Controlling crystallization and its absence: proteins, colloids and patchy models†

Jonathan P. K. Doye,*[a] Ard A. Louis,[b] I-Chun Lin,[c] Lucy R. Allen,[d] Eva G. Noya,[e] Alex W. Wilber,[a] Hoong Chwan Kok[e] and Rosie Lyus[e]

Received 16th October 2006, Accepted 22nd December 2006
First published as an Advance Article on the web 23rd January 2007
DOI: 10.1039/b614955c

The ability to control the crystallization behaviour (including its absence) of particles, be they biomolecules such as globular proteins, inorganic colloids, nanoparticles, or metal atoms in an alloy, is of both fundamental and technological importance. Much can be learnt from the exquisite control that biological systems exert over the behaviour of proteins, where protein crystallization and aggregation are generally suppressed, but where in particular instances complex crystalline assemblies can be formed that have a functional purpose. We also explore the insights that can be obtained from computational modelling, focussing on the subtle interplay between the interparticle interactions, the preferred local order and the resulting crystallization kinetics. In particular, we highlight the role played by "frustration", where there is an incompatibility between the preferred local order and the global crystalline order, using examples from atomic glass formers and model anisotropic particles.

1. Introduction

Controlling crystallization is a subject of considerable importance both from a fundamental and an applied perspective. Chemical physicists want to understand how the interparticle interactions determine the most favoured crystal structure and the ease with which crystallization can occur. Biochemists want to know the best recipe to crystallize the protein in which they are interested, so that they can then determine its structure. Nanotechnologists want to know how to get colloids or nanoparticles to self-assemble into a given target structure, such as a diamond lattice because of its potential importance in photonics. Metallurgists want to be able to predict which alloys will most readily avoid crystallization, and instead form a metallic glass.[1]

In this paper we want to take a theoretical and computer simulation perspective on the factors that control crystallization, and its absence, in these kinds of systems. So far in the literature, isotropic models have been the starting point for much of the theoretical work in this area, and although there have been considerable successes, this approach has its limits. For example, the experimentally observed "crystallization slot" for globular proteins, where the interprotein interactions, as measured by the second virial coefficient, are sufficiently attractive to encourage crystallization, but not so attractive as to lead to irreversible aggregation,[2] has been rationalised using isotropic model potentials.[3] Another fruitful idea is that if a protein solution lies near to a metastable critical point associated with separation into protein-rich and solvent-rich phases, the associated concentration fluctuations could enhance crystal nucleation.[4] However, proteins are not isotropic, but are anisotropic both in their shape and in their interactions. One reflection of the latter is that protein crystals typically have much lower packing fractions[5] than the close-packed structures that are favoured by isotropic potentials.

Similarly, for colloids and nanoparticles there are now a considerable array of different crystals that have been obtained just using isotropic particles, particularly through the use of binary mixtures. For example, just by varying the relative sizes of the two types of particles, surprisingly complex crystals can be formed.[6,7] Very recently, the use of particles with charges of opposite signs has led to colloidal[8,9] and nanoparticle[10–12] analogues of ionic crystals, even relatively low-density structures such as zinc blende.[12] However, there are limits to the structures that can be assembled from isotropic particles,[13] and so there has been considerable recent interest in developing methods to produce colloidal particles that are anisotropic in shape[14–16] or in their interactions.[15,17–21] For all these reasons, there has been much theoretical interest in beginning to explore simple anisotropic models and their effects on crystallization[22–29] and self-assembly.[30–33]

Here we provide a particular viewpoint on this broad area of controlling crystallization and its absence. In Section 2 we outline some of our perspectives on protein crystallization, particularly why it is important to take into account the

[a] *Physical and Theoretical Chemistry Laboratory, Oxford University, South Parks Road, Oxford, UK OX1 3QZ. E-mail: jonathan.doye@chem.ox.ac.uk; Fax: +44 (0)1865 275410; Tel: +44 (0)1865 275426*
[b] *Rudolf Peierls Centre for Theoretical Physics, 1 Keble Road, Oxford, UK OX1 3NP*
[c] *Institute of Chemical Sciences and Engineering, BCH 4109 Swiss Federal Institute of Technology EPFL, CH-1015 Lausanne, Switzerland*
[d] *School of Physics and Astronomy, University of Leeds, UK LS2 9JT*
[e] *Department of Chemistry, University of Cambridge, Lensfield Road, Cambridge, UK CB2 1EW*
† The HTML version of this article has been enhanced with colour images.

evolutionary origins of proteins when trying to understand their crystallization behaviour. In Section 3 we explore how the concept of frustration can provide physical insight into the structural origins of the glass-forming ability for two of the most common models for simulating supercooled liquids. Finally, in Section 4, we introduce some of our recent results that use model patchy particles to explore how the geometry of the interparticle interactions can both influence and be used to control the crystallization and self-assembly behaviour.

2. Protein crystallization

The question concerning protein crystallization that we have been particularly seeking to address[34] is: Why are globular proteins seemingly so hard to crystallize? Although it is not true of all proteins—lysozyme and insulin provide notable examples of proteins that crystallize relatively easily—there is plenty of anecdotal evidence that proteins are often very difficult to crystallize. Furthermore, the rise of structural genomics initiatives has now made it possible to start to quantify this difficulty. For example, the success rate for producing X-ray quality crystals has been estimated to be roughly 20% for those prokaryotic proteins that can be expressed in soluble form.[35,36] Moreover, the crystallization approaches used are mainly empirical. For example, screening kits allow large numbers of solution conditions, which have previously been found to be useful for crystallization, to be quickly tested.

Our perspective on this question[34] is that the need for proteins to function within the crowded cellular environment places major constraints on the surface properties of proteins. In particular, not only must a protein interact correctly with its binding partners, but it must also make sure that it does not stick to anything else in the cell, including other copies of itself. Failure to do so is likely to be deleterious to the cell, as in the many protein aggregation diseases, of which the most relevant to this paper are those associated with native state aggregation and crystallization. For example, sickle cell anaemia involves the ordered aggregation of haemoglobin into fibrils, and some forms of anaemia and cataracts are the result of the crystallization of mutant forms of the haemoglobin[37] and γ-crystallin[38] proteins, respectively.

Thus, our hypothesis is that the surface properties of proteins have been selected in order to prevent native-state aggregation and crystallization *in vivo* (we term this 'engineering out' of unwanted properties negative design) and that, because of the robustness of the mechanisms used, many proteins are difficult to crystallize even in the far-from-physiological conditions explored by the protein crystallographer. Furthermore, there is a relatively simple way to test this hypothesis. If negative design is present, then random mutagenesis of surface amino acids should on average lead to a protein that is easier to crystallize, and the two such mutagenesis studies in the literature do indeed find such a correlation.[39,40]

More useful, though, would be to identify the mechanisms used to achieve this negative design, since then this raises the possibility of developing rational strategies to overcome the negative design in order to make protein crystallization easier. Interestingly, bioinformatic analyses have provided a 'smoking gun' for the potential role of surface lysine residues. Lysine is the most common surface amino acid, but yet is the most underrepresented at functional interfaces[41,42] and at contacts between proteins in crystals.[43] This of course begs the question: What is it doing there, if it is only reluctantly involved in interactions? One possible answer is that lysine helps to prevent unwanted interactions.

