UNIVERSITY NAME (IN BLOCK CAPITALS)

Alex's Thesis

by

A thesis submitted in partial fulfillment for the
degree of Doctor of Philosophy

in the
Faculty Name
Department or School Name

May 2009

May 2009

Chapter 1

Abbreviations

HPV Human papillomavirus HBV Hepatitis B virus CCMV Cowpea chlorotic mottle virus CPMV Cowpea mottle virus BMV Brome mosaic virus MC (Metropolis) Monte Carlo DMD Discontinuous molecular dynamics TBSV Tomato bushy stunt virus TCV Turnip crinkle virus HSV Herpes simplex virus SEM Scanning electron microscopy cryo-EM Cryo-electron microscopy MD Molecular dynamics

Chapter 2

Introduction

2.1 Overview

The goal of this thesis is to advance the understanding of systems which self-assemble to produce monodisperse (uniformly sized) products. In the systems we will be considering the assembly is an emergent consequence of the interactions of the component particles, with no external ordering influence. All of the information required to specify the assembly product, and preclude the formation of competing products, must therefore be encoded in the design of the particles. However, the required features of an appropriate design can be far from obvious.

Our approach is to study self-assembly by means of simulations using a minimal computational model, eschewing the inclusion of the detail of particular biological or synthetic systems in the hope of obtaining good quality statistics which may have wide applicability. Our aims in this work are as follows:

- To examine the considerations that must be made when designing particles to assemble into a given target structure, and what features of biological systems are important to the assembly process.
- To investigate the traps and pitfalls which can prevent successful assembly.
- To compare our models to the details of biological viruses, in the hope that the analogy will provide insight in both directions.
- To explore the possibilities of what structures are likely to be possible to self-assemble, and what interactions would be required for them to do so.

The canonical example of monodisperse self-assembling systems in nature is the assembly of virus capsids, the protein shells which encapsulate the DNA or RNA of viruses.

Viruses present an astonishing example of assembly, with large numbers of proteins (from 60 to thousands) coming together to form highly ordered structures. The biology and structure of viruses are described in detail in Section 2.2. There are many potential advantages to an improved understanding of virus assembly, some of which are discussed in Section 2.3, while some complicating issues which may frustrate understanding are described in Section 2.4.

There is also some hope that in future it may be possible to create systems of synthetic colloids or nanoparticles which self-assemble into monodisperse structures. A fundamental requirement for such systems is the production of anisotropic subunits - isotropic subunits will always tend to form structures of infinite size. However, controlled anisotropy is very difficult to achieve, and work in this area is still at quite an early stage. Once the synthesis of nanoparticles or colloids with controllable anisotropy becomes a reality, we hope that the work done here may act as a guide towards designing self-assembling systems. Progress on the synthesis of such anisotropic particles is described in Section 2.5

Finally, previous experimental and theoretical studies of monodisperse self-assembly (predominantly that of viruses) are discussed in Section 2.6, and our own approach and the structure of the remainder of the thesis are described in Section 2.7.

2.2 Biology, structure and self-assembly of viruses

2.2.1 An introduction to viruses

Richard Dawkins suggested the idea of a *meme*[1], an item of information such as an idea or a joke which self-replicates in the environment of human minds. The meme's fitness can be taken to be the efficiency with which it propagates through the human population, and persists with time. Memes adapt to their environment; ideas and jokes may constantly mutate through mishearings, misunderstandings or intentional modifications, and if the mutated versions are fitter within a particular culture or time period, those versions will come to dominate.

The meme is a strong analogy for the gene. Both are essentially packets of information which have evolved to excel at self-replication and have thus been propagated through time and space. This viewpoint is a potent one in understanding biology, but of course this underlying theme is often obscured by the details of the mechanisms that genes cause to be constructed so as to further their reproduction. Viruses, however, embody this idea of self-replicating information through their startling minimalism. They have no cells and do not eat or respire; they do not satisfy many common definitions of life. Rather than copying their genome themselves, they rely on parasitising the mechanisms of living cells to do so, co-opting the natural mechanisms of their host cells they infect

in order to produce new viral particles. Nonetheless they are extremely successful, with a recent estimate suggesting that they are second only to prokaryotes in terms of total biomass on Earth[2], and viruses have even been discovered in the boiling acid environments of Yellowstone National Park[3].

Viruses consist of a bundle of nucleic acid (either DNA or RNA) packaged into a protein shell called a capsid. In some viruses, this capsid is itself wrapped in a lipid membrane known as a viral envelope. The capsid is generally either roughly spherical (for around half of the known virus families) or cylindrical, though the capsids of some viruses, notably bacteriophages, are more complex in shape. The surface of the virus is typically studded with proteins that bind with certain cell surface proteins, allowing specific infection of particular cells. Viruses are much smaller than bacteria, with the diameter of most spherical viruses lying in the approximate range of 20-200 nm while the size of bacteria tends to lie in the range of 0.5-5 μm. Indeed, viruses were first discovered as an infectious agent that could not be removed by filtration, and were thought for a time to be a liquid form of life.

2.2.2 The viral life cycle

The viral life cycle begins with the binding of the virus to a host cell via the interaction of proteins on the virus surface with specific receptors embedded in the membrane of the host cell. The virus then enters the cell by one of a number of possible mechanisms. "Naked" viruses (those without envelopes) may inject their nucleic acids through the cell membrane (as in bacteriophages), simply pass through the membrane intact, or be taken up into a vesicle inside the host cell, in which case they then fuse with the vesicle wall and pass into the cell proper. Enveloped viruses may also enter via vesicles, or their envelopes may fuse with the cell membrane depositing the capsid directly into the cell.

The details of what follows depend strongly on the specific type of virus. The viral particle may take advantage of intracellular transport mechanisms in eukaryotes before being broken down, releasing its nucleic acids either into the cytoplasm or the nucleus. The viral nucleic acids may be either DNA or RNA, they may be double stranded or single stranded, linear or looped, and single stranded nucleic acids may either code directly for proteins (in "positive strand" viruses) or may be the complementary ("negative") strand and require copying before proteins can be produced. In any case, by one mechanism or another proteins begin to be produced in the host cell based on the viral genome. These proteins will begin the process of forming new virus particles, by rapidly copying the viral genome (or causing it to be copied by host proteins). The viral proteins may also have a number of effects on the host cell, for example preventing the transcription of host DNA or even cutting up the host DNA so that it does not compete with the viral genome for transcription. Smaller viruses, which contain less of their own genetic data, are often reliant on the continued function of many host processes in order to facilitate

their reproduction, while the genomes of larger viruses may themselves code for most of the proteins required for successful reproduction.

A critical step in the production of new viruses, and of course the step most relevant for this thesis, is the assembly of capsids for the offspring virus particles. We shall return in Section 2.2.4 to consider this remarkable process in more detail, but for now we shall press ahead to conclude the description of the viral life cycle. Suffice to say that the capsids, composed of proteins coded for by the viral genome, assemble spontaneously and in their finished state form closed structures containing new copies of the viral nucleic acids.

The viral life cycle is completed by the departure of the new viral particles from the host cell. In many cases this occurs by lysis - once the cell is sufficiently full of completed viral particles its membrane ruptures, releasing the particles to begin a new round of infection. In the case of enveloped viruses the viruses instead migrate to the cell membrane which then forms vesicles around the virus particles, in a process assisted by viral proteins which will already have migrated to the membrane. The viruses then bud off from the host cell, their envelopes formed from the host cell membrane. In this case cell lysis may not occur, and the cell may survive for some time, in some cases indefinitely.

2.2.3 Viral structure

Viruses are characterised by the short lengths of nucleic acid they are able to carry, leading to a need for very efficient use of their own genetic information along with a tendency towards profligate use of host resources. For example many viral capsids may consist of many copies of the same protein, so that only a single gene is required to code for the capsid structure. In other cases, generally for larger viruses with room for more DNA/RNA, the capsid may consist of several different proteins. Scaffolding proteins, which are not present in the assembled virus, may also be required for assembly.

Around half of all virus families have "spherical" capsids, and it is this group that we shall be exclusively considering. Despite the name, spherical capsids in fact generally possess icosahedral symmetry. However, while there are only 60 equivalent sites available under icosahedral symmetry (three for each face of the twenty faces of the icosahedron), viruses are found to contain $60T$ asymmetric units, with T an integer, each of which consists of one or more proteins.

The arrangement of proteins in an icosahedral capsid was first explained by Caspar and Klug in 1962 with their quasiequivalence theory[4]. The underlying postulate of quasiequivalence theory is that the capsid proteins can form both pentamers and hexamers using the same intersubunit contacts. The surface of the capsid can be considered to be composed of triangles, each consisting of three asymmetric units. The corners of

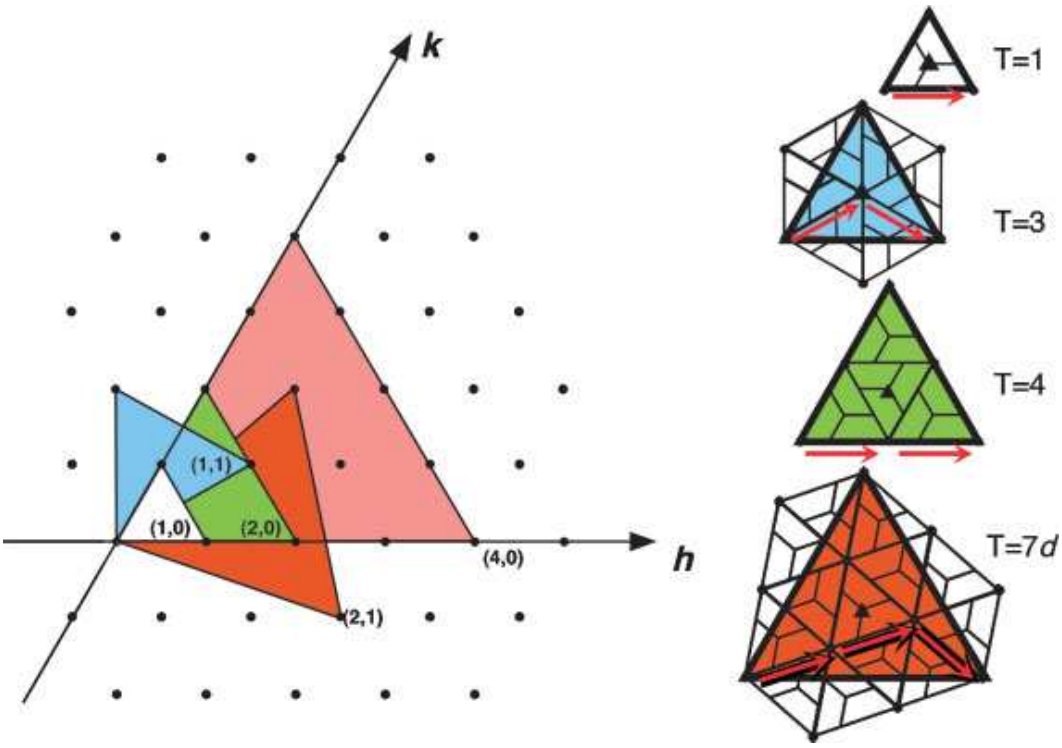

FIGURE 2.1: Facets for icosahedra with quasi-symmetry. The facets are equilateral triangles with an integral number of subunits. The geometry of each facet is described by the hexagonal coordinate system (left). On the right are selected facets, drawn with trapezoidal subunits. The arrows indicate a system of determining the h and k indices by counting vertices from fivefold to fivefold. Reproduced from Ref. [6].

these triangles can join at either sixfold or fivefold sites. The sixfold sites provide no curvature, while the fivefold sites provide curvature such that a closed capsid must always contain exactly twelve fivefold sites. Thus the smallest possible icosahedral capsid has no sixfold sites and twelve fivefold sites, giving a straightforward icosahedron with just one set of three asymmetric units per face for a total of 60 asymmetric units. However, it is also possible to fit more sets of three asymmetric units per icosahedral face, as shown in Fig. 2.1. The T number gives the number of sets of three asymmetric units per icosahedral face, and is given by $T = h^2 + hk + k^2$, where h and k are non-negative integers, so that $T = 1, 3, 4, 7 \ldots$. The sites of the asymmetric units are quasiequivalent because in general some are involved in sixfold sites while others contribute to fivefold sites, so that their positions are not strictly equivalent. A complete capsid will contain a total of $10(T-1)$ hexamers and 12 pentamers. The T number may take values as high as 219[5] and probably beyond; the size of viral capsids varies enormously.

2.2.4 Capsid self-assembly

Viral capsids are stabilised by interactions between the individual capsid proteins that constitute them. The proteins generally interact through large regions of their surfaces

which fit together well, and consist of a higher than average proportion of hydrophobic residues. Consequently when they come together water is excluded from the vicinity of these hydrophobic residues, which is entropically favourable (since water forms ordered, ice-like structures near hydrophobic surfaces). These regions of the protein surfaces can therefore be considered as "sticky patches". The proteins are of an appropriate shape to bind with a number of neighbours in the correct orientations to form the capsid, or at least conform to such a shape as the capsid is assembled.

In vivo the capsid proteins are produced by ribosomes in the host cell and then may either be localised to a cell compartment by cell transport mechanisms, or simply be released into the cytoplasm. Then, somehow, these individual proteins must arrange themselves into large monodisperse structures, with high yield and with few errors if the virus is to be viable.

A few possibilities come to mind immediately. Perhaps a hierarchical scheme is employed, in which the capsid proteins (also known simply as subunits) first arrange themselves into groups, which then themselves assemble into complete capsids. In fact this is known to be the case for some viruses, for example the subunits of the bacteriophage HK97 first form pentamers and hexamers. These pre-assembled groups of subunits are known as capsomers. Another possibility might be that the nucleic acid acts as a template, forming a core around which the subunits arrange themselves. Again, this has been found to be true for some viruses[**?**]. Unfortunately a recurring theme in virus assembly is that the behaviour and methods employed are hugely diverse; at best, we can hope to identify and explain common trends, but exceptions will always abound.

Beyond the first ideas of possible assembly mechanisms described here - which fall far short of explaining assembly fully - it is necessary to turn to experiments and simulations. A summary of the understanding gleaned so far from such work is presented in Section 2.6.

A crucial issue in forming viable viruses is that of how the nucleic acid ends up inside the capsid. Viruses take one of two approaches - either the capsid is assembled around the nucleic acid, or the capsid is assembled empty and then nucleic acid is pumped in. Most viruses containing single-stranded nucleic acids take the former approach. Assembly requires the simultaneous presence of the capsid subunits and the nucleic acid, and the nucleic acid is thought to act as a nucleus for assembly.[**?**]. As evidence, highly basic amino acid residues, which would be expected to interact favourably with nucleic acids, are often located near to the inner capsid surface.

Many viruses containing double-stranded nucleic acids, by contrast, take the latter approach and assemble via a two-step mechanism involving a portal motor. In the first step an empty capsid assembles. During the assembly of the capsid a single portal motor protein is incorporated into it (raising questions as to how the addition of further portal

motors is prevented[7]). In the second step the motor then pumps the nucleic acid into the capsid, using energy in the form of ATP in the process.

Incidentally, the different mechanisms of genome packing are thought to be important in determining the mechanism of infection. For example bacteriophages pack their genome using a portal motor, and infect host cells by injecting the genome through the cell membrane, making use of the high osmotic pressure inside the capsid built up during packing.

As well as the possible dependence on the presence of nucleic acids, some viruses depend on scaffolding proteins for successful assembly[8–10]. Scaffolding proteins are proteins which act to assist in the assembly of well-formed capsids, but are eventually lost and are not present in mature virus particles. They are believed to act either by binding groups of capsid proteins together to facilitate assembly, or by inducing conformational changes in capsid proteins.

While some viruses certainly need nucleic acids or scaffolding proteins to assemble, a number of viruses such as cowpea chlorotic mottle virus (CCMV)[11], hepatitis B virus (HBV)[12] and human papillomavirus (HPV)[13] have been shown to assemble *in vitro* from purified protein, in the absence of nucleic acids or any other possible assembly assistants. The importance of this point for our own work can hardly be overstated. It demonstrates that it is possible to encode all of the information necessary for assembly in the structure of a capsid protein alone, and that the proteins are able to come together and form correct structures in high yield without the guidance of any other molecules. This evidence that capsid assembly does not always rely on the complexities of the cellular environment implies that, at least in the simplest cases, it may represent a relatively tractable target for study via theory and simulation.

2.3 Potential applications of understanding virus assembly

Possibly the most most obvious application of improved understanding of viral assembly would be the development of drugs to interfere with the process[14]. Viral assembly has not yet been routinely targeted in the development of antiviral drugs. This is partly because protein-protein interfaces have traditionally not been considered viable targets for modification by small molecules, although there is growing evidence to show that it is possible,[15], and partly because of difficulties in screening for potential drugs in vitro. The historically poor understanding of viral assembly is also likely to have been an important contributing factor. However, drugs which interfere with assembly ultimately offer great promise in the possibility of strong therapeutic effects with few side effects. A general problem in the design of anti-pathogen drugs is in finding pathogenic processes which have no analogue in healthy cells, so that the processes may disrupted without danger to the patient. Antiviral drugs thus far have tended to focus on viral proteins

such as polymerases and proteases, which present difficulties in some cases because of the existence of related proteins in healthy cells. However, healthy cells produce no analogue to viral capsid proteins, making them a highly suitable target for drug action.

There has been some experimental progress in this area, although at present the number of compounds investigated remains very limited. The molecule bis-ANS has been found to bind to dimers of hepatitis B virus (HBV) capsid protein[16]. Dimers thus affected no longer contribute to capsid assembly, although they do aggregate into noncapsid polymers. bis-ANS does not appear to block the subunit-subunit binding interfaces[17], but rather is thought to act as a "molecular wedge", altering the angle between the binding surfaces of the dimers so that they can no longer assemble correctly. Another class of molecules, heteroaryldihydropyrimidines (HAPs) have also been found to be effective in inhibiting HBV replication[18]. HAP promotes the formation of sixfold as opposed to fivefold sites in capsid protein polymers, leading to reduced curvature. At low concentrations this causes small deformations in capsids which can generally be tolerated. However, at larger concentrations the capsid proteins tend instead to form spiral structures, tubes or sheets. In the presence of high concentrations of HAP even completed capsomers may dissociate to form lower curvature structures.

Zlotnick and coworkers have recently published experimental protocols for the performance of in vitro assays for the screening of compounds which may interfere with assembly[15, 19]. Hopefully the combination of such screening approaches with increasing understanding of the assembly process may lead to valuable antiviral drugs, with great implications for the alleviation of human suffering.

Virus capsids find some use as vaccines, for example empty capsids of human papillomavirus (HPV) are used as vaccines aimed at reducing the incidence of cervical cancer, which is believed to often be caused by the virus[20]. The efficacy of the vaccine is found to depend strongly on the extent of assembly of the capsids[21]. However, the assembly of capsids cannot yet be effectively controlled in the laboratory, leading to the formation of many misformed structures as well as correctly formed capsids[22]. Better understanding of assembly would certainly be helpful in improving this and other related laboratory procedures.

However, an understanding of capsid assembly may prove to be just as important in aiding the assembly and disassembly of highly modified capsids, adapted for new purposes, as it is in fighting disease-causing viruses in the wild. A number of emerging areas of research propose to use capsids as tools in medicine and materials science[23–25]. In almost all of these cases the assembly of the capsids is a critical step, and an improved understanding of how it can be controlled, and in what ways the capsid can be modified without disrupting assembly, may be useful for many of them.

Viruses have a number of features which make them ideal for use as highly specialised containers, especially for medical applications. They are naturally adapted to encapsulate cargo (viral nucleic acid), and to release it under certain conditions. The cargo can be replaced by synthetic cargoes, such as drugs[26], DNA or RNA for gene therapy[27] or contrast agents for visualisation techniques such as MRI[28]. Viruses naturally enter and release their cargo into specific cell types, and they can be modified to target different cells. A huge number of possible medical applications can be envisaged, such as delivering drugs or contrast agents to tumors, delivering wild type genes to the lungs of cystic fibrosis sufferers[**?**] or even delivering sources of free radicals to bacteria in the body[**?**].

The chemical surfaces of capsids are exceptionally easy to modify because of the possibility of genetic as well as chemical modification. Specific residues can be placed as desired on the interior or exterior surface of a capsid, or even the interfaces between subunits, and may then be used as a starting point for chemical modification. Further, considerable changes have been found to be possible without disrupting the structure of the capsids. Each of the capsid surfaces offers its own potential for unique applications.

The interior interface of the capsid can be used for directing the encapsulation[29**?**] or even synthesis[30, 31] of both organic and inorganic materials. Wild type viruses have positively charged interiors, suited to interactions with negatively charged nucleic acids, but the positively charged residues can easily be replaced by neutral or negatively charged residues, each allowing for the packaging of different varieties of cargoes. The interior surface also provides a rich and highly repetitive surface for the encapsulation of cargo molecules through covalent attachment to modified residues. Medically relevant small molecules such as drugs and imaging agents can be chemically attached to these reactive functional groups[32–34].

The external surfaces of many capsids naturally display peptides responsible for cell-specific recognition. These peptides may be modified or replaced with other targeting peptides to tailor the specificity of the capsid to the desired target cell types. Such modifications have been shown to successfully impart specificity[35, 36]. Other biologically useful molecules such as nucleic acids[37], carbohydrates[38] and fluorophores[32, 34] have also been successfully added to viral exterior surfaces.

Finally, the interfaces between capsid subunits may be modified, although of course the potential for modification here is far more limited since the virus must remain stable and capable of self-assembly. As an example, in the capsid of cowpea chlorotic mottle virus (CCMV) there are 180 metal-binding sites which normally bind Ca^{2+} at the points where three proteins come into contact[39]. Gd^{3+} can instead be bound in these sites, and the resulting modified capsids show exceptional properties as MRI imaging agents, producing the highest water proton relaxivities yet reported for a molecular

paramagnetic material[40] (although the Gd^{3+} was not sufficiently tightly bound for use *in vivo* and so further work is still needed).

These possibilities for surface modification also tie into a wide range of opportunities in materials science. For example, capsids may be used for the formation of ordered 2D or 3D arrays[41–43], as templates for the synthesis of highly monodisperse nanoparticles[31, 44], to encapsulate quantum dots[45], as enzyme nanoreactors[46] and even as binary memory storage devices[47, 48]. While this thesis concentrates on spherical viruses, helical viruses such as tobacco mosaic virus offer a great many other potential applications such as the templating of nanowires[30] and the formation of light-harvesting devices[49].

The medical implications of being able to inject therapeutics or imaging agents which will target specific tissues while remaining biologically inaccessible to the rest of the body are clearly extraordinary. Since viruses by their nature remain stable in physiological conditions until they come into contact with specific receptors, at which point they enter the cell and deliver their contents, they are ideal for such applications. Highly toxic chemicals, for instance cell-killing drugs, could be delivered to tumor cells with few side effects. However, a number of problems still stand in the way of such treatments. One of the major difficulties is the reaction of the body's immune system to the virus particles. In the case of animal viruses such as human adenovirus, the injection of large doses of the particles can cause harmful side effects including liver damage, even when the viruses are incapable of infecting cells[**?**]. Preliminary studies indicate that such side effects may be avoided by using plant viruses or bacteriophages[50]. Apart from causing side effects, the immune system may also destroy many of the virus particles before they can reach their destination. However, it is expected that this issue may be overcome in a number of different ways, for instance by masking the surface of the virus with humanised proteins or polyethylene glycol (pegylation)[38]. The immune response may even be useful in clearing up those viral particles which have not successfully delivered their cargoes to target cells. Other difficulties include the challenge of producing large quantities of noninfectious viral capsids; current laboratory techniques have been found to be difficult to scale up to an industrial level. Possible approaches include culturing recombinant microbes, or even the harvesting of large masses of infected plant matter. In any case, these issues are not expected to be insurmountable, and the potential rewards in the alleviation of human suffering are very great indeed.

2.4 Complicating issues in virus assembly

For the majority of this thesis we shall consider virus assembly to be a process taking place within at least moderately simple systems. Taking inspiration from the successful *in vitro* assembly of a number of viruses in the absence of nucleic acid or other proteins,

we shall for the most part focus on the complexities of the emergent behaviour which arises even from such apparently straightforward single-component systems. Other computational work on the subject (described in detail in Section 2.6.2) has tended to take a similar approach. However, it is important to bear in mind that for many viruses the assembly process is considerably more complex, with many more species involved. We give a few examples here to illustrate the great variety present in the field, and the huge challenges that lie ahead in extending theoretical and computational models beyond the simplest cases.

The polyomavirus provides an initial example of a somewhat surprising interaction between viral and cellular proteins. *In vivo* the polyomavirus capsid is found to assemble only in the nucleus and will not form in the cytoplasm of a host cell, presumably to ensure that the capsid forms in the presence of viral DNA. The mechanism for this is believed to involve the cellular heat-shock chaperonin hsc70[51]. The chaperonin binds to pentamers of the capsid protein and blocks assembly by steric interference. It accompanies the pentamer into the nucleus where, presumably because of changed conditions, assembly occurs. Assembly of polyomavirus capsid *in vitro* from purified protein is found to produce capsids of very variable structural accuracy, whereas in the presence of hsc70 efficient assembly is observed. The virus therefore appears to rely on the cellular hsc70 protein both to prevent assembly in the cytoplasm and to assist accurate assembly in the nucleus. hsc70, of course, has its own entirely different function in the cell. This is an instance of the surprisingly complex interprotein interactions which appear to be common to many virus assembly processes and which are likely to frustrate complete understanding via simple models.

An important role in the assembly of many viruses is played by proteins loosely grouped together as "scaffolding proteins". Scaffolding proteins are defined as proteins which are involved in the assembly process of the capsid, but are not present in the mature capsid. Scaffolding proteins may act by a specific catalysis of the binding of capsid proteins, or by nucleating and promoting assembly. They may also act to prevent the formation of incorrect protein-protein interactions, in an analogous role to cellular chaperonins (which act to prevent protein misfolding). Scaffolding proteins can be roughly divided into two major groups: icosahedrally orded, external scaffolds such as those of bacteriophage P4 and ϕX174, internal core-like scaffolds such as those of P22, herpes and λ phage. Internal scaffolds, which are often disordered and fail to share the symmetry of the capsid, appear to be more common but are somewhat harder to study. Very helpful review papers on scaffolding proteins have been written by Dokland[52] and by Fane and Prevelige[53].

As described in Section 2.2.1, a large number of viruses pump their genome into a complete capsid by means of a portal motor incorporated in the capsid. However, the incorporation of the portal motor has not yet to our knowledge been considered in any computational model. As an example of this process and its possible relevance to the whole assembly process we will consider the case of herpes simplex virus type 1

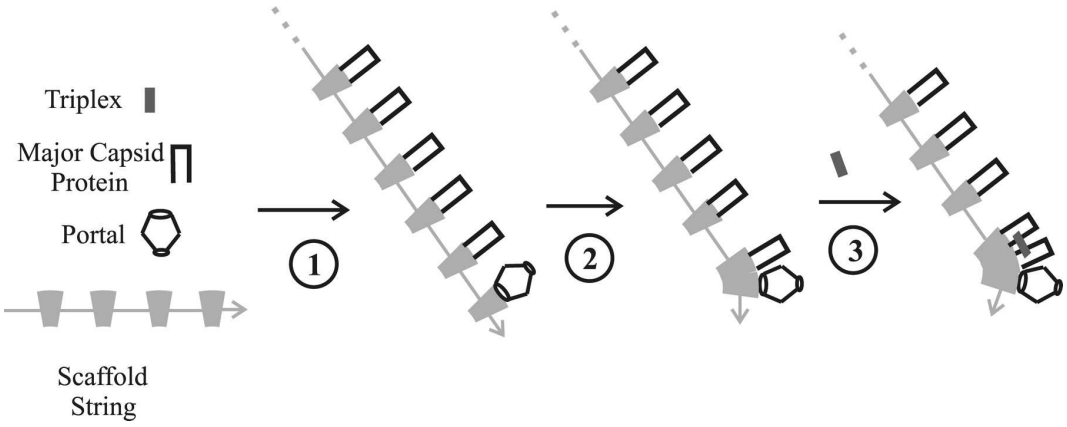

FIGURE 2.2: A schematic diagram indicating a proposed mechanism by which portal incorporation is limited to, but not necessary for, the initiation step of capsid assembly. The model postulates that (i) the scaffolding protein forms a head-to-tail polymer and (ii) the portal attaches only at a specific site at one end of the polymer (step 1), with the remaining capsid protein molecules bound to all other scaffold binding sites. Procapsid formation is suggested to occur by successive steps in which one capsid protein molecule at a time is delivered to a growing edge of the nascent procapsid (steps 2 and 3). Reproduced from Ref. [7].

(HSV-1). The capsid of HSV-1 consists of 162 capsomers (rapidly assembled groups of subunits): 150 hexons that make up the faces and edges of the icosahedron, 11 pentons occupying vertices, and a single portal found at the remaining vertex. When *in vitro* assembly experiments are carried out in the presence of the portal protein, correctly formed capsids are produced with at most one portal per capsid[7]. If portal protein is absent well-formed capsids are still produced, but with a penton in place of the missing portal. In experiments where the portal protein was added at a time t after the beginning of the experiment it was found that the eventual yield of portal-containing capsids fell off with increasing t up to 15 to 30 minutes, after which point the value of t had no impact, even though by that early stage there were few visible part-formed capsids (the total allowed assembly time was 20 hours). A suggested explanation of this rather difficult set of observations is summarised in Fig. 2.2. It is proposed that a scaffolding protein forms extended chains by joining head-to-tail, and binds the major capsid proteins along its length. The portal protein is bound selectively at the end site. The proteins bound to the scaffold then begin to come together to form a capsid nucleus. Thus the portal protein has to be present during the early stages of the process in order to occupy its unique site on the scaffold. In other cases, such as that of the bacteriophage T4, the portal protein is known to be involved in the initiation of capsid formation[54, 55], while in yet others the portal appears to be important in regulating the size and shape of the capsid[56]. It seems that this unique vertex may have an important role in the assembly processes of many viruses.

The subunits in any capsid larger than a $T = 1$ structure occupy a set of quasiequivalent positions which do not experience the same local environments. In general it is believed

that the proteins occupying different sites are in different conformations, such that they able to fit well and stably into the different environments. This immediately raises the questions of how the proteins come to exist in the different conformations in the correct proportions, how they come together in the precise spatial arrangement required, and whether the details of the process of conformational change are important to assembly. It is believed that in many cases the different conformations of the capsid proteins exist in equilibrium in solution, and that interactions with the growing edge of a capsid determine which conformation may bind there. In the cases of Tomato Bushy Stunt Virus (TBSV), Turnip Crinkle Virus (TCV) and CCMV, the N terminals of the capsid proteins have been found to interact with neighboring proteins in such a way as to provide a "switch" to promote the binding of only certain conformations depending on the position within the capsid[39, 57]. In general the differences in interaction energies resulting from changes in conformation are believed to be small, raising questions as to whether energetic factors are sufficient to explain the ordering of conformations. However, recent computational work by Elrad and Hagan[58] showed that energy differences as low as the thermal energy were sufficient to produce capsid-like structures with robust ordering of conformations in the appropriate sites. Intriguingly, it has been shown that in some cases capsid proteins are able to assemble into capsids of different T-numbers depending on the cargo they are encapsidating[59]. This implies that the choice of conformation for each protein depends on the relative energetic favourability of the different states, and that the presence of cargoes affects the relative energies such that capsids of an appropriate curvature are favoured.

All computational models of assembly thus far which explicitly include the positions of particles have assumed that assembly occurs in solution, and generally include only capsid proteins and solvent (explicit or implicit). This representation omits two potentially significant features of assembly in vivo: (i) the cellular environment is actually extremely crowded, which affects the dynamics of the diffusion of the proteins, and (ii) many capsids assemble in restricted environments, such as on membranes or in "virus factories", bodies within the cell which contain high concentrations of viral proteins.

Happily, some simple known examples of capsid assembly avoid all of these issues. The aforementioned *in vitro* assembly of CCMV, HBV, HPV and others does not require the presence of nucleic acids, scaffolding proteins, cell membranes or any other molecules, is completely symmetric (no portal protein) and is not known to involve any conformational changes. These examples justify the use of relatively simplistic approaches towards understanding virus assembly, which is fortunate since at present such approaches are the only viable option. A complete understanding of these simple examples would give considerable insight into the underlying principles of self-assembly processes and might have significant application for both synthetic self-assembling systems and for understanding more complex viral systems. The issues raised in this section, however, show that there are enormous obstacles to understanding or even modelling the full range of

viral assembly processes. In some cases, it is likely that the details may not in fact be too important, and models which ignore them may still be very helpful. In others, however, such as that of the HSV-1 scaffolding chain, the enormous impact on the assembly mechanism may render simplistic approaches entirely unhelpful.

2.5 A second realm of applications: Synthetic monodisperse self-assembling systems

2.5.1 Background, objectives and applications of the field

The synthesis and applications of colloids have been areas of significant interest for several decades. These colloidal particles have traditionally been spherical and possessed isotropic interactions, not least because of the difficulties in generating controlled anisotropy. However, over the last two decades there has been a drive towards more complex colloidal and nanoparticle structures leading to an incredible diversity of structures, including spheres with different functionality on each hemisphere[60], rods[61], cubes[62–64], icosahedra[65], tetrahedra[66], triangular prisms[67–70], tripods[71], tetrapods[61, 71–73], stars[74], X shapes[72], divalent spheres[75] and many more being synthesised. Fig. 2.3 shows many examples from the literature.

A significant focus of colloid science has been the exploration of the possible structures of colloidal crystals. By default, colloids tend to form close-packed fcc crystals. However, through work on tuning the interactions between colloids the range of possible crystals has been greatly extrended. The isotropy of interactions, however, remains a major limiting factor and the addition of anisotropic particles to the experimental toolkit would greatly increase the options for those studying colloidal crystals. This has been a major driving factor in the development of more complex colloids. A particular goal has been the assembly of colloidal diamond, an array of colloids with the same crystal structure as diamond, which has an optical band gap analogous to silicon's electronic band gap. Colloidal diamond is expected to have many applications in the field of optronics, but for now its synthesis remains elusive (although samples have been produced by the laborious methods of direct drilling[90] and layer-by-layer lithography[91]). A Brownian dynamics simulation by the Glotzer group has indicated that colloidal diamond may feasibly assemble from particles with four attractive patches located with tetrahedral symmetry, especially if seed crystals are introduced or if some kind of dependence on the torsional angle is included the interaction potential[92].

While recent developments in synthesising anisotropic colloids have been largely driven by the potential assembly of crystalline structures, they may also enable the assembly of finite-sized clusters. Some approaches have recently succeeded in producing patterns over the entire surface of particles, e.g. four patches arranged with tetrahedral symmetry.

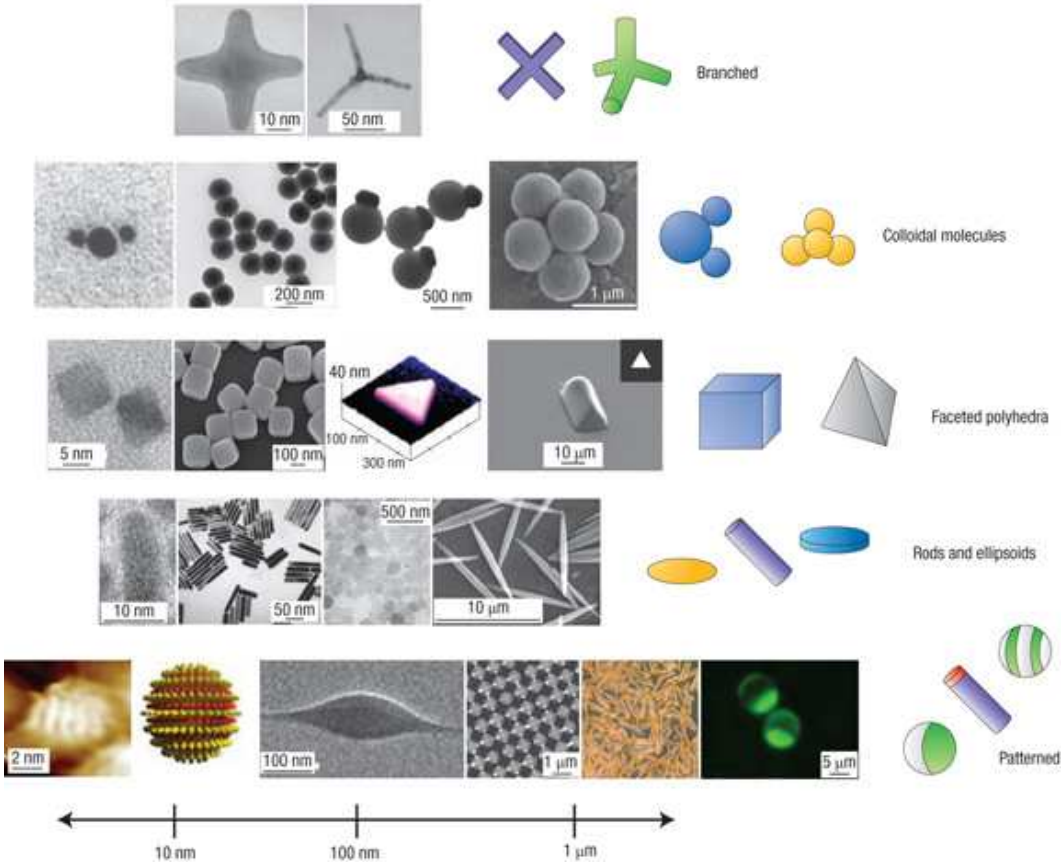

FIGURE 2.3: An array of recently synthesised anisotropic colloids and nanoparticles, organised into rough categories. The particles increase in size from left to right in rough correspondence with the scale at the bottom of the figure. From left to right, top to bottom: gold[71] and CdTe[73] tetrapods; DNA-linked gold nanocrystals[76], silica dumbells[77], asymmetric spherical oxide dimers[78], and fused clusters of polystyrene microspheres[79]; PbSe[64] and silver[62] cubes, gold[68] and polymer[70] triangular prisms; rods and ellipsoids composed of CdSe[80], gold[81], gibbsite[82] and polymer latex[83]; striped spheres[84], biphasic rods[85], patchy spheres with gold dots[86], Au-Pt nanorods[87] and bifunctional spheres[60, 88]. Reproduced from Ref. [89].

However, in order to form finite-sized clusters it will be necessary to produce particles with patches on only one side. There has so far been very little progress in this area. When a group does succeed in producing such particles and assembling an ordered finite-sized structure it will be likely to attract a great deal of attention. Unfortunately the applications of such structures are not yet clear - the current focus of most research groups has been on improving synthetic techniques rather than considering potential applications, since such considerations would in any case be highly hypothetical. Most likely the particles may be used as an alternative to virus capsids in carrying cargoes, for example drugs or imaging agents, or perhaps as templates in the synthesis of advanced materials.

In the following section we will describe a few examples of synthetic approaches to the manufacture of anisotropic particles which seem to have some potential in assembling

finite-sized clusters, or are otherwise of exceptional interest. Note that this is a vast and rapidly growing field, with literally hundreds of papers recently published demonstrating new syntheses. Here we will only attempt to give a brief summary of some of the more remarkable or promising approaches which have been taken.

2.5.2 Experimental progress and outlook for synthetic anisotropic particles

The development of anisotropic colloids and nanoparticles can be divided into two major approaches, focussing respectively on shape-anisotropy and on anisotropy in surface material or functionalisation.

We first focus briefly on the synthesis of shape-anisotropic particles. The development of shape anisotropy in itself may not represent a promising approach towards assembling ordered finite-sized structures; we wish to control the directions in which particles can interact, and so anisotropic functionalisation may appear to be a more promising route. However, the formation of shape-anisotropic structures may go hand-in-hand with the creation of variations in surface functionality, or the shape may allow uneven functionalisation after synthesis.

Manna et al.[73] have demonstrated a reliable method for the production of inorganic tetrapods of the order of 100 nm across (as shown in the first row of Fig. 2.3). The tetrapods consist of two different crystal structures of the same material, CdTe, with the central junction and the arms comprising zinc blende and wurtzite crystal structures respectively. Conditions are chosen such that nucleation produces zinc blende crystals, and growth then proceeds with a wurtzite structure. Phosphonic acid molecules bind selectively to the sides of the wurtzite arms, promoting extension rather than broadening of the arms. Since the terminal facets of the tetrapods remain unprotected, it may be possible to functionalise tetrapods with a "sticky" molecule, for example DNA, providing a possible building block for a diamond-like crystal.

Showcasing an entirely different, top-down approach, a very attractive method of synthesising two dimensionally shaped polymer particles on the scale of tens of micrometers has recently been developed[70]. The technique makes use of microfluidic techniques to overcome limitations associated with batch processes. Fig. 2.4(a) shows how flowing photocurable monomer is exposed to bursts of UV radiation, with each burst screened by a photomask. The transparent regions of the photomask comprise a row of identical shapes. As a result each burst of UV irradiates a set of corresponding shapes in the monomer flow, causing polymerisation. The flow moves on such that the next burst will produce a separate set of polymer shapes. As a result very large numbers of well-formed

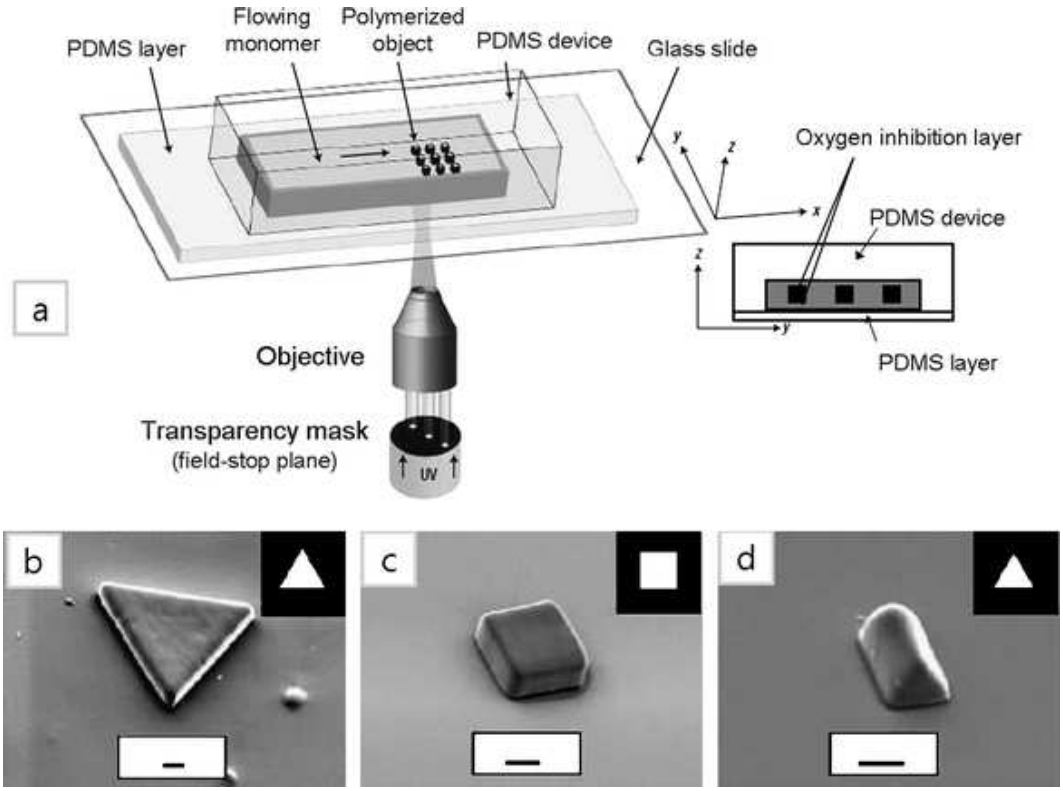

FIGURE 2.4: The production of polymer prisms in flow by UV curing. (a) A schematic illustration of the method. (b-d) SEM images of triangle, square and triangular rod particles. The insets in (b-d) represent the feature shapes of the corresponding photomasks, while the scale bars are 10 μm. Reproduced from Ref. [93], adapted from Ref. [70].

shaped particles are produced, as shown in Fig. 2.4(b-d). However, the resulting particles are relatively large, their shapes are limited to prisms, and they do not immediately suggest suitability for any self-assembly process producing finite size targets.

We now turn to the production of colloids and nanoparticles with patterned surfaces, whether via the anisotropic functionalisation of spherical particles, or the coassembly of several components into a single, roughly spherical particle.

Named after the two-faced Roman god of doors, gates, beginnings and endings, Janus particles are particles where one face (e.g. one half of a spherical particle) has different properties (such as material or functionalization) to the other. Many different methods have been developed to synthesise Janus particles, as shown in Fig. 2.5. The most well-studied approach is that of partial protection, where one face of the particle is protected, often by a film or gel, while the other face is functionalised[94–96]. If a stable directional flux of the modifying agent is achievable, as for example in the deposition of a metal layer, no protection is necessary[97, 98]. These methods are shown in Fig. 2.5(a).

An alternative approach was demonstrated in 2005 by Perro et al.[99], and is depicted

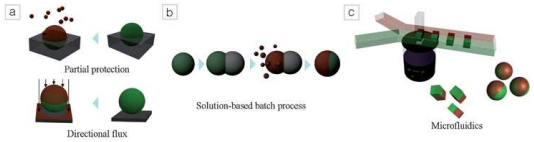

FIGURE 2.5: A schematic of some major approaches to synthesising Janus particles. (a) Synthesis on a surface, in which functionalisation is limited to one side of the particles either by partial protection, or by use of a directional flux. (b) A solution-based batch process in which the formation of a snowman-like structure protects one side of a particle. (c) Two polymer flows running side by side in a microfluidic system can act as a source of Janus particles either by lithography like that depicted in Fig. 2.4, or by jetting through a nozzle to create spherical particles. Adapted from Ref. [93].

in Fig. 2.5(b). 80 nm silica nanoparticles are used as components in snowman-like structures consisting of adjoined spheres of silica and polystyrene, produced using a process known as emulsion polymerisation. The exposed regions of the silica surface are then functionalised, while regions in contact with the polystyrene are protected. The polystyrene is then removed by ultrasonication and ultracentrifugation. Further treatment allows modification of the surface chemistries of the two faces of the resulting Janus properties to achieve desired properties. This method allows the batch production of large quantities of Janus particles.

A very exciting new method of producing Janus particles is offered by microfluidic devices, in which distinct polymer solutions flow side by side, as depicted in Fig. 2.4(c). UV polymerisation can then be used to produce shaped particles made of two different materials as with the single flow system described above[70], or alternatively the flow may pass into a nozzle[100]. In the latter case the polymer solutions can be ejected from the nozzle at high speeds, such that the liquid forms into spherical droplets in flight. Moreover, three-component colloids could be prepared using a sytem with three parallel flows[101], which may ultimately allow among other things the formation of particles with ring-shaped regions of functionality.

The groups of David Pine and Sueng-Man Yang have collaborated to develop an attractive method for producing composite colloids which may act as patchy particles[**? ?**]. A water-in-oil emulsion is mixed with both microspheres (either functionalised polystyrene or silica) and nanoparticles (any of polystyrene, silica or gold). The microspheres and nanoparticles segregate into the water droplets, which are then slowly evaporated. As the droplets evaporate the microspheres and nanoparticles come together to eventually form a single composite particle in each droplet. The larger microspheres tend to adopt highly symmetric configurations, while the nanoparticles fill in the gaps to create relatively smooth particles, as shown in Fig. 2.6. The technique naturally produces a wide range of cluster sizes, which can be separated by ultracentrifugation in a density gradient. If the particles are differently functionalised the so-called "noses" of the protruding microspheres could act as sticky patches. Many variations have been produced, such as using a soluble polymer instead of nanoparticles to fill the gaps between the microspheres,

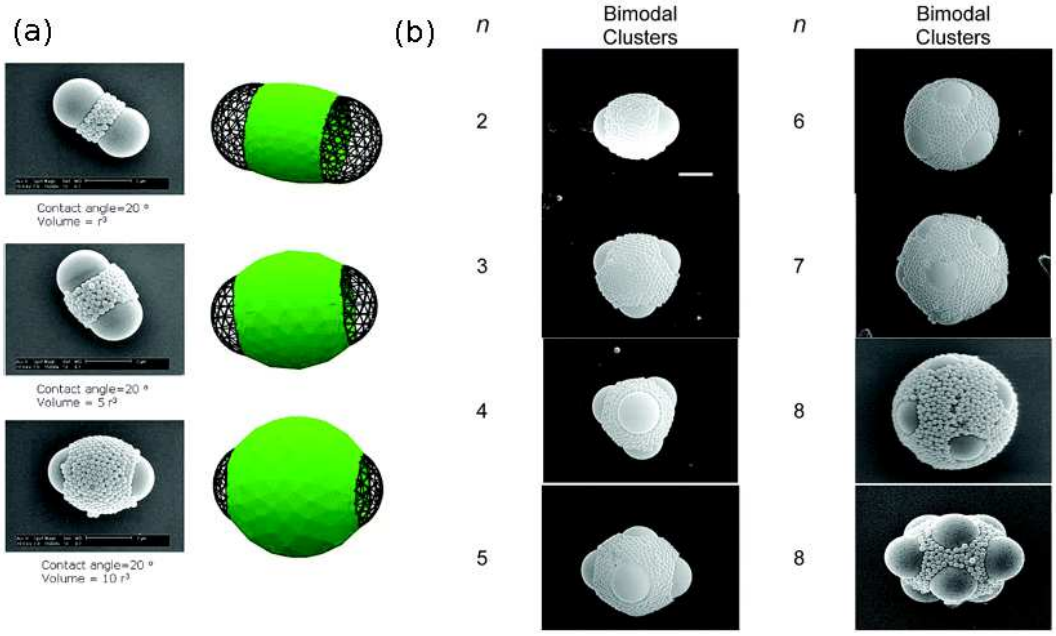

FIGURE 2.6: Clusters formed from mixtures of silica microspheres and nanospheres in a water-in-oil emulsion. (A) SEM images and simulated structures of bimodal clusters for $n = 2$, where n is the number of microspheres in the cluster, and different numbers of nanospheres. (B) SEM images for clusters with larger values of n. The scale bar is $2\,\mu$m. Reproduced from Ref. [102].

or etching away the microspheres to leave behind cagelike nanoparticle structures. The major disadvantages of this approach are the labour-intensity of the separation process and the relatively small quantities of clusters produced. The functionalisation of the microsphere noses has also turned out to be surprisingly difficult.

The relative positions of a particle's neighbours in a colloidal crystal provide a natural source of anisotropy. Capitalising on this, Wang et al. produced a colloidal crystal from micrometer-sized polystyrene spheres, annealed it to form connections between adjacent particles, and treated the crystal with sulphuric acid to allow sulphonation on the exposed surfaces of the particles[103]. The crystal was then washed and broken up, after which the particles were used as seeds in a polymerisation process, causing polymer to grow from the sulphonated regions of the surface. A similar process was used with silica spheres, with extra silica growing only on those regions of the surface which had not been functionalised. A summary of the process and scanning electron microscope images of example particles are shown in Fig. 2.7.

An alternative method making use of the relative positions of particles in colloidal crystals has been demonstrated by Zhang et al.[86]. They arranged polystyrene microspheres into colloidal crystals consisting of four or five layers on a flat substrate and then deposited gold atoms from above, with the intention of using the first and second layers as masks for the third layer, which they intended to pattern. In fcc crystals a single gold spot was produced in the third layer via the interstices in the layers above. In hcp

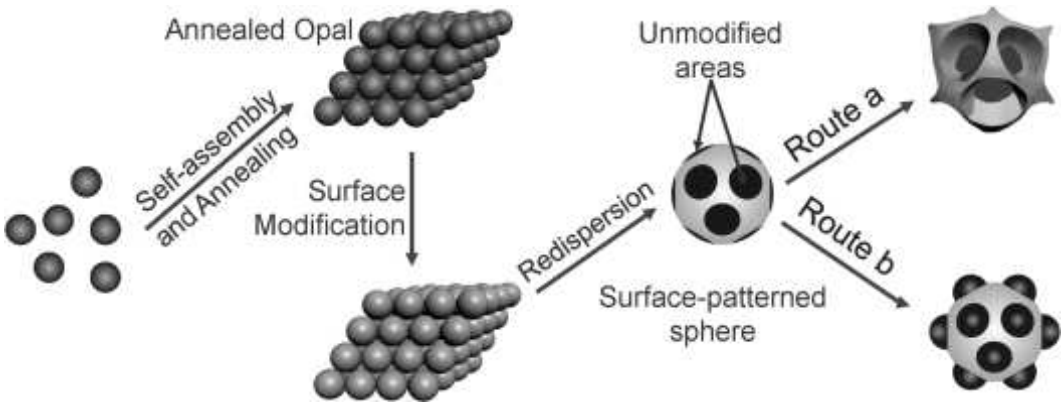

FIGURE 2.7: Patterning the surface of microspheres by making use of crystal contacts, and two routes to fabricating nonspherical particle using the surface patterned microspheres. Route a: Production of particles with protruding edges by seeded polymerisation which occurs mainly on the modified areas of the surface. Route b: Production of particles with protruding nodules by growth on the unmodified areas of the surface. Reproduced from Ref. [103].

crystals no gold was able to reach the third layer, until etching with O_2 plasma was used to reduce the size of the colloids in the top layers. Following etching, gold deposition produced three spots on the third layer in hcp, or four in an fcc crystal oriented with a (100) plane parallel to the substrates. Bizarrely, in all cases an extra gold spot was found on the underside of the particles; as yet this remains unexplained. In any case these asymmetric patchy particles are very attractive as starting points for self assembly, although they suffer from a number of problems. For example, they are hard to produce in large quantities, control over patch positioning is limited, and as of publication the gold spots were found to immediately detach when the colloids were suspended in water.

DeVries et al. used an entirely different symmetry-breaking approach to produce divalent spheres[75]. They coated gold nanoparticles in thiolated molecules. On a flat surface these molecules would tend to rest at an equilbrium tilt angle so as to maximise their intermolecular interactions. On the surface of the gold spheres they similarly tend to adopt a tilted configuration; however according to the "hairy ball" theorem it is impossible to form a 2D crystal on the surface of a sphere without producing two defects, similar to a crown in human hair. These defects tend to form directly opposite each other. The group found that these points were susceptible to exchange reactions with linker molecules, forming divalent nanoparticles which could then be assembled into chains. The theoretical topology of particles on a spherical surface has been discussed in more detail by Nelson, who highlights the fact that in a system where the particles are rod-like but lie flat on the surface (i.e. both ends of the rods are equivalent) the system will tend to form four defects, arranged in a tetrahedral configuration. If this can be exploited experimentally it may offer a very promising route to producing tetravalent particles which may then be used to synthesise colloidal diamond.

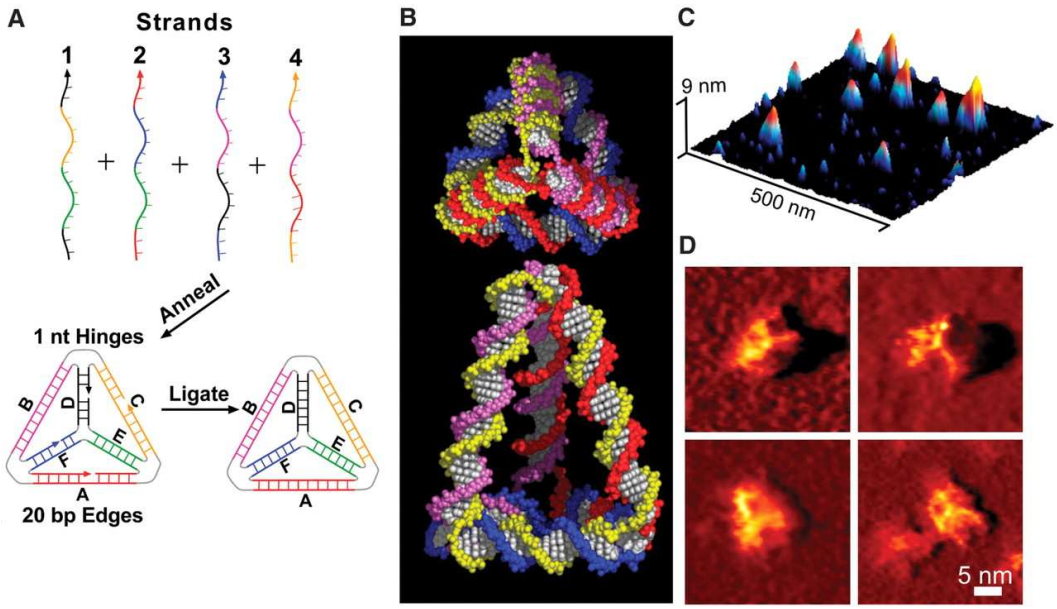

FIGURE 2.8: DNA tetrahedra. (A) The design of the DNA strands used to assemble DNA tetrahedra by annealing. Complementary sequences are given the same colour. (B) Two space-filling views of a tetrahedron. The tetrahedron shown here has three 20 base pair edges, and three 30 base pair edges. (C) AFM image showing several tetrahedra on a mica surface. (D) AFM images recorded with ultrasharp tips, showing four tetrahedra. The upper three edges are resolved. Reproduced from Ref. [106].

As a final note in this section, it is important to discuss recent work on the construction of self-assembling structures from DNA. Although the DNA from which these structures are constructed is a very different class of material from the colloids and nanoparticles discussed above, this work represents the most successful experimental progress thus far in the self-assembly of nano-scale monodisperse structures. In addition, although physically very different from the model of spherical patchy particles to be presented in this thesis, it is likely that the behaviour of such DNA systems is closely related to that of our models.

Polyhedra have been assembled from DNA since the early 1990s[104, 105]. However, the procedure for doing so was very long and arduous (synthesis of truncated octahedra took approximately two worker-years) and yields were around $1 in Oxford published results from a new method$ Fig. 2.8(a) shows how the tetrahedra are designed, with each tetrahedron built from four complementary strands, while Fig. 2.8(b-d) show idealised and experimental images of the structures. The base sequences were chosen to promote hierarchical assembly, such that two opposite edges would form at a high temperature, and then be bound together by weaker interactions along the other edges as the temperature fell further. This was believed to be an important factor in the success of the assembly. However, more recently the group has come to believe that this hierarchicity is not important, leading to some uncertainty as to the origin of the robust nature of the process.

The DNA tetrahedra have a number of potentially useful properties. Compression with

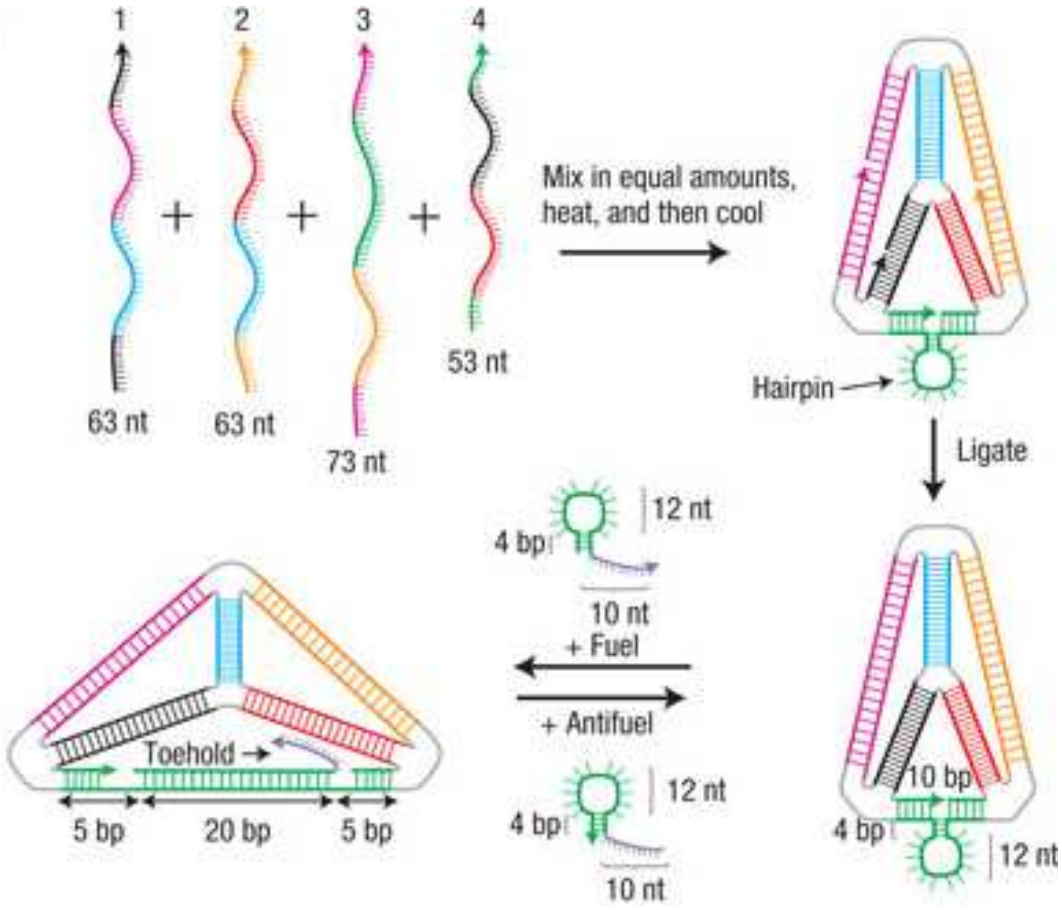

FIGURE 2.9: Scheme for the synthesis of a tetrahedron with a configurable edge. This edge is only 10 base pairs (bp) long, compared to 20 bp for the other edges, and includes a hairpin. On the addition of a "fuel" DNA strand, itself a hairpin, the fuel can hybridize with the hairpin in the short edge, allowing the edge to extend to a length of 30 bp. An antifuel hairpin may then be added which hybridises first to the fuel's single-stranded "toehold" region and proceeds to hybridise entirely with the fuel, returning the tetrahedron to its original state. Adapted from Ref. [107].

an atomic force microscope shows that they are physically strong, which may make them useful as building blocks in constructing nanoscale structures. The nature of DNA allows for many clever modifications to be made to promote certain behaviours. For example, if the design is such that a short section of the tetrahedron is left single-stranded, linker sections of DNA can be added to solution to bind together pairs of tetrahedra. Hairpins can be included in some of the edges which extend to their full lengths when interacting with added lengths of DNA (termed "fuel" DNA)[107], such that the size of the tetrahedra can be increased as shown in Fig. 2.9. "Anti-fuel" may then be added to pair with the fuel DNA and shorten the edges again. Cargo may also be encapsulated in the tetrahedron[108], preventing interaction with other large molecules such as enzymes until release and suggesting medical applications.

In 2008 a different approach led to the synthesis of complex DNA polyhedra as shown in Fig. 2.10, certainly the most impressive example of synthetic monodisperse self-assembly

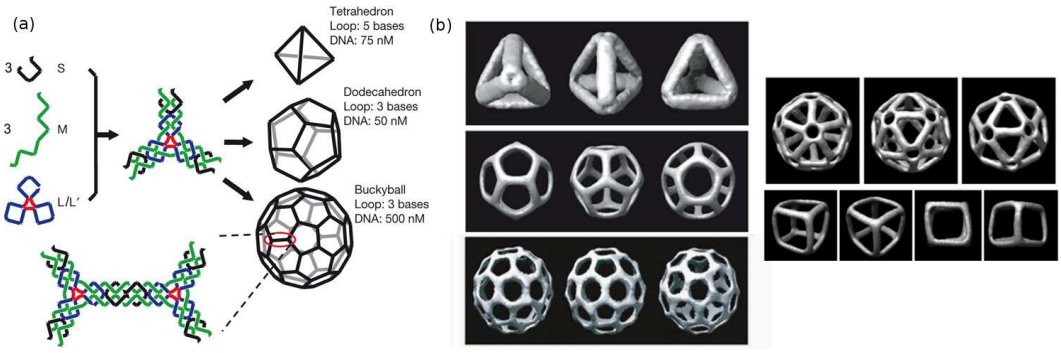

FIGURE 2.10: (a) Scheme for the synthesis of tetrahedra, dodecahedra and buckyballs. Three different sequences of single-stranded DNA assemble into symmetry three-point-star motifs and then into polyhedra in a one-pot process. By varying the concentration and the length of the flexible single stranded loop, coloured red, different polyhedra are obtained. (b) Projections of the completed polyhedra reconstructed from cryo-EM images. Adapted from Refs. [109–111].

thus far[109]. The polyhedra are of the order of tens of nanometres in size. The building blocks consist of three-point-star motifs with each point consisting of a pair of DNA double helices with overhanging "sticky" single-stranded ends. The blocks are constructed from seven strands with only three unique DNA sequences: a long repetitive strand which forms the central loop (blue-red and labelled L/L' in the figure), three medium strands (green, M) and three short peripheral strands (black, S). At the centre of the motif are three single stranded loops (red). The length of these loops determines the flexilibity of the motif. The major product of self-assembly of the motifs depends only on this flexibility and on the concentration of DNA. Long single-stranded loops allow the formation of tetrahedra (since the motifs need to bend a great deal to produce the required curvature). Shorter loops lead to dodecahedra or buckyball (truncated icosahedron) structures, with lower concentrations favouring the smaller structure, presumably because the slower growth of the structure gives more time for it to close via intrastructure binding. More recently two further papers have demonstrated the assembly of icosahedra[110] and cubes[111]. The assembly of icosahedra from five-pointed star motifs was attempted on the rationale that triangular faces are a helpful feature in a self-assembly target, since they increase its structural integrity, and hence assembly of icosahedra might be expected to be more efficient than that of dodecahedra or buckyballs. The cubes were assembled by a slightly different scheme, in which two different types of three-pointed star motifs were used. Each type was able to bond to the other, but not to motifs similar to itself. This forced each face to include an even number of vertices, leading to good yields of cubes by excluding tetrahedra, octahedra and buckyballs as computing structures. Fig. 2.10(b) shows projections of all the completed structures reconstructed from cryo-EM images. While yields of the largest of the structures, the buckyball, are not high, there is no doubt that these structures represent an astonishing feat of self-assembly, especially given the very limited set of experimental variables which needed to be varied by the researchers in order to control the structure of the product.

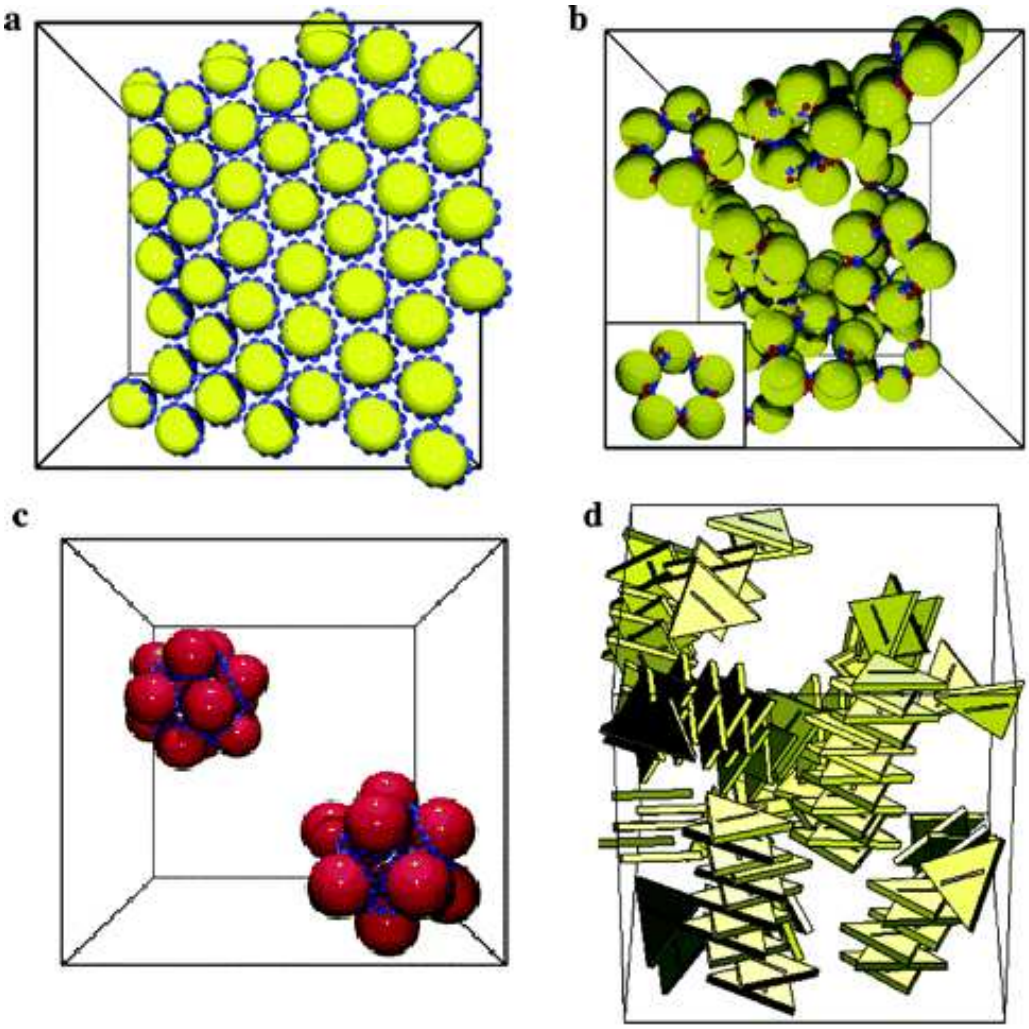

FIGURE 2.11: Structures formed following simulation of patchy particles. While the particles consist of rigid sets of soft spheres, the images show the overall shape of the particles, with only the attractive subparticles individually visible. (a) A sheet formed by spheres with a ring of attractive sites. (b) Pentameric rings formed by particles with two attractive regions, each able to bond only to the other type (and not to another like itself). (c) Icosahedra formed from particles with rings of attractive sites. (d) Columns of triangles formed from individual triangular particles with attractive lines with different orientations on top and bottom. Adapted from Ref. [112].

2.5.3 A computational study

There have been very few computational studies of the self-assembly of patchy particles. No discussion of patchy particle assembly would be complete, however, without referencing work by Zhang and Glotzer briefly surveying some of the structures which might be assembled from patchy particles[112]. A number of different particles were designed, each consisting of a rigid set of soft spheres, some of which were able to form attractive interactions, and systems of particles were simulated using Brownian dynamics. Fig. 2.11 shows a representative selection of the structures formed. Even this rather cursory study gives an idea of the considerable range of structures that may be possible.

2.6 Experimental and simulation progress on self-assembly

2.6.1 Experimental progress

There have been a large number of experimental studies on the assembly of icosahedral capsids, e.g. [11–13, 57, 113**?** –125]. However, since the issue is a complicated one involving many complex interactions, most of these studies have tended to focus on particular aspects of assembly, for example interactions with nucleic acids[116, 124**?** , 125], cooperative binding between subunits[115], or the effects of scaffolding proteins[118, 122]. A relatively small number of studies[11–13, 57, 113] consider the assembly process as a whole in order to examine the dynamics of assembly. These studies make use of light scattering[11–13, 113], electron microscopy[11, 13, 57] or size exclusion chromatography[11, 12] to track the progress of assembly with time, under various reaction conditions.

Such studies collectively reach a number of conclusions, with each conclusion observed by multiple studies. Firstly, the yield of complete capsids is found to be a sigmoidal function of time, with an initial lag period during which no capsids are produced, followed by a period of rapid completion before the yield reaches a plateau. Hysteresis in capsid formation is found, i.e. capsids remain stable in conditions under which they would not form. The assembly process is generally found to begin with the formation of a nucleus, for example a subunit pentamer. Under fast growth conditions such as high subunit concentrations, assembly is seen to be vulnerable to kinetic traps, in which an excessive number of nuclei are formed. Each of these nuclei grow, consuming subunit monomers in the process and leading to a rapid depletion of the monomer population ("monomer starvation") and the eventual production of a large number of part-formed capsids, with few reaching completion. The only way such a trap may be escaped is by breakup of some clusters to provide monomers for the completion of others, which in many cases may be such a slow process as to render the trapping essentially permanent. This kinetic trapping may be avoided by slow nucleation, or escaped by highly reversible capsid growth. Reversible capsid growth is also found to reduce the occurrence of misformed capsids, since non-native interactions can be broken up. Consequently slow nucleation and reversible growth both appear to be crucial features of successful assembly. Several of these conclusions are only reached by comparison with the predictions of kinetic models by Zlotnick and coworkers[12, 126**?**], which will be covered in detail in Section 2.6.2.1.

Many capsids are found to be formed by subunits which are pre-grouped (or group rapidly) into capsomers, e.g. pentamers or dimers. For example, the $T = 7$ bacteriophage HK97 is formed from pentamers and hexamers[**?**], human papillomavirus (HPV) is formed from pentamers (HPV is unusual in that despite possessing a $T = 7$ structure it contains no hexamers; pentamers occupy sites with six neighbours), and hepatitis

B virus (HBV)[12], cowpea chlorotic mottle virus (CCMV)[11] and turnip crinkle virus (TVC)[57] all form from dimers. Consequently, assembly nuclei tend to consist of groups of these capsomers. For example, the nuclei for assembly of CCMV[11], HPV[13] and HBV[12] are thought to be a pentamer of dimers, a dimer of pentamers and a trimer of dimers respectively.

At present, findings of the kind described here appear to be at the limit of what is experimentally accessible with regard to overall assembly mechanisms (although of course constant and considerable progress is being made in elucidating the details of interactions between capsomers, nucleic acids and scaffolding proteins, processes of capsid maturation and so forth). The difficulty is that the intermediate species of assembly are numerous and small in number at any one time (populations tend to be dominated by monomers and complete capsids, even during growth). Given the extraordinary range of possible intermediates and assembly pathways, the observation of a small number does not shed much light on the process. For this reason computational models provide a valuable tool in extending our understanding.

2.6.2 Theoretical and simulation studies of monodisperse self-assembly

Because of the extreme difficulty in studying the details of virus assembly by experiment (and the virtual absence of extant synthetic systems self-assembling into monodisperse structures), along with the great potential benefits to medicine and materials science if progress is made, the field of monodisperse self-assembly represents an attractive subject for simulation. Viruses provide the classic model system, both because of the many applications of understanding them better and because they provide the best example systems known, combining high-fidelity assembly of large structures with, at least in some cases, an approachable degree of simplicity.

As desribed in Section 2.4, there are a large number of complicating factors than may be involved in virus assembly and are beyond current modelling ability, and so the best approach at present is to restrict simulation for the most part to the simplest examples, i.e. the *in vitro* assembly of capsids from purified capsid protein. Even then there are many difficulties in simulating virus assembly, which lead to a need for sacrifices of one kind of another in either the detail of the model or the timescales simulated. We shall briefly consider these issues here.

The first and most immediate problem is the difficulty of representing the virus in a model in the first place. Results have recently been published for the first all-atom simulation of a virus[127]. Although a small $T = 1$ virus was chosen, more than 1 million atoms had to be simulated. Consequently the simulation was extremely expensive (128 "Altix" nodes progressed the simulation by around 0.7 ns per day), simulated only a single capsid and covered a timescale of only around 50 nanoseconds. Since assembly

typically takes place over timescales of seconds to days, clearly this approach is not viable for the study of assembly, although it may ultimately have some application in studying the stability of capsids.

A practical approach to improving the situation is to attempt to coarse-grain the simulation, such that each model particle represents a number of atoms. This has been done in further work by the same group[128], in which a specialised algorithm is used to group sections of the molecule to be represented by single particles, such that the key features of the protein structures are retained. An average of around 200 atoms are represented by a single coarse-grained particle. This coarse-graining allows an extension of the simulation to capsids of up to around ten times the diameter, and to timescales in the microsecond range. While this is a great improvement, it still falls far short of what will be required for assembly simulations. Consequently much more dramatic coarse-graining is required, such as replacing the capsid proteins with simple rigid structures (e.g. work by Rapaport[129, 130] and Nguyen et al.[131, 132]) or even single spherical particles (e.g. Hagan et al.[133] and our own work[134**? ?**]).

A further difficulty arises from the need to obtain good quality statistics. The assembly process can potentially proceed via any of a very large number of intermediate states, leading to an extraordinary number of possible assembly routes, especially if the possibility of reversible subunit binding is taken into account. As such, data on the assembly pathways of a small number of model capsids may not be representative and care should be taken not to overestimate the significance of individual results. The way to overcome this problem is to obtain sufficiently large quantities of data that general trends in the assembly mechanisms become apparent. This requirement further greatly increases the computational expense.

A recurring issue in developing computational models of assembly is the lack of experimental data to check results against. Because of the great experimental difficulties in making statistically significant observations of transient intermediates, little information is available about assembly mechanisms (which is, of course, one of the reasons that computational models promise to play an important role). Much of the large amount of detailed information that experimentalists have amassed about capsid assembly may not be applicable because it applies on a level of detail which has been lost in the process of coarse-graining or abstracting the problem (e.g. information about protein binding sites, protein folding and conformations etc.). As a result the only available checks for model systems are that they reproduce the key observable features of assembly, for example yields which are sigmoidal functions of time with initial lag periods, kinetic traps at high concentrations, and so forth.

Because of the computational issues involved it is necessary to make many assumptions and simplifications in constructing a tractable model of assembly. Different researchers have approached this problem in different ways. There have in fact been surprisingly

few serious computational studies studies of virus assembly, with as few as six research groups making major contributions to the field. As a result, and since their results will be directly relevant to our own work, we will consider them in some detail. We divide the simulations into three groups: kinetