynamics of the structural evolution. DNA-coated Janus colloids thus provide a simple but effective model of fiber formation and interactions with substrates, which

underscores their utility as a model system for exploring the dynamics of structure formation at a level that remains difficult to achieve with biological macromolecules.

7.4 Method and supporting information

7.4.1 Molecular dynamics simulation method

Molecular dynamics (MD) simulations are performed to simulate the self-assembly behavior of the DNA coated Janus particles. The MD method implemented in the HOOMD-blue simulation package (*9*, *55*) is available as open source, which can be downloaded at http://glotzerlab.engin.umich.edu/hoomd-blue/. HOOMD-blue runs both on central processing units (CPUs) and graphics processing units (GPUs) and is optimized for parallel computing with multiple CPUs and GPUs using MPI domain decomposition. Most simulations of the present work were run on Intel Xeon E5-2680v3 CPUs and NVIDIA Tesla K20X GPUs.

A minimal model is developed to simulate a large number of Janus particles in reasonable timescale. The Weeks-Chandler-Andersen (WCA) potential (*121*) is applied on whole sphere to create a rigid core, and the Morse pair potential(*199*) is applied on the patch region to reproduce the self-complementary DNA hybridization (Figure 7.3a and 7.3b). Note that, we used the Morse pair potential in this study because it is convenient to control the range of attraction, not physical reason. The depletion effect caused by surfactants is modeled in implicit way by applying the Lennard-Jones (LJ) pair potential (*200*) on the Janus particles.

The MD simulations were performed with N = 1000 particles for most self-assembly systems, which is large enough size to form multiple superstructures which are usually composed of dozens of particles, but small enough to form those structures in a reasonable timescale. Typically, 48 GPU hours are taken for 10^8 MD timesteps (dt = 0.005) with Brownian dynamics

for each state point. We used periodic boundary conditions in two dimensions and one non-periodic boundary for creating a bottom where particles sediment and assemble. Gravitational force (F_G) is applied toward the bottom, which is converted from the experimental gravitational height ($k_BT/F_G \approx 0.3\ \mu$m) (*201*) of the present work (Figure 7.3c). Packing density (ϕ) is kept constant as 0.1 for all system, which is approximately calculated by $N \times v_0/V_{assemble}$, where N and v_0 are the number of particles and volume of a particle, respectively. The volume of the assembly space ($V_{assemble}$) is calculated by multiplying an assembled height ($\sim 2\sigma$) and two other simulation box lengths, where σ is the unit length of simulation and the diameter of a Janus particle.

7.4.2 Simulation model of DNA-coated Janus particles

DNA hybridization between patches is modeled by Morse potential (Figure 7.3a),

$$V_{\mathrm{P}} = D_0[\exp\big(-2\alpha(r - r_0)\big) - 2\exp\big(-2\alpha(r - r_0)\big)] \qquad \text{(Eq. 7.1)}$$

where D_0 and α is depth and width of the potential well with the unit of ε and $1/\sigma$, respectively, and r_0 is position of the potential minimum with σ unit. The ratio (χ) of DNA-patch is angularly modulated (*202*) as follow:

$$V_{\mathrm{AM\_P}} = \frac{V_{\mathrm{P}}}{(1+\exp\big(-w(\hat{n}_i \cdot \hat{r}_{ij} - \chi)\big))(1+\exp\big(-w(\hat{n}_j \cdot \hat{r}_{ji} - \chi)\big))} \qquad \text{(Eq. 7.2)}$$

where $\hat{n}_i$ is unit vector from particle *i* center to its patch center, $\hat{r}_{ij}$ is the center-to-center unit vector of particle *i* and *j*, and *w* is the angular sharpness ranging from 0 to infinity. When *w* is infinity, this potential is identical to the Kern-Frenkel model (*203*). The particle contacts via non-patch regions are modeled by the Weeks-Chandler-Andersen (WCA) potential (*121*) (Figure 7.3b) to simulate volume exclusion :

$$V_{\mathrm{NP}}(r) = \begin{cases} 4\varepsilon_{\mathrm{NP}}\left[\left(\frac{\sigma_{\mathrm{NP}}}{r}\right)^{12} - \left(\frac{\sigma_{\mathrm{NP}}}{r}\right)^{6}\right] - 4\varepsilon_{\mathrm{NP}}\left[\left(\frac{\sigma_{\mathrm{NP}}}{r_{\mathrm{cut}}}\right)^{12} - \left(\frac{\sigma_{\mathrm{NP}}}{r_{\mathrm{cut}}}\right)^{6}\right}, & r < r_{\mathrm{cut}} \\ 0, & r \geq r_{\mathrm{cut}} \end{cases} \quad \text{(Eq. 7.3)}$$

where $r_{\mathrm{cut}} = \sigma_{\mathrm{NP}} \times 2^{1/6}$. The depletion attraction (V_{dep}) between monopatch particles induced by depletants is implicitly modeled (*200*) via the Lennard-Jones potential:

$$V_{\mathrm{dep}}(r) = 4\varepsilon_{dep}\left[\left(\frac{\sigma_{\mathrm{dep}}}{r}\right)^{12} - \left(\frac{\sigma_{\mathrm{dep}}}{r}\right)^{6}\right] \quad \text{(Eq. 7.4)}$$

The ε_{dep} can vary depending on the concentration of depletant. The combined effect of depletion attraction and patch or non-patch contacts is achieved by summation. The potential parameters are listed in the Table below.

Table 7.1: Potential parameters for the MD simulation. Angular modulation parameter (w) is 20.0 for all system.

Morse potential (V_{P})		WCA potential (V_{NP})		LJ potential (V_{dep})	
Parameter (unit)	Value	Parameter (unit)	Value	Parameter (unit)	Value
D_0 (ε)	5.0	$\varepsilon_{\mathrm{NP}}$ (ε)	1.0	$\varepsilon_{\mathrm{dep}}$ (ε)	0.0 – 0.14
α ($1/\sigma$)	7.0	σ_{NP} (σ)	1.0	σ_{dep} (σ)	1.0
r_0 (σ)	1.0				

7.4.3 Simulation protocol

For the self-assembly simulations, each run follows the following steps: (i) Locate the Janus particles randomly in 3-D simulation box without overlapping between particles and give a random orientation to each particle. (ii) To create a gravitational effect, apply a constant force (F_G) uniformly to the whole system toward a uniform direction. Note that the simulation box should be non-periodic along the F_G direction to sediment particles, but periodic to the two other directions. (iii) Running the system at constant temperature (T*) by checking the potential energy of the whole

system in each $10^3\tau$ to make sure that the system is equilibrated. Typically, $10^8\tau$ was enough to equilibrate the system. Snapshots for each stage are shown in Figure 7.3c.

For investigating the depletion effect on the transformation of dimer chain into the ring and the compact chain, a simulation is initialized from an ideally constructed dimer chain (N=36) and the depletion effect is applied to both particles and bottom. By varying strength of the depletion effect ($\varepsilon_{\mathrm{dep}}$= 0 – 0.14, $\Delta\varepsilon_{\mathrm{dep}}$=0.01) and the patch ratio, we check the results observed within the $10^8\tau$ simulation time. The temperature (T*) is kept constant for each state point.

For the phase diagram (Figure 7.6f and 7.7g), each superstructure (dimer chain, trimer chain and bilayer) is built and initialized from a low enough temperature ($T^* \leq 0.3$) to keep the structure stable. Then, we gradually increase the temperatures and obtained the melting temperature of each phase. In the presence of the depletion effect (Figure 7.7g), the ring phase and the compact chain (CC) phase are determined when the initial dimer chain transform into the structure.

CHAPTER 8

Conclusions and Outlook

8.1 Summary of results

In summary, this dissertation explored various assembly behaviors of hard particles (Chapters 2, 3, 4) and DNA-programmable colloid (Chapters 5, 6, 7) systems. Assembly of hard particle systems is driven by entropy maximization with particle shape, and DNA-programmable colloids assemble by maximizing the hybridization of DNA ligands on the particle surface. Within the two topics, each study provided a fundamental understanding of particle attributes that induce the complex phase behavior found in colloidal systems.

In the first study (Chapter 2), we showed three instances of two-step crystallization of hard particle fluids, where interaction is dictated solely by building-block shape and thus entropy. Crystallization in these systems proceeds via a high-density precursor fluid phase with prenucleation motifs in the form of clusters, fibers and layers, and networks, respectively. We categorize our two-step crystallization pathways based on the dimension of the prenucleation motifs and discuss possible comparisons to other crystallization processes. The crystals that form are complex, including, notably, a crystal with 432 particles in the cubic unit cell. Our results establish the existence of complex crystallization pathways in entropic systems and reveal prenucleation motifs of various dimensions.

In the second study (Chapter 3), we demonstrated entropy compartmentalization of single component hard particle systems. This work showed that the truncation of the hard triangular bipyramids creates a cavity at the center of a cage-like motif, and the cavity can be occupied by guest particles if the size of the cavity is comparable to that of the guest. Using free energy calculations, we showed that crystallization occurs via entropy compartmentalization, where the entropy of the system is decomposed into low entropy (host particles) and high entropy (guest particles) subsystems. In addition, we identified three different rotational motions of the guest (free rotation, rotation around a fixed axis and quantized rotation), which depends on the shape correlation between the guest and the cavity.

In the third study (Chapter 4), we demonstrated that the formation of fivefold and icosahedral twin clusters could be entropically engineered. The thermodynamic stability of cubic and hexagonal diamond crystals can be entropically controlled by the proper design of edge and vertex truncation of the hard tetrahedra. When the particle shape is designed to have a negligible free-energy difference between the cubic and the hexagonal diamonds, the formation of twin boundaries or stacking faults is easily induced, which enables the entropy-driven formation of fivefold and icosahedral twin crystals. We showed that the fivefold and icosahedral twin clusters could be entropically stabilized in a dense fluid due to a strong fluid-crystal interfacial tension.

In the fourth study (Chapter 5), we showed that DNA-programmed triangular bipyramids self-assemble into clathrate architectures, one of the most complex structures made via programmable assembly. Their formation can be understood based on the shape of the nanoparticle building blocks and the maximization of DNA hybridization because the particle shape is allowed to form a tetrahedral network by maximizing face-to-face contacts. We developed a simple and efficient DNA-shell model by appropriately arranging patches, which interact via an effective pair

potential, on the surface of the particles. The simulation model successfully reproduced the formation of clathrates and allowed us to analyze structural details that could not be done in experiments. This work suggests that highly complex crystals that have not been reported in colloidal systems could be obtained by a proper combination of particle shape and DNA linker.

In the fifth study (Chapter 6), we demonstrated that the anisotropic distribution of electron equivalents in colloidal crystals engineered with DNA can be used to make low symmetry crystalline phases. We showed that the strength of local DNA bonding interactions can be regulated by two means: temperature and the EE-PAE sticky end ratio. Under appropriate conditions, the spatial distribution of the EEs breaks the symmetry of isotropic PAEs, leading to a set of well-defined coordination geometries and access to low symmetry crystalline phases, including a triple double-gyroid structure that has no known natural equivalent. This symmetry breaking strategy relies on the establishment of the new concept of "valence electrons" in the PAEs and EEs that comprise colloidal crystals, which could be used to discover more low symmetry and complex colloidal architectures.

In the sixth study (Chapter 7), we showed the formation of various superstructures, including colloidal micelles, chains and bilayers, by changing the size of an attractive patch of Janus colloids. When the Janus balance is small enough, small aggregates are formed. Once the Janus balance becomes larger than 30%, dimer chains are formed through cooperative polymerization. On the other hand, the formation of trimer chains, which is found when the Janus balance is larger than 45%, follows a two-stage process, first by cooperative polymerization into disordered aggregates followed by condensation into more ordered stiff trimer chains. Introducing substrate binding through depletion catalyzes dimer chains to form nonequilibrium rings that otherwise do not form. This work provides a simple but effective model of DNA-programmable

colloids that can form rich morphologies by simply changing the size of the DNA-coated area of the particle.

8.2 Concluding remarks

This dissertation studies the use of hard particles with shape anisotropy and DNA-programmable techniques to find interesting assembly behavior of colloids. We have strived to show that diverse material classes, which have been found only in molecular systems, are accessible in colloidal systems via the appropriate design of colloidal particles. Not only that, as we conducted this work, we found a rich set of new behaviors in colloidal crystals, which are not possible to realize in molecular systems. We believe that this dissertation will contribute a new viewpoint to the study of colloidal assembly.

This dissertation is mainly focusing on fundamental understanding of assembly behavior driven by hard particles and DNA-programmable colloid systems. We showed that complex phase behavior, which has been found only in atomic or molecular systems, can be accessible at nanoscale using colloidal systems via the appropriate design of colloidal particles. These behaviors include multi-step crystallization, compartmentalization of entropy, multiple twinning, the formation of complex open-caged structures, stabilization of crystals through mobile particles, and formation of self-limiting structures like fibers and bilayers. Our results have two implications. First, entropic bonds can be as effective as traditional chemical bonds in producing complex assembly behaviors and crystal structures (Chapters 2, 3, 4). This means that even in a system with little energetic interaction between particles, sufficiently diverse structural and dynamic properties can be expected. Second, the tunability of colloidal particles will allow us to find a rich set of new colloidal crystal applications (Chapters 5, 6, 7), which are not possible to realize in molecular systems. Nanosize cavities can be used to store biological materials for biomedical applications,

and the interparticle distances of colloidal crystals (~ a few hundreds of nanometers) can be used for optical applications. The design phase space of colloidal particles is now highly multi-dimensional; thus, it is very challenging to decide on particle design direction to target a specific property in experiments. The simulation models and approaches suggested in this dissertation will provide effective guidance on designing colloidal particles to realize the desired properties.

BIBLIOGRAPHY

1. G. van Anders, D. Klotsa, N. K. Ahmed, M. Engel, S. C. Glotzer, *Proc. Natl. Acad. Sci.* **111**, E4812–E4821 (2014).
2. E. S. Harper, G. van Anders, S. C. Glotzer, *Proc. Natl. Acad. Sci.* **116**, 16703–16710 (2019).
3. D. Frenkel, *Nat. Mater.* **14**, 9–12 (2015).
4. P. F. Damasceno, M. Engel, S. C. Glotzer, *Science (80-.).* **337**, 453–457 (2012).
5. P. K. Bommineni, N. R. Varela-Rosales, M. Klement, M. Engel, *Phys. Rev. Lett.* **122**, 128005 (2019).
6. M. R. Jones, N. C. Seeman, C. A. Mirkin, *Science (80-.).* **347**, 1260901–1260901 (2015).
7. N. P. King *et al.*, *Science (80-.).* **336**, 1171–1174 (2012).
8. L. Onsager, *Ann. N. Y. Acad. Sci.* **51**, 627–659 (1949).
9. J. Glaser *et al.*, *Comput. Phys. Commun.* **192**, 97–107 (2015).
10. X. Ye *et al.*, *Nat. Commun.* **6**, 10052 (2015).
11. A. Dreyer *et al.*, *Nat. Mater.* **15**, 522–528 (2016).
12. W. Liu *et al.*, *Science (80-.).* **351**, 582–586 (2016).
13. R. J. Macfarlane *et al.*, *Science (80-.).* **334**, 204–208 (2011).
14. M. N. O'Brien, M. R. Jones, B. Lee, C. A. Mirkin, *Nat. Mater.* **14**, 833–839 (2015).
15. C. Knorowski, S. Burleigh, A. Travesset, *Phys. Rev. Lett.* **106**, 215501 (2011).
16. T. I. N. G. Li, R. Sknepnek, R. J. Macfarlane, C. A. Mirkin, M. Olvera de la Cruz, *Nano Lett.* **12**, 2509–2514 (2012).
17. S. Lee, E. G. Teich, M. Engel, S. C. Glotzer, *Proc. Natl. Acad. Sci.*, 201905929 (2019).
18. P. F. Damasceno, M. Engel, S. C. Glotzer, *ACS Nano*. **6**, 609–614 (2012).
19. S. Auer, D. Frenkel, *Nature*. **409**, 1020–1023 (2001).
20. U. Gasser, E. R. Weeks, A. Schofield, P. N. Pusey, D. A. Weitz, *Science (80-.).* **292**, 258–262 (2001).

21. P. Rein ten Wolde, D. Frenkel, *Science (80-.).* **277**, 1975–1978 (1997).

22. P. G. Vekilov, *Cryst. Growth Des.* **4**, 671–685 (2004).

23. S. Sastry, C. Austen Angell, *Nat. Mater.* **2**, 739–743 (2003).

24. C. Desgranges, J. Delhommelle, *J. Am. Chem. Soc.* **133**, 2872–2874 (2011).

25. G. Franzese, G. Malescio, A. Skibinsky, S. V. Buldyrev, H. E. Stanley, *Nature*. **409**, 692–695 (2001).

26. F. Smallenburg, L. Filion, F. Sciortino, *Nat. Phys.* **10**, 653–657 (2014).

27. V. Agarwal, B. Peters, (2014; http://doi.wiley.com/10.1002/9781118755815.ch03), pp. 97–160.

28. F. Smallenburg, F. Sciortino, *Phys. Rev. Lett.* **115**, 015701 (2015).

29. D. Gebauer, A. Völkel, H. Cölfen, *Science (80-.).* **322**, 1819–1822 (2008).

30. D. Zahn, *ChemPhysChem*. **16**, 2069–2075 (2015).

31. J. J. De Yoreo *et al.*, *Science (80-.).* **349**, aaa6760 (2015).

32. T. Li, D. Donadio, L. M. Ghiringhelli, G. Galli, *Nat. Mater.* **8**, 726–730 (2009).

33. A. Haji-Akbari, P. G. Debenedetti, *Proc. Natl. Acad. Sci.* **114**, 3316–3321 (2017).

34. A. Dey *et al.*, *Nat. Mater.* **9**, 1010–1014 (2010).

35. H. D. Nguyen, C. K. Hall, *Proc. Natl. Acad. Sci.* **101**, 16180–16185 (2004).

36. L. C. Jacobson, W. Hujo, V. Molinero, *J. Am. Chem. Soc.* **132**, 11806–11811 (2010).

37. M. Lauricella, S. Meloni, N. J. English, B. Peters, G. Ciccotti, *J. Phys. Chem. C*. **118**, 22847–22857 (2014).

38. H. Lin *et al.*, *Science (80-.).* **355**, 931–935 (2017).

39. M. Lenz, T. A. Witten, *Nat. Phys.* **13**, 1100–1104 (2017).

40. G. Bergman, J. L. T. Waugh, L. Pauling, *Acta Crystallogr.* **10**, 254–259 (1957).

41. N. Jaussaud *et al.*, *Solid State Sci.* **5**, 1193–1200 (2003).

42. Y. G. Kuznetsov, A. J. Malkin, A. McPherson, *J. Cryst. Growth*. **232**, 30–39 (2001).

43. F. S. Bates, P. Wiltzius, *J. Chem. Phys.* **91**, 3258–3274 (1989).

44. H. Jiang, P. G. Debenedetti, A. Z. Panagiotopoulos, *J. Chem. Phys.* **150**, 124502 (2019).

45. D. James *et al.*, *J. Chem. Phys.* **150**, 074501 (2019).

46. J. Vatamanu, P. G. Kusalik, *Phys. Chem. Chem. Phys.* **12**, 15065 (2010).

47. J. C. Palmer *et al.*, *Nature*. **510**, 385–388 (2014).

48. J. Israelachvili, *Langmuir*. **10**, 3774–3781 (1994).

49. J. R. Davis, A. Z. Panagiotopoulos, *J. Chem. Phys.* **129**, 194706 (2008).

50. A. F. Wallace *et al.*, *Science (80-.).* **341**, 885–889 (2013).

51. T. Eichner, S. E. Radford, *Mol. Cell*. **43**, 8–18 (2011).

52. H. E. Stanley, O. Mishima, *Nature*. **396**, 329–335 (1998).

53. E. R. Chen, D. Klotsa, M. Engel, P. F. Damasceno, S. C. Glotzer, *Phys. Rev. X*. **4**, 011024 (2014).

54. J. A. Anderson, M. Eric Irrgang, S. C. Glotzer, *Comput. Phys. Commun.* **204**, 21–30 (2016).

55. J. A. Anderson, C. D. Lorenz, A. Travesset, *J. Comput. Phys.* **227**, 5342–5359 (2008).

56. W. F. Claussen, *J. Chem. Phys.* **19**, 259–260 (1951).

57. E. D. Sloan Jr., C. A. Koh, C. A. Koh, *Clathrate Hydrates of Natural Gases* (CRC Press, 2007; https://www.taylorfrancis.com/books/9781420008494).

58. J.-A. Dolyniuk, B. Owens-Baird, J. Wang, J. V. Zaikina, K. Kovnir, *Mater. Sci. Eng. R Reports*. **108**, 1–46 (2016).

59. H. Lin *et al.*, *Science (80-.).* **355**, 931–935 (2017).

60. W. L. Mao, *Science (80-.).* **297**, 2247–2249 (2002).

61. E. V. Shevchenko, D. V. Talapin, N. A. Kotov, S. O'Brien, C. B. Murray, *Nature*. **439**, 55–9 (2006).

62. A. C. Lawson, C. E. Olsen, J. W. Richardson, M. H. Mueller, G. H. Lander, *Acta Crystallogr. Sect. B Struct. Sci.* **44**, 89–96 (1988).

63. O. Degtyareva, M. I. McMahon, D. R. Allan, R. J. Nelmes, *Phys. Rev. Lett.* **93**, 205502 (2004).

64. S. Lee, M. J. Bluemle, F. S. Bates, *Science (80-.).* **330**, 349–353 (2010).

65. E. D. S. Jr, *Nature*. **426**, 353–359 (2003).

66. D. Frenkel, *Phys. A Stat. Mech. its Appl.* **263**, 26–38 (1999).

67. F. C. Frank, J. S. Kasper, *Acta Crystallogr.* **12**, 483–499 (1959).

68. D. Frenkel, A. J. C. Ladd, *J. Chem. Phys.* **81**, 3188–3193 (1984).

69. A. Haji-Akbari, M. Engel, S. C. Glotzer, *J. Chem. Phys.* **135** (2011), doi:10.1063/1.3651370.

70. W. Shen *et al.*, *Soft Matter*. **15**, 2571–2579 (2019).

71. A. Haji-Akbari, M. Engel, S. C. Glotzer, *J. Chem. Phys.* **135**, 194101 (2011).

72. D. Frenkel, B. Smit, *Understanding molecular simulation: From algorithms to applications* (Academic Press, 2001; http://linkinghub.elsevier.com/retrieve/pii/B9780122673511500092).

73. Blavatska, von Ferber, Holovatch, *Condens. Matter Phys.* **14**, 33701 (2011).

74. H.-S. Nam, N. M. Hwang, B. D. Yu, J.-K. Yoon, *Phys. Rev. Lett.* **89**, 275502 (2002).

75. M. R. Langille, J. Zhang, M. L. Personick, S. Li, C. A. Mirkin, *Science (80-.).* **337**, 954–957 (2012).

76. B. de Nijs *et al.*, *Nat. Mater.* **14**, 56–60 (2014).

77. J. Wang *et al.*, *Nat. Commun.* **9**, 5259 (2018).

78. D. Wang *et al.*, *Nat. Phys.* (2020), doi:10.1038/s41567-020-1003-9.

79. K. Kobayashi, L. M. Hogan, *Philos. Mag. A*. **40**, 399–407 (1979).

80. J. Y. Wu, S. Nagao, J. Y. He, Z. L. Zhang, *Nano Lett.* **11**, 5264–5273 (2011).

81. M. R. Langille, J. Zhang, C. A. Mirkin, *Angew. Chemie Int. Ed.* **50**, 3543–3547 (2011).

82. M. Song *et al.*, *Science (80-.).* **367**, 40–45 (2020).

83. S. Patala, L. D. Marks, M. Olvera de la Cruz, *J. Phys. Chem. Lett.* **4**, 3089–3094 (2013).

84. A. Haji-Akbari *et al.*, *Nature*. **462**, 773–777 (2009).

85. D. Frenkel, A. J. C. Ladd, *J. Chem. Phys.* **81**, 3188–3193 (1984).

86. H. Hofmeister, *Cryst. Res. Technol.* **33**, 3–25 (1998).

87. A. Müller *et al.*, *Angew. Chemie Int. Ed.* **41**, 3604–3609 (2002).

88. D. Klotsa, E. R. Chen, M. Engel, S. C. Glotzer, *Soft Matter*. **14**, 8692–8697 (2018).

89. A. Howie, L. D. Marks, *Philos. Mag. A*. **49**, 95–109 (1984).

90. R. L. Davidchack, J. R. Morris, B. B. Laird, *J. Chem. Phys.* **125**, 094710 (2006).

91. Z. Ou, L. Yao, H. An, B. Shen, Q. Chen, *Nat. Commun.* **11**, 4555 (2020).

92. E. G. Teich, G. van Anders, S. C. Glotzer, *Nat. Commun.* **10**, 64 (2019).

93. J. A. Anderson, J. Glaser, S. C. Glotzer, *Comput. Mater. Sci.* **173**, 109363 (2020).

94. V. Ramasubramani *et al.*, *Comput. Phys. Commun.* **254**, 107275 (2020).

95. C. A. Mirkin, R. L. Letsinger, R. C. Mucic, J. J. Storhoff, *Nature*. **382**, 607–609 (1996).

96. S. Angioletti-Uberti, B. M. Mognetti, D. Frenkel, *Nat. Mater.* **11**, 518–522 (2012).

97. L. Feng, R. Dreyfus, R. Sha, N. C. Seeman, P. M. Chaikin, *Adv. Mater.* **25**, 2779–2783 (2013).

98. M. T. Casey *et al.*, *Nat. Commun.* **3**, 1209 (2012).

99. M. R. Jones *et al.*, *Nat. Mater.* **9**, 913–917 (2010).

100. W. B. Rogers, W. M. Shih, V. N. Manoharan, *Nat. Rev. Mater.* **1**, 16008 (2016).

101. R. J. Macfarlane, M. N. O'Brien, S. H. Petrosko, C. A. Mirkin, *Angew. Chemie Int. Ed.* **52**, 5688–5698 (2013).

102. F. Lu, K. G. Yager, Y. Zhang, H. Xin, O. Gang, *Nat. Commun.* **6**, 6912 (2015).

103. H. Xing *et al.*, *ACS Nano*. **6**, 802–809 (2012).

104. Y. Ke, L. L. Ong, W. M. Shih, P. Yin, *Science (80-.).* **338**, 1177–1183 (2012).

105. Y. Wang *et al.*, *Nature*. **491**, 51–55 (2012).

106. M. Grünwald, P. L. Geissler, *ACS Nano*. **8**, 5891–5897 (2014).

107. F. Romano, F. Sciortino, *Nat. Commun.* **3**, 975 (2012).

108. M. R. Walsh, C. a Koh, E. D. Sloan, A. K. Sum, D. T. Wu, *Science*. **326**, 1095–1098 (2009).

109. M. Engel, P. F. Damasceno, C. L. Phillips, S. C. Glotzer, *Nat. Mater.* **14**, 109–116 (2015).

110. W. F. Claussen, *J. Chem. Phys.* **19**, 259–260 (1951).

111. L. Pauling, R. E. Marsh, *Proc. Natl. Acad. Sci.* **38**, 112–118 (1952).

112. G. B. Adams, M. O'Keeffe, A. A. Demkov, O. F. Sankey, Y.-M. Huang, *Phys. Rev. B*. **49**, 8048–8053 (1994).

113. A. M. Guloy *et al.*, *Nature*. **443**, 320–323 (2006).

114. Y. Xia, Y. Xiong, B. Lim, S. E. Skrabalak, *Angew. Chemie Int. Ed.* **48**, 60–103 (2009).

115. M. N. O'Brien, M. R. Jones, K. A. Brown, C. A. Mirkin, *J. Am. Chem. Soc.* **136**, 7603–7606 (2014).

116. M. L. Personick *et al.*, *J. Am. Chem. Soc.* **133**, 6170–6173 (2011).

117. M. Spellings, R. L. Marson, J. A. Anderson, S. C. Glotzer, (2016).

118. A. V. Shevelkov, K. Kovnir, (2010; http://link.springer.com/10.1007/430_2010_25), pp. 97–142.

119. K. Momma *et al.*, *Nat. Commun.* **2**, 196 (2011).

120. F. C. Frank, J. S. Kasper, *Acta Crystallogr.* **11**, 184–190 (1958).

121. J. D. Weeks, D. Chandler, H. C. Andersen, *J. Chem. Phys.* **54**, 5237 (1971).

122. C. Speranza, S. Prestipino, G. Malescio, P. V. Giaquinta, *Phys. Rev. E*. **90**, 012305 (2014).

123. A. K. Sinha, in *Progress in Materials Science* (Pergamon Press, 1972), vol. 15, pp. 79–185.

124. G. N. Lewis, *J. Am. Chem. Soc.* **38**, 762–785 (1916).

125. L. Pauling, *The Nature of the Chemical Bond, and the Structure of Molecules and Crystals; An Introduction to Modern Structural Chemistry* (Cornell University Press, Ithaca, NY, ed. 2., 1940).

126. A. M. Kalsin *et al.*, *Science (80-.).* **312**, 420–424 (2006).

127. M. A. Boles, M. Engel, D. V. Talapin, *Chem. Rev.* **116**, 11220–11289 (2016).

128. S. C. Glotzer, M. J. Solomon, *Nat. Mater.* **6**, 557–562 (2007).

129. P. F. Damasceno, M. Engel, S. C. Glotzer, *Science (80-.).* **337**, 453–457 (2012).

130. M. Huang *et al.*, *Science (80-.).* **348**, 424–428 (2015).

131. C. Yi *et al.*, *Science (80-.).* **369**, 1369–1374 (2020).

132. M. He *et al.*, *Nature*. **585**, 524–529 (2020).

133. S.-J. Park, A. A. Lazarides, J. J. Storhoff, L. Pesce, C. A. Mirkin, *J. Phys. Chem. B*. **108**, 12375–12380 (2004).

134. S. Y. Park *et al.*, *Nature*. **451**, 553–556 (2008).

135. D. Nykypanchuk, M. M. Maye, D. van der Lelie, O. Gang, *Nature*. **451**, 549–552 (2008).

136. M. Girard *et al.*, *Science (80-.).* **364**, 1174–1178 (2019).

137. S. Wang *et al.*, *J. Am. Chem. Soc.* **141**, 20443–20450 (2019).

138. E. Auyeung *et al.*, *Nature*. **505**, 73–77 (2014).

139. E. Auyeung, R. J. Macfarlane, C. H. J. Choi, J. I. Cutler, C. A. Mirkin, *Adv. Mater.* **24**, 5181–5186 (2012).

140. C. R. Laramy, M. N. O'Brien, C. A. Mirkin, *Nat. Rev. Mater.* **4**, 201–224 (2019).

141. P. Akcora *et al.*, *Nat. Mater.* **8**, 354–359 (2009).

142. G. Pólya, *Mathematics and Plausible Reasoning* (Oxford University Press, London, 1954).

143. C. R. Iacovella, A. S. Keys, M. A. Horsch, S. C. Glotzer, *Phys. Rev. E*. **75**, 040801 (2007).

144. S. T. Hyde, M. O'Keeffe, D. M. Proserpio, *Angew. Chemie Int. Ed.* **47**, 7996–8000 (2008).

145. W. Longley, T. J. McIntosh, *Nature*. **303**, 612–614 (1983).

146. I. Prasad, H. Jinnai, R.-M. Ho, E. L. Thomas, G. M. Grason, *Soft Matter*. **14**, 3612–3623 (2018).

147. M. Saba, M. D. Turner, K. Mecke, M. Gu, G. E. Schröder-Turk, *Phys. Rev. B*. **88**, 245116 (2013).

148. M. Saba *et al.*, *Phys. Rev. Lett.* **106**, 103902 (2011).

149. J. J. K. Kirkensgaard, M. E. Evans, L. de Campo, S. T. Hyde, *Proc. Natl. Acad. Sci.* **111**, 1271–1276 (2014).

150. C. X. Du, G. van Anders, R. S. Newman, S. C. Glotzer, *Proc. Natl. Acad. Sci.* **114**, E3892–E3899 (2017).

151. L. Sandoval, H. M Urbassek, P. Entel, *New J. Phys.* **11**, 103027 (2009).

152. S. Lee, C. Leighton, F. S. Bates, *Proc. Natl. Acad. Sci.* **111**, 17723–17731 (2014).

153. B. Li, D. Zhou, Y. Han, *Nat. Rev. Mater.* **1**, 15011 (2016).

154. M. Hanfland, K. Syassen, N. E. Christensen, D. L. Novikov, *Nature*. **408**, 174–178 (2000).

155. D. Hobbs, J. Hafner, D. Spišák, *Phys. Rev. B*. **68**, 014407 (2003).

156. M. I. McMahon, R. J. Nelmes, *Chem. Soc. Rev.* **35**, 943 (2006).

157. J. M. Seddon, *Biochim. Biophys. Acta - Rev. Biomembr.* **1031**, 1–69 (1990).

158. C. R. Iacovella, A. S. Keys, S. C. Glotzer, *Proc. Natl. Acad. Sci.* **108**, 20935–20940 (2011).

159. H.-J. Sun, S. Zhang, V. Percec, *Chem. Soc. Rev.* **44**, 3900–3923 (2015).

160. P. Ziherl, R. D. Kamien, *J. Phys. Chem. B*. **105**, 10147–10158 (2001).

161. T. D. Nguyen, C. L. Phillips, J. A. Anderson, S. C. Glotzer, *Comput. Phys. Commun.* **182**, 2307–2313 (2011).

162. J. Glaser, X. Zha, J. A. Anderson, S. C. Glotzer, A. Travesset, *Comput. Mater. Sci.* **173**, 109430 (2020).

163. J. Towns *et al.*, *Comput. Sci. Eng.* **16**, 62–74 (2014).

164. C. Knorowski, S. Burleigh, A. Travesset, *Phys. Rev. Lett.* **106**, 215501 (2011).

165. P. J. Steinhardt, D. R. Nelson, M. Ronchetti, *Phys. Rev. B*. **28**, 784–805 (1983).

166. J. S. Oh, S. Lee, S. C. Glotzer, G.-R. Yi, D. J. Pine, *Nat. Commun.* **10**, 3936 (2019).

167. J. A. Marsh, S. A. Teichmann, *Annu. Rev. Biochem.* **84**, 551–575 (2015).

168. S. E. Ahnert, J. A. Marsh, H. Hernández, C. V. Robinson, S. A. Teichmann, *Science (80-.).* **350**, 1331 (2015).

169. L. Rossi *et al.*, *Soft Matter*. **7**, 4139–4142 (2011).

170. S. Sacanna, W. T. M. Irvine, P. M. Chaikin, D. J. Pine, *Nature*. **464**, 575–578 (2010).

171. A. J. Simon *et al.*, *Nat. Chem.* **11**, 204–212 (2019).

172. G. Munaò, Z. Preisler, T. Vissers, F. Smallenburg, F. Sciortino, *Soft Matter*. **9**, 2652 (2013).

173. S. N. Fejer, D. Chakrabarti, H. Kusumaatmaja, D. J. Wales, *Nanoscale*. **6**, 9448–9456 (2014).

174. Z. Preisler, T. Vissers, G. Munaò, F. Smallenburg, F. Sciortino, *Soft Matter*. **10**, 5121–5128 (2014).

175. Z. W. Li, Z. Y. Lu, Z. Y. Sun, L. J. An, *Soft Matter*. **8**, 6693–6697 (2012).

176. D. J. Beltran-Villegas, B. A. Schultz, N. H. P. Nguyen, S. C. Glotzer, R. G. Larson, *Soft Matter*. **10**, 4593–4602 (2014).

177. E. Bianchi, A. Z. Panagiotopoulos, A. Nikoubashman, *Soft Matter*. **11**, 3767–3771 (2015).

178. G. Avvisati, T. Vissers, M. Dijkstra, *J. Chem. Phys.* **142**, 084905 (2015).

179. Q. Chen *et al.*, *Science*. **331**, 199–202 (2011).

180. Y. Wang *et al.*, *Nat. Commun.* **6**, 7253 (2015).

181. W. B. Rogers, J. C. Crocker, *Proc. Natl. Acad. Sci.* **108**, 15687–15692 (2011).

182. B. M. Mladek, J. Fornleitner, F. J. Martinez-Veracoechea, A. Dawid, D. Frenkel, *Phys. Rev. Lett.* **108**, 268301 (2012).

183. L. Tang, L. Stith, E. K. Jaffe, *J. Biol. Chem.* **280**, 15786–15793 (2005).

184. K. Sakurai, H. Shimada, T. Hayashi, T. Tsukihara, *Acta Crystallogr. Sect. F Struct. Biol. Cryst. Commun.* **65**, 80–83 (2009).

185. R. Kalia *et al.*, *Nature*. **558**, 401–405 (2018).

186. S. Sacanna *et al.*, *Nat. Commun.* **4**, 1688 (2013).

187. J. S. Oh, Y. Wang, D. J. Pine, G. R. Yi, *Chem. Mater.* **27**, 8337–8344 (2015).

188. Y. Wang *et al.*, *J. Am. Chem. Soc.* **137**, 10760–10766 (2015).

189. N. J. Agard, J. A. Prescher, C. R. Bertozzi, *J. Am. Chem. Soc.* **126**, 15046–15047 (2004).

190. E. R. Miraldi, P. J. Thomas, L. Romberg, *Biophys. J.* **95**, 2470–2486 (2008).

191. T. Vissers, F. Smallenburg, G. Munaò, Z. Preisler, F. Sciortino, *J. Chem. Phys.* **140**, 144902 (2014).

192. M. Song, Y. Ding, M. A. Snyder, J. Mittal, *Langmuir*. **32**, 10017–10025 (2016).

193. S. Kumar, M.-J. Lee, V. K. Aswal, S.-M. Choi, *Phys. Rev. E*. **87**, 042315 (2013).

194. P. D. Kaplan, J. L. Rouke, A. G. Yodh, D. J. Pine, *Phys. Rev. Lett.* **72**, 582–585 (1994).

195. S. I. A. Cohen *et al.*, *Proc. Natl. Acad. Sci.*, in press, doi:10.1073/pnas.1218402110.

196. C. Kang, A. Honciuc, *J. Phys. Chem. Lett.* **9**, 1415–1421 (2018).

197. C. Kang, A. Honciuc, *ACS Nano*. **12**, 3741–3750 (2018).

198. W. Xi, U. H. E. Hansmann, *Sci. Rep.* **7**, 6588 (2017).

199. P. M. Morse, *Phys. Rev.* **34**, 57–64 (1929).

200. É. Ducrot, M. He, G.-R. Yi, D. J. Pine, *Nat. Mater.* **16**, 652 (2017).

201. C. P. Royall, R. van Roij, A. van Blaaderen, *J. Phys. Condens. Matter*. **17**, 2315–2326 (2005).

202. A. A. Shah, B. Schultz, W. Zhang, S. C. Glotzer, M. J. Solomon, *Nat. Mater.* **14**, 117 (2014).

203. N. Kern, D. Frenkel, *J. Chem. Phys.* **118**, 9882–9889 (2003).