Consistent with this suggestion, the group of Zygmunt Derewenda have shown that systematically mutating surface lysine residues, particularly to alanine,[44] almost invariably enhances the crystallizability of a protein.[36,45] The long lysine side-chain has substantial conformational entropy, which would be lost if constrained at a protein–protein interface. Thus, it has been suggested that lysine stabilizes the free protein in solution by providing a 'surface entropy shield',[36] and replacing lysines with the compact alanine is expected to make crystallization thermodynamically more favourable.

One possible objection to our negative design hypothesis is that, given that proteins are irregular objects subject to dynamical fluctuations, it is not surprising that they are hard to crystallize. However, further evidence that the low crystallizability of most proteins is not an intrinsic property, but one that arises due to selection, comes from instances when protein crystallization does occur *in vivo*, because it achieves some functional purpose. Nature has no problem getting proteins to crystallize, even in the 'dirty' environment of the cell, when there is a positive selection pressure for this.

There are many fascinating examples of this *in vivo* protein crystallization, as catalogued in a recent review.[46] Such crystallization represents a rather neglected area of biological self-assembly, and one which is worthy of further study. Here we give just a few examples. Certain genera of viruses that infect insect larvae coopt the cell to express large quantities of specific proteins during the late stages of infection. These then crystallize around the viral particles to provide a protective environment for the viruses after the death of the insect larvae, as illustrated in Fig. 1(b). On ingestion of the crystals by a new insect larva, the protein crystals dissolve in the alkaline environment of the gut, releasing the virus to infect the new host.[47]

A number of examples of crystalline proteins in the cell correspond to proteins stored in secretory granules. Fig. 1(a) illustrates one particularly specialized example from the protist (a type of single-celled eukaryotic organism) *Paramecium*.[49] Large numbers of these granules, which are called trichocysts, are attached to the outer membrane of this organism, and are assembled from three different families of closely related polypeptides, each localized to different regions of the trichocyst to form a structure that is of the order of 3–4 μm in size. In response to an external stimulus the trichocysts can be simultaneously extruded from the cell. In this process, the crystalline trichocysts expand in length by a factor of roughly eight. The purpose seems to be defensive. The explosive release of the trichocysts can push *Paramecium* away from a potential predator, giving it a chance to escape.[51]

In both these examples, a high level of control is asserted over the crystallization process in order to determine the location, size and shape of the crystal. The mechanical

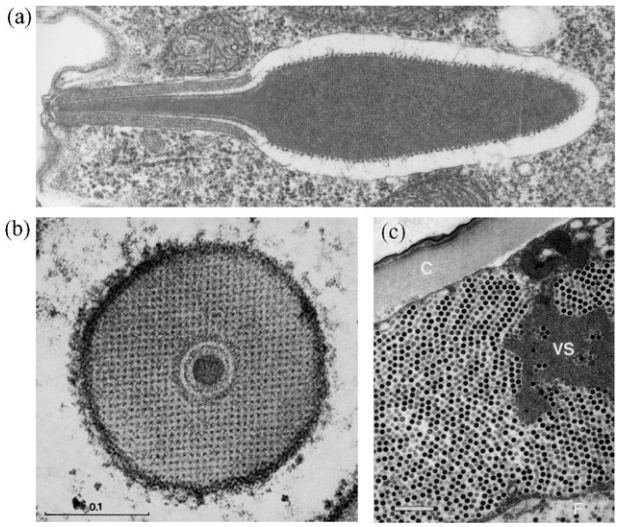

Fig. 1 Some examples of protein crystallization occurring in the cell. (a) A trichocyst attached to the outer membrane of *Paramecium*. (b) An encapsulated rod virus of the granulosis virus of *Plodia interpunctella*. (c) Paracrystalline arrays of an iridescent virus in the epidermis of a larva of *Culicoides variipennis sonorensis*. The scale bars in (b) and (c) correspond to 0.1 μm and 1 μm, respectively. Reproduced from ref. (a) [48] and (b) [49] and (c) [50] (© 1968, 2000 and 1999, respectively, with permission from Elsevier).

properties are also important for their function, with the virus-encapsulating crystals required to be tough in order to provide a protective environment, and the trichocysts able to undergo an irreversible expansion triggered by the presence of calcium ions.

Our final example is of the crystallization of virus particles themselves (Fig. 1(c)), and has perhaps more to do with an absence of a selection pressure to prevent crystallization (why should it bother the virus?) than a functional purpose for the crystals. The most dramatic examples come from iridoviruses, which have large icosahedral capsids. This type of virus was first identified due to the iridescent colours imparted to insect larvae (hence the name)[52] due to the Bragg scattering of visible light by the crystals.[53]

3. Glass formation

There has been much work trying to understand the unusual dynamic properties of supercooled liquids as the glass transition is approached.[54] To do this from a simulation perspective, model systems have been developed that are robust glass-formers and show no tendency to crystallize. Such models also allow one to probe the structural determinants of a system's glass-forming ability. A key notion is that of 'frustration',[55,56] which is said to occur when the preferred local order is incompatible with global crystalline order. Typically, the liquid structure reflects the preferred local order, and so when frustration is present, significant structural reorganization will be necessary to nucleate a crystalline phase.

To apply these ideas, one needs to be able to identify the preferred local order, but, because of the inherent disorder associated with a liquid, sometimes this can be hard to achieve directly from analysing liquid configurations. An alternative approach is to instead look at the structures of isolated clusters, since such a cluster can adopt the preferred structure without having to mould itself to any surrounding medium. This approach was first applied by Charles Frank to rationalize why small liquid metal droplets could be substantially supercooled.[57]

One of the potentials most commonly used by simulators is the Lennard-Jones potential. However, the one-component Lennard-Jones fluid is not a good system to look at the behaviour of supercooled liquids, since crystallization into a close-packed solid occurs relatively easily. To generate a good glass-former, Dzugutov modified this potential by adding a barrier in the potential at roughly $\sqrt{2}$ times the position of the potential minimum (Fig. 2(a)), hence energetically disfavouring close-packed crystals because of their octahedral interstices.[58] Instead, polytetrahedral order,[59] and local icosahedral order in particular, is preferred. As can be seen in Fig. 2(a), the distances present in the 13-atom icosahedral cluster avoid sampling this bump in the potential.

Even though there is a stable body-centred-cubic phase,[60] the resulting Dzugutov liquid can be easily supercooled. Interestingly, the Dzugutov system is one of the only liquids in which an increase in a structural length scale has been detected on increased supercooling, in particular domains with local icosahedral order grow in size as the temperature is decreased.[61] However, these domains are not compact, but are ramified in structure. These results raise the questions: What are the origins of the non-compactness, and might this shape be a generic property of domains of enhanced local order within supercooled liquids? An analysis of the structures of isolated clusters is able to provide some answers. This type of ordering is also clearly seen in the clusters, but an analysis

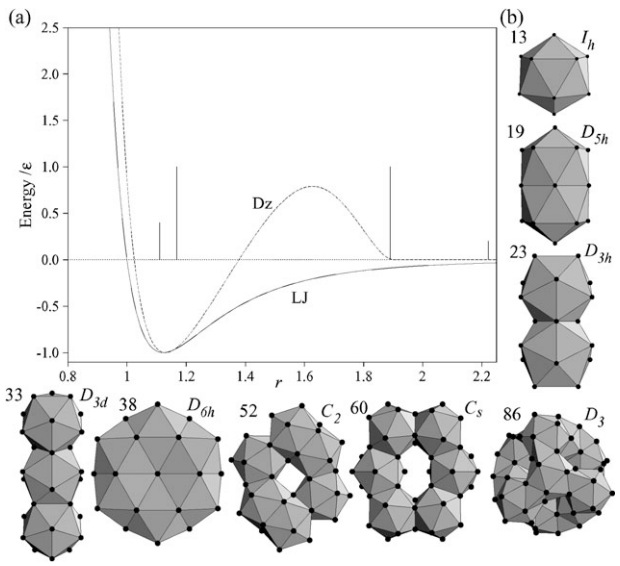

Fig. 2 (a) The Dzugutov (Dz) and Lennard-Jones (LJ) potentials compared. The vertical lines show the positions of the pair distances in the 13-atom icosahedron and their heights are proportional to the numbers of pairs with that distance. (b) Some of the particularly stable structures for Dzugutov clusters. Each cluster is labelled by its size and point group.

of the energetics shows that it is a consequence of the unusual shape of the Dzugutov potential. Therefore, this domain shape is unlikely to be universal.

The lowest-energy Dzugutov clusters are all aggregates of interpenetrating and face-sharing icosahedra (Fig. 2(b)).[62] As the cluster size increases first chains, then disks, then rings, and finally 3-dimensional porous networks of icosahedra are seen. The energetic causes for this ordering are also clear. The bump in the potential promotes local icosahedral order. However, icosahedral structures are inherently strained, *e.g.* the distance between adjacent atoms on the surface of the regular 13-atom icosahedron is 5% longer than that to the central atom (Fig. 2(a)). This strain, and the associated energetic penalty, grows rapidly for compact polyicosahedral structures. The resulting non-compact icosahedral aggregates represent a compromise that maintains the local icosahedral coordination, but which does not involve excessive strains. Similar behaviour has also been seen for variants of the Dzugutov potential.[63]

A second approach used to generate a good glass-former is to introduce two atom types. Partly this is because the formation of a compositionally ordered crystal can require the atoms to diffuse significantly further to find the correct environment than for the one-component case. However, there is more to it than this, since the interactions between the particles and the composition also need to be tuned to reach those regions of the parameter space where crystallization is particularly difficult.[64,65] The most commonly used model of this type is the Kob–Andersen binary Lennard-Jones mixture,[66] *i.e.* the potential is

$$V_{\mathrm{BLJ}}(r_{ij}) = 4\varepsilon_{\alpha\beta}\left[\left(\frac{\sigma_{\alpha\beta}}{r_{ij}}\right)^{12} - \left(\frac{\sigma_{\alpha\beta}}{r_{ij}}\right)^{6}\right], \qquad (3.1)$$

where α and β are the atom types of atoms i and j, and $\varepsilon_{\alpha\beta}$ and $2^{1/6}\sigma_{\alpha\beta}$ are the pair well depth and equilibrium pair separation, respectively, for the interaction between atoms i and j. The Kob–Andersen parameters, $\sigma_{\mathrm{AB}} = 0.8\sigma_{\mathrm{AA}}$, $\sigma_{\mathrm{BB}} = 0.88\sigma_{\mathrm{AA}}$, $\varepsilon_{\mathrm{AB}} = 1.5\varepsilon_{\mathrm{AA}}$, and $\varepsilon_{\mathrm{BB}} = 0.5\varepsilon_{\mathrm{AA}}$, are non-additive and strongly favour mixing. At the canonical composition A_4B the ground state is a coexisting pure A face-centred-cubic crystal, and an AB CsCl-type crystal,[64] with low-lying layered A_4B crystals also possible.[67] However, crystallization has never been seen in a simulation.

Here, we use isolated clusters to analyse the preferred coordination environments of A atoms around B atoms, since at the glass-forming composition B atoms are in a minority. In particular, we locate the particular stable clusters with one, two, three and four B atoms to see how these preferred environments can organize into larger structures (Fig. 3). For coordination of a single B atom, clusters with 8, 9 and 10 atoms have similar stability. All three structures are based on a square antiprism of A atoms, but with possible capping atoms over the two square faces. This stability of the square antiprism is also seen for A_8B_2 and A_8B_3, where the additional capping atoms are now B atoms. One way of combining such coordination environments is illustrated by $A_{12}B_2$ and $A_{16}B_3$ and involves the sharing of the square-faces of the square antiprisms to form linear aggregates. Similar columns of square anti-prisms are found in the Al_2Cu crystal. $A_{18}B_3$

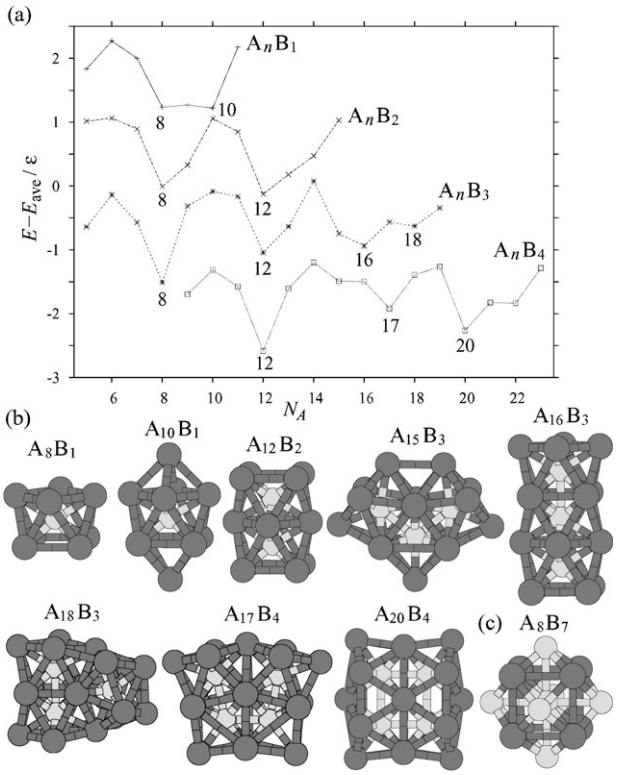

Fig. 3 (a) The energies of the global minima for four series of Kob–Andersen BLJ clusters with different numbers of B atoms. To make particularly stable clusters stand out, the energies are measured with respect to E_{ave}, a fit to the energies of the clusters. For clarity, the curves have also be displaced with respect to each other. (b) Some of the particularly stable structures identified in (a) with all B atoms completely coordinated. (c) A particularly stable cluster at near equimolar composition.

illustrates another possibility and involves the sharing of two triangular faces of $A_{12}B_2$ with a bicapped square antiprism.

Another possible coordination environment for the B atoms is the tricapped trigonal prism. Although this structure is not stable for A_9B (it can relax to a monocapped square antiprism by a single diamond-square process[68]) it becomes a common environment for larger clusters. This is illustrated by the structures $A_{15}B_3$, $A_{17}B_4$ and $A_{20}B_4$, which have one, two and four tricapped trigonal prismatic environments, respectively, albeit with B atoms in some of the capping sites.

These results tie in well with studies of the local structure in liquid configurations, which have found that square antiprism, and trigonal prismatic environments increasingly dominate the local structure in the liquid as the temperature is decreased.[69]

So how do these results help to rationalize the system's ability to avoid crystallization. Firstly, these preferred local environments are not present in the lowest-energy crystal structures. However, there are still a number of crystalline structures that are significantly lower in energy than the liquid that do involve these environments.[64,65,69] Secondly, the diversity of environments and ways that these can pack is likely to frustrate the formation of a uniform crystal.

In contrast to the A_4B system, at equimolar compositions the Kob–Andersen BLJ model easily crystallizes into a

CsCl-type structure.[64] Again, isolated clusters can help to understand this behaviour. Examining A_8B_n clusters, we found A_8B_7 to be particularly stable and to exhibit the bulk crystal structure (Fig. 3(c)), hence showing the absence of frustration at this composition.

4. Patchy models

In the above models with isotropic interactions, the connection between the preferred local structure and the form and parameters of the potential can be quite subtle. To probe the connection between crystallizability and local structure further, it would be useful to have a model where the local structure can be more directly controlled. To achieve this goal in a simple and flexible way, here we use 'patchy' particles, where the particles are spherical and only interact strongly when the patches on adjacent particles are aligned, but we should note that other ways of using anisotropic interactions to control the local structure have recently been explored.[70,71] There has been much recent interest in such patchy models from the perspective of protein crystallization,[22,23,25,26,28] self-assembly[29,30] (particularly into monodisperse clusters in a way similar to virus self-assembly[30,31,33]) and the dynamics of supercooled liquids,[72,73] but it is still an area that is very much in its infancy with much to be discovered. Our interest is to probe how the anisotropy of the interparticle interactions can be used to control a system's crystallization and self-assembly behaviour, and, in particular, through its effect on the degree of frustration. As well as being of fundamental interest, the hope is that the results will also be useful for understanding the crystallization behaviour of proteins and anisotropic colloids.

4.1. Potential

Here, we model the patchy particles with a single-site potential, *i.e.* each particle is represented by just a single site, but which has both position and orientation. The potential consists of an isotropic repulsion, which is based on the Lennard-Jones potential

$$V_{\mathrm{LJ}}(r) = 4\varepsilon\left[\left(\frac{\sigma_{\mathrm{LJ}}}{r}\right)^{12} - \left(\frac{\sigma_{\mathrm{LJ}}}{r}\right)^{6}\right] \qquad (4.2)$$

but where the attraction is modulated by an orientational dependent term, V_{ang}. Thus, the complete potential is

$$V_{ij}(\boldsymbol{r}_{ij}, \boldsymbol{\Omega}_i, \boldsymbol{\Omega}_j) = \begin{cases} V_{\mathrm{LJ}}(r_{ij}) & r < \sigma_{\mathrm{LJ}} \\ V_{\mathrm{LJ}}(r_{ij}) V_{\mathrm{ang}}(\hat{\boldsymbol{r}}_{ij}, \boldsymbol{\Omega}_i, \boldsymbol{\Omega}_j) & r \geq \sigma_{\mathrm{LJ}}, \end{cases} \qquad (4.3)$$

where $\boldsymbol{\Omega}_i$ is the orientation of particle i. The patches are specified by a set of patch vectors, $\{\boldsymbol{p}_i\}$, as illustrated in Fig. 4. V_{ang} has the form

$$V_{\mathrm{ang}}(\hat{\boldsymbol{r}}_{ij}, \boldsymbol{\Omega}_i, \boldsymbol{\Omega}_j) = \exp\left(-\frac{\theta^2_{k_{\min} ij}}{2\sigma^2}\right) \exp\left(-\frac{\theta^2_{l_{\min} ji}}{2\sigma^2}\right) \qquad (4.4)$$

where σ is the standard deviation of the Gaussian, θ_{kij} is the angle between the patch vector $\boldsymbol{k}$ on particle i and the interparticle vector $\boldsymbol{r}_{ij}$, and $k_{\min}$ is the patch that minimizes the magnitude of this angle. Hence, only the patches on each particle that are closest to the interparticle axis interact with each other, and the potential is continuous as a function of the orientations of the particles. $V_{\mathrm{ang}} = 1$ when two patches point directly at each other, but falls off as the patches deviate further from the perfect alignment.

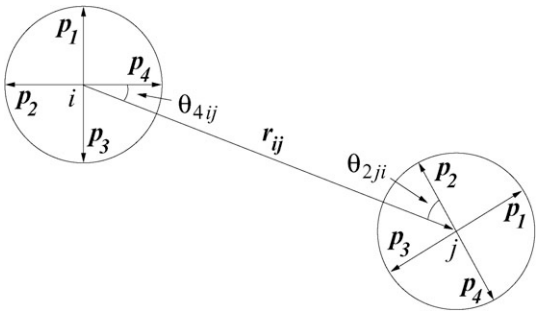

Fig. 4 The geometry of the interactions between the model particles. In this example, there are four patches regularly arranged on the disks, with their directions described by the patch vectors, $\boldsymbol{p}_i$. Patch 4 on particle i interacts with patch 2 on particle j because they are closest to the interparticle vector.

One of the nice features of this single-site potential is that, given a set of patch vectors, the potential has only one parameter σ, which determines the widths of the patches. Furthermore, as $1/\sigma \to 0$ the isotropic Lennard-Jones potential is recovered. Hence, it is possible to systematically study the behaviour of the model as a function of the degree of anisotropy with the well-characterized Lennard-Jones model as one limit.

4.2. Two-dimensional crystallization

We first discuss the application of this model to crystallization in two dimensions to illustrate the effects of the geometrical arrangement of the patches, since visualization is easier than in three dimensions. We choose to study particles with 4, 5 or 6 patches arranged regularly on the surface of the disks. Particularly interesting will be the 5-patch system as the local five-fold symmetry of the patches is incommensurate with global crystalline order.

The two-dimensional Lennard-Jones reference system crystallizes into a close-packed crystal with a hexagonal arrangement of neighbours around each particle. Unsurprisingly, for particles with a regular hexagonal array of patches, the anisotropy reinforces this behaviour. Crystallization is easy (Fig. 5(c)) and the close-packed crystal is lowest in energy for any combination of pressure and σ.

In the 4-patch system there is competition between a low-density square crystal in which each patch points directly at a nearest neighbour, and a higher-density hexagonal crystal. The latter is orientationally ordered with two opposite patches pointing at nearest neighbours, and the other two at next neighbours. As the square crystal has more pairs of nearest-neighbour patches pointing directly at each other, it is energetically preferred when the patches are sufficiently narrow (Fig. 6(a)). However, because of its more open structure, it becomes destabilised relative to the hexagonal crystal as the pressure is increased. Crystallization to the square crystal occurs relatively easily in the appropriate region of the phase diagram (Fig. 5(a)).

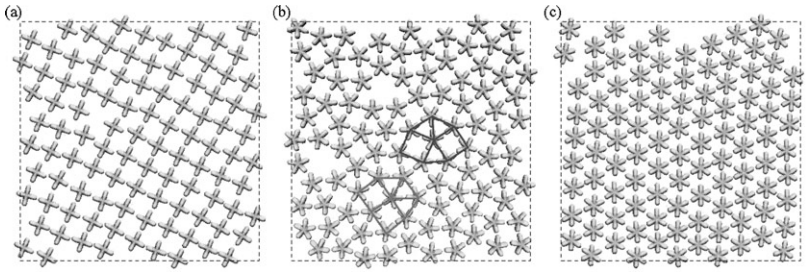

Fig. 5 Structures resulting from Monte Carlo cooling simulations for two-dimensional particles with (a) four (b) five and (c) six regularly spaced patches. In each case, there are 100 particles and $\sigma = \pi/12 \approx 0.262$. The pressures are (a) 0.1, (b) 0.2 and (c) 1.0 $\varepsilon/\sigma_{LJ}^3$. The dashed lines show the periodic boundary conditions. In (a) and (c) crystallization clearly occurs. The crystals are not quite perfect, because unless crystallization occurs at the correct orientation with respect to the boundary conditions and with the correct number of lattice points in each direction, there will not be the right number of particles to form a perfect crystal, and so some defects are necessarily present. In (b) there is no overall crystalline order, but two common motifs are highlighted.

For the 5-patch system the situation is more complex. No crystal phases are possible where all the patches point directly at neighbouring particles. For example, on cooling this system in the region of the phase diagram where the patches prefer to be aligned, the resulting configuration has no overall crystalline order. However, most of the particles have one of the two local environments highlighted in Fig. 5(b). From these two motifs, two crystals can be constructed where every particle has an identical environment (Fig. 6(b)). In both crystals each particle has five nearest neighbours, and the crystals can be thought of in terms of tilings of squares and equilateral triangles, where every vertex is surrounded by three triangles and two squares. The nomenclature for these semi-regular tilings is (3^2.4.3.4) and (3^3.4^2), and derives from the ordering of polygons around each vertex.[74] However, we will use the notation 'σ' and 'H' by analogy to the Frank–Kasper phases of these names that can be envisaged as square-triangle tilings in two of their three dimensions.[75]

When the patches are sufficiently narrow, the H- and σ-phases represent the best compromise between the five-fold symmetry of the particles and crystalline order. However, the patches cannot point directly at each other in these crystals, and so the system is frustrated. The mean deviation of the patches away from the nearest-neighbour interparticle vectors is smaller for the σ-phase (7.2°) than for the H-phase (9.6°), making the former lowest in energy for intermediate values of σ at low pressure. However, at small σ, the energetic penalty associated with these deviations becomes significant, and causes the H-phase to distort so that three of the five patches can point directly at each other. This distortion reduces the aspect ratio of the unit cell and the triangles deviate significantly from the ideal equilateral geometry. No such similar distortion is possible for the σ-phase, and so the H-phase becomes lowest in energy at small σ (Fig. 6(b)).

The H- and σ-phase represent only two of the possible tilings of squares and triangles. These tilings need not be crystalline, and dodecagonal quasicrystalline packings are also possible. Indeed, these quasicrystals have been seen in both metallic alloys[75] and macromolecular systems,[76] under circumstances near to where crystalline square-triangle phases are stable. Therefore, a stable quasicrystalline state remains an intriguing possibility for the current system.

The more dense hexagonal crystal is most stable at higher pressure, and closer to the isotropic limit. Although positionally ordered, the crystal is orientationally disordered, because of the incompatibility of the five-fold symmetry of the particles and the six-fold symmetry of the crystal (Fig. 6(b)).

An interesting comparison to the current system, which has five-fold symmetry in the attractive interactions, is provided by a study of hard pentagons.[77] Although the shape of these particles does frustrate crystallization to some extent, the lattice formed by the particle centres in the only stable orientationally-ordered crystalline phase, is based on a slightly distorted hexagonal lattice.

Also relevant is a recent study of a two-dimensional system of particles interacting with a Lennard-Jones potential plus an anisotropic term that favours the formation of isolated five-fold rings of particles.[70] The presence of such order again frustrates crystallization, and, interestingly, on varying the strength of the anisotropic term in the potential, the dynamic properties of the liquid, such as the fragility, can be changed substantially.

4.3. Three-dimensional crystallization

In our applications of this model to three dimensions, we again look at how low-density crystalline phases can be stabilized by the patchy interactions. The two systems that we consider are 4- and 6-patch particles with a regular tetrahedral and octahedral arrangement of the patches, respectively. For these systems, the stable crystalline phases at sufficiently low pressure and σ would be expected to be a diamond and a simple cubic lattice, respectively, with each patch directly pointing at a neighbouring particle. These have maximum packing fractions of 34% and 52% compared to 74% for the face-centred-cubic crystal favoured by the isotropic Lennard-Jones potential.

One might have expected these two systems to behave quite similarly in the regime where the open structures are preferred, but in fact their crystallization behaviour is quite different. The 6-patch system is able to crystallize easily, leading to a step-like decrease in the energy on crystallization (Fig. 7(a)), sometimes even giving a perfect defect-free crystal. By contrast for the 4-patch system at best only partial crystallization is observed, and on ordering the energy does not exhibit a step, but instead changes continuously.

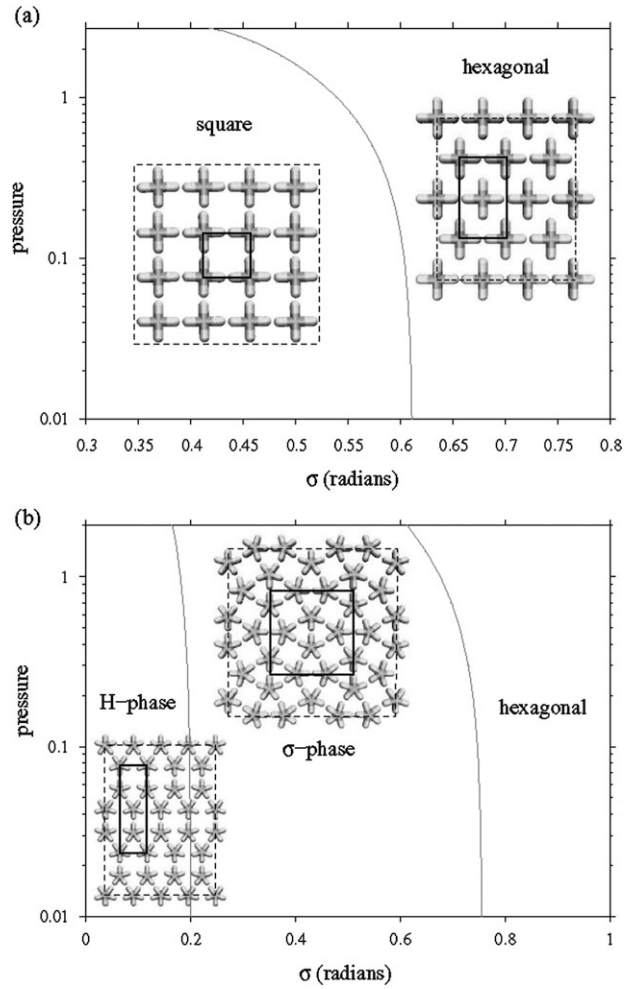

Fig. 6 Structural phase diagram showing how the ground-state structure in two dimensions depends on patch width and pressure (in units of $\varepsilon/\sigma_{LJ}^3$) for particles with (a) four and (b) five regularly spaced patches. The unit cells of the crystalline structures are depicted with thick lines, and the dashed lines show the periodic boundary conditions.

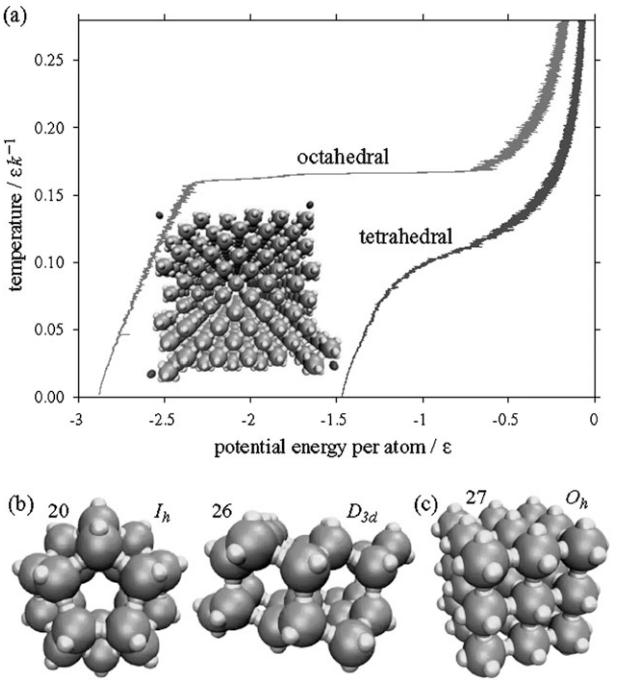

Fig. 7 (a) The caloric curves for cooling the 4-patch tetrahedral particles and the 6-patch octahedral particles from high temperature. The number of particles is 512, the patch width σ is 0.3 radians, and the pressure is $0.1\varepsilon/\sigma_{LJ}^3$. The resulting simple cubic crystal for the 6-patch system is also illustrated (the dots represent the corners of the simulation box). (b) and (c) Some of the global minima of isolated clusters for the 4- and 6-patch systems, respectively. Each cluster is labelled by the number of atoms and point group.

As in Section 3, examining the structure of isolated clusters can give us some clues as to the physical origins of this behaviour, because the clusters provide a picture of the preferred local order for the system. For both systems, when the patches are sufficiently narrow, as expected, the lowest-energy clusters exhibit open structures where the maximum coordination number is equal to the number of patches. For the octahedral particles, for $\sigma \lesssim 0.5$ most of the clusters are cuboidal nanocrystals with the simple cubic structure (*e.g.* the $3 \times 3 \times 3$ cube in Fig. 7(c)). The system is unfrustrated and it is unsurprising that crystallization is relatively easy.

By contrast, the global minima of the tetrahedral system for $0.2 \lesssim \sigma \lesssim 0.45$ do not exhibit the structure of the stable crystalline phase, but instead are based on dodecahedral cages (the 20-particle dodecahedron is shown in Fig. 7(b)). Only for $\sigma \lesssim 0.2$ are clusters with the diamond structures lowest in energy, *e.g.* the 26-particle structure in Fig. 7(b). In diamond the particles form six-fold rings, but are puckered, whereas the dodecahedra are characterized by planar five-fold rings, where the 108° bond angles are very close to the 109.57° angles between the tetrahedral patches. The effect of this difference is small until the patches become very narrow, and so for intermediate values of σ, the dodecahedral clusters are preferred because they have a greater number of bonds, *e.g.* for the 20-atom dodecahedron there is only one unused patch per particle, whereas the particles on the surface of diamond clusters often have two unused patches. Thus, the results for these clusters suggest that the tendency of the liquid to form structures involving five-fold rings is one of the reasons underlying the much greater difficulty of crystallization in the tetrahedral system.

Furthermore, as well as this competition between five- and six-fold rings, in this model the diamond and hexagonal diamond (the one-component equivalent of wurtzite) structures are practically degenerate. The structural difference between these two crystals is that all the six-fold rings in diamond have a form analogous to the chair isomer of cyclohexane, whereas in hexagonal diamond, some of these rings are analogous to the boat form of cyclohexane. Again, this further variety of local structural forms is not going to aid crystallization.

These results are consistent with the work of Zhang *et al.* who attempted to crystallize diamond using similar patchy particles with the same tetrahedral patch geometry.[29]

Crystallization to the diamond structure only reliably occurred when a seed crystal was introduced, or when an additional term in the potential was added that favoured staggered over eclipsed torsional configurations, and hence disfavoured hexagonal diamond.

This work on model anisotropic particles is particularly timely given the rapid recent advances in synthesizing anisotropic colloids,[14–16] and because one target for research in this field is to produce particles with tetrahedral symmetry to assemble into a diamond-like crystal.[78] Our results indicate that the crystallization of such tetrahedral colloids might not be so straightforward, because of the potentially frustrating effects of the variety of local structures possible in the liquid phase. However, it is not clear how an additional potential term similar to that used by Zhang *et al*. could be introduced into the intercolloidal potential to alleviate this. Instead, an alternative strategy would be to use a binary system of oppositely-charged tetrahedral colloids, as this would penalize the formation of rings with an odd number of particles. In particular, this change would reduce the tendency for pentagons to form, hence removing some of the frustration and making crystallization more likely.

An interesting comparison to the present results is provided by a recent study that took the Stillinger–Weber silicon potential[79] and varied the strength of the anisotropic 3-body term, finding that it has a significant effect on the system's glass-forming ability.[71] For the original silicon potential crystallization into a diamond structure occurs relatively easily, but as the strength of the 3-body term is decreased the system becomes a good glass-former in the region of parameter space where there is a crossover in the stability of the diamond and body-centred-cubic crystals, and it has been suggested that this is partially due to the structural dissimilarity between the liquid and the possible crystals. Similarly, it would be interesting to look at how varying the patch width affected the dynamics of crystallization for our systems.

5. Conclusions

Through the examples of crystallization (and its absence) considered in this paper, we hope to have shown the subtle interplay between the interparticle potential, the preferred local structure and the kinetics of crystallization. In particular, if the interactions favour a local structure that is incompatible with the global crystalline order, or almost equally favour a variety of different local environments, then crystallization is likely to be frustrated. Furthermore, we have provided further empirical support that examining the structures of isolated clusters, as first performed by Frank,[57] can potentially provide a clearer picture of the preferred local order, and hence the possible presence of frustration.

An important question concerning the type of structural approach advocated here is how it relates to other approaches to understanding the kinetics of crystallization, such as classical nucleation theory. In the latter, one of the key parameters in determining the ease of nucleation is the surface free energy associated with the crystal–liquid interface. The connection to the current approach is that this interfacial free energy is particularly sensitive to the degree of structural dissimilarity between the crystal and the liquid, *i.e*. the greater the structurally dissimilarity, the greater the interfacial free energy, and hence the harder crystallization becomes.

In such instances where the liquid and crystal are structurally very different, one way of circumventing the large free-energy barriers to direct nucleation of the crystal is to undergo a two-step nucleation process, *i.e*. first nucleate a metastable phase, and then nucleate the final crystal form from within that phase.[80,81] For example, for the Stillinger–Weber silicon potential mentioned earlier, crystallization to the diamond structure only occurs at temperatures below that for a liquid–liquid phase transition, which leads to the nucleation of a lower-density liquid phase that more closely resembles the solid.[71]

The patchy models introduced in Section 4 are particularly useful for studying crystallization, since the geometry of the patches allows the system's structural propensities to be directly controlled in a simple and flexible manner. One of our original intentions for these models was to also use them to illuminate aspects of the crystallization of proteins. Therefore, a key question is how relevant these models are to real proteins. Clearly, the interactions between proteins are strongly anisotropic in character, *i.e*. in order for two proteins (in their native state) to come together, the regions of their respective surfaces that come into contact, the ''patches'' if you like, must be correctly aligned and oriented. Furthermore, the patchy models, like proteins, are also able to form relatively low-density crystals. However, actual interprotein interactions are much more complex—proteins are anisotropic in shape, their surfaces are very heterogeneous, and the interactions can depend sensitively on solution conditions, such as pH and the concentrations of other ions.[82] Thus, how much of this complexity one needs to capture in order to reproduce or gain insight into the crystallization of proteins is an open, but important, question.

The examples described in Section 2 illustrate the exquisite control that biological systems are able to exert over the interaction properties of proteins, in particular their ability to avoid native state aggregation or crystallization in the dense cellular environment, and the ability of particular proteins to form into complex crystalline assemblies. There are some indications of how this control is achieved, *e.g*. the role of lysine residues in preventing unwanted interactions, but there is much still to be learnt. Such knowledge will be particularly important if rational methods to make protein crystallization *in vitro* easier are to be developed that aim to overcome the evolutionary selection of the surface properties of proteins to prevent native state aggregation.

In some ways, our patchy particles are probably somewhat closer to some of the anisotropic colloids now beginning to be generated.[14–16] Therefore, it will be particularly interesting, once it becomes possible to generate such colloids in sufficient number and quality, and with their surfaces appropriately functionalized, to be able to probe their phase behaviour and phase transformation kinetics. Hopefully, some of the insights gained from our patchy models will help to guide the experimentalists in their quest to create crystals with useful photonic properties.

Acknowledgements

The authors are grateful to the Royal Society, the Ramón Areces Foundation and the Engineering and Physical Sciences Research Council for financial support.

References

1 D. B. Miracle, *Nat. Mater.*, 2004, **3**, 697.
2 A. George and W. W. Wilson, *Acta Crystallogr., Sect. D: Biol. Crystallogr.*, 1994, **50**, 361.
3 D. F. Rosenbaum, P. C. Zamora and C. F. Zukoski, *Phys. Rev. Lett.*, 1996, **76**, 150.
4 P. R. ten Wolde and D. Frenkel, *Science*, 1997, **277**, 1975.
5 B. W. Matthews, *J. Mol. Biol.*, 1968, **33**, 491.
6 P. Bartlett, R. H. Ottewill and P. N. Pusey, *Phys. Rev. Lett.*, 1992, **68**, 3801.
7 M. D. Eldridge, P. A. Madden and D. Frenkel, *Nature*, 1993, **365**, 35.
8 P. Bartlett and A. I. Campbell, *Phys. Rev. Lett.*, 2005, **95**, 128302.
9 M. E. Leunissen, C. G. Christova, A.-P. Hynninen, C. P. Royall, A. I. Campbell, A. Imhof, M. Dijkstra, R. van Roij and A. von Blaaderen, *Nature*, 2005, **437**, 235.
10 E. V. Shevchenko, D. V. Talapin, N. A. Kotov, S. O'Brien and C. B. Murray, *Nature*, 2006, **439**, 55.
11 D. Frenkel, *Nat. Mater.*, 2006, **5**, 85.
12 A. M. Kalsin, M. Fialkowski, M. Paszewski, S. K. Smoukov, K. J. M. Bishop and B. A. Grzybowski, *Science*, 2006, **312**, 420.
13 C. N. Likos, N. Hoffmann, H. Löwen and A. A. Louis, *J. Phys.: Condens. Matter*, 2002, **14**, 7681.
14 V. N. Manoharan, M. T. Elsesser and D. J. Pine, *Science*, 2003, **301**, 483.
15 Y.-S. Cho, G.-R. Yi, J.-M. Lim, S.-H. Kim, V. N. Manoharan, D. J. Pine and S.-M. Yang, *J. Am. Chem. Soc.*, 2005, **127**, 15968.
16 A. van Blaaderen, *Nature*, 2006, **439**, 545.
17 A. M. Jackson, J. W. Myerson and F. Stellacci, *Nat. Mater.*, 2004, **3**, 330.
18 K.-H. Roh, D. C. Martin and J. Lahann, *Nat. Mater.*, 2005, **4**, 759.
19 C. E. Snyder, A. M. Yake, J. D. Feick and D. Velegol, *Langmuir*, 2005, **21**, 4813.
20 Z. F. Li, D. Y. Lee, M. F. Rubner and R. E. Cohen, *Macromolecules*, 2005, **38**, 7876.
21 K.-H. Roh, D. C. Martin and J. Lahann, *J. Am. Chem. Soc.*, 2006, **128**, 6796.
22 R. P. Sear, *J. Chem. Phys.*, 1999, **111**, 4800.
23 N. M. Dixit and C. F. Zukoski, *J. Chem. Phys.*, 2002, **117**, 8540.
24 X. Song, *Phys. Rev. E*, 2002, **66**, 011909.
25 N. Kern and D. Frenkel, *J. Chem. Phys.*, 2003, **118**, 9882.
26 J. Chang, A. M. Lenhoff and S. I. Sandler, *J. Chem. Phys.*, 2004, **120**, 3003.
27 J. Chang, A. M. Lenhoff and S. I. Sandler, *J. Phys. Chem. B*, 2005, **109**, 19507.
28 V. Talanquer, *J. Chem. Phys.*, 2005, **122**, 084704.
29 Z. Zhang, A. S. Keys, T. Chen and S. C. Glotzer, *Langmuir*, 2006, **21**, 11547.
30 Z. Zhang and S. C. Glotzer, *Nano Lett.*, 2004, **4**, 1407.
31 M. F. Hagan and D. Chandler, *Biophys. J.*, 2006, **91**, 42.
32 K. Van Workum and J. F. Douglas, *Phys. Rev. E*, 2006, **73**, 031502.
33 A. W. Wilber, J. P. K. Doye, A. A. Louis, E. G. Noya, M. A. Miller and P. Wong, http://arxiv.org/pdf/cond-mat/0606634.
34 J. P. K. Doye, A. A. Louis and M. Vendruscolo, *Phys. Biol.*, 2004, **1**, P9.
35 R. Hui and A. Edwards, *J. Struct. Biol.*, 2003, **142**, 154.
36 Z. S. Derewenda and P. G. Vekilov, *Acta Crystallogr., Sect. D: Biol. Crystallogr.*, 2006, **62**, 116.
37 P. G. Vekilov, A. R. Feeling-Taylor, D. N. Petsev, O. Galkin, R. L. Nagel and R. E. Hirsch, *Biophys. J.*, 2002, **83**, 1147.
38 A. Pande, J. Pande, N. Asherie, A. Lomakin, O. Ogun, J. King and G. B. Benedek, *Proc. Natl. Acad. Sci. U. S. A.*, 2001, **98**, 6116.
39 H. E. McElroy, G. W. Sisson, W. E. Schoettlin, R. M. Aust and J. E. Villafranca, *J. Cryst. Growth*, 1992, **122**, 265.
40 A. D'Arcy, M. Stihle, D. Kostrewa and G. Dale, *Acta Crystallogr., Sect. D: Biol. Crystallogr.*, 1999, **55**, 1623.
41 S. Jones and J. M. Thornton, *Proc. Natl. Acad. Sci. U. S. A.*, 1996, **93**, 13.
42 L. Lo Conte, C. Chothia and J. Janin, *J. Mol. Biol.*, 1999, **285**, 2177.
43 S. Dasgupta, G. H. Iyer, S. H. Bryant, C. E. Lawrence and J. A. Bell, *Proteins: Struct., Funct., Genet.*, 1997, **28**, 494.
44 K. L. Longenecker, S. M. Garrard, P. J. Sheffield and Z. S. Derewenda, *Acta Crystallogr., Sect. D: Biol. Crystallogr.*, 2001, **57**, 679.
45 Z. S. Derewenda, *Structure*, 2004, **12**, 529.
46 J. P. K. Doye and W. C. K. Poon, *Curr. Opin. Colloid Interface Sci.*, 2006, **11**, 40.
47 K. M. Smith, *Virus-Insect Relationships*, Longmans, London, 1976.
48 H. J. Arnott and K. M. Smith, *J. Ultrastruct. Res.*, 1968, **21**, 251.
49 L. Vayssié, F. Skouri, L. Sperling and J. Cohen, *Biochimie*, 2000, **82**, 269.
50 B. A. Mullens, R. K. Velten and B. A. Federici, *J. Invertebr. Pathol.*, 1999, **73**, 231.
51 G. Knoll, B. Haackebell and H. Plattner, *Eur. J. Protistol.*, 1991, **27**, 381.
52 R. Williams and K. D. Smith, *Nature*, 1957, **179**, 119.
53 A. Klug, R. E. Franklin and S. P. F. Humphreys-Owen, *Biochim. Biophys. Acta*, 1959, **32**, 203.
54 P. G. Debenedetti and F. H. Stillinger, *Nature*, 2001, **410**, 259.
55 J.-F. Sadoc and R. Mosseri, *Geometric Frustration*, Cambridge University Press, Cambridge, 1999.
56 G. Tarjus, S. A. Kivelson, Z. Nussinov and P. Viot, *J. Phys.: Condens. Matter*, 2005, **17**, R1143.
57 F. C. Frank, *Proc. R. Soc. London, Ser. A*, 1952, **215**, 43.
58 M. Dzugutov, *Phys. Rev. A*, 1992, **46**, R2984.
59 D. R. Nelson and F. Spaepen, *Solid State Phys.*, 1989, **42**, 1.
60 J. Roth and A. R. Denton, *Phys. Rev. E: Stat. Phys., Plasmas, Fluids, Relat. Interdiscip. Top.*, 2000, **61**, 6845.
61 M. Dzugutov, S. I. Simdyankin and F. H. M. Zetterling, *Phys. Rev. Lett.*, 2002, **89**, 195701.
62 J. P. K. Doye, D. J. Wales and S. I. Simdyankin, *Faraday Discuss.*, 2001, **118**, 159.
63 J. P. K. Doye, D. J. Wales, F. H. Zetterling and M. Dzugutov, *J. Chem. Phys.*, 2003, **118**, 2792.
64 J. R. Fernández and P. Harrowell, *Phys. Rev. E*, 2003, **67**, 011403.
65 J. R. Fernández and P. Harrowell, *J. Chem. Phys.*, 2004, **120**, 9222.
66 W. Kob and H. C. Andersen, *Phys. Rev. Lett.*, 1994, **73**, 1376.
67 T. F. Middleton, J. Hernández-Rojas, P. N. Mortenson and D. J. Wales, *Phys. Rev. B: Condens. Matter*, 2001, **64**, 184201.
68 W. N. Lipscomb, *Science*, 1966, **153**, 373.
69 J. R. Fernández and P. Harrowell, *J. Phys. Chem. B*, 2004, **108**, 6850.
70 H. Shintani and H. Tanaka, *Nat. Phys.*, 2006, **6**, 200.
71 V. Molinero, S. Sastry and C. A. Angell, *Phys. Rev. Lett.*, 2006, **97**, 075701.
72 C. de Michele, S. Gabrielli, F. Sciortino and P. Tartaglia, *J. Phys. Chem. B*, 2006, **110**, 8064.
73 E. Bianchi, J. Largo, P. Tartaglia, E. Zaccarelli and F. Sciortino, *Phys. Rev. Lett.*, 2006, **97**, 168301.
74 B. Grünbaum and G. C. Shephard, *Tilings and Patterns*, W. H. Freeman, New York, 1987.
75 D. P. Shoemaker and C. B. Shoemaker, in *Introduction to Quasicrystals*, ed. M. V. Jaric, Academic Press, London, 1988, pp. 1–57.
76 X. Zeng, *Curr. Opin. Colloid Interface Sci.*, 2005, **9**, 384.
77 T. Schilling, S. Pronk, B. Mulder and D. Frenkel, *Phys. Rev. E*, 2005, **71**, 036138.
78 T. T. Ngo, C. M. Liddell, M. Ghebrebrhan and J. D. Joannopoulos, *Appl. Phys. Lett.*, 2006, **88**, 241920.
79 F. H. Stillinger and T. A. Weber, *Phys. Rev. B: Condens. Matter*, 1985, **31**, 5262.
80 W. Ostwald, *Z. Phys. Chem.*, 1897, **22**, 289.
81 P. R. ten Wolde and D. Frenkel, *Phys. Chem. Chem. Phys.*, 1999, **1**, 2191.
82 E. Allahyarov, H. Löwen, A. A. Louis and J. P. Hansen, *Europhys. Lett.*, 2002, **57**, 731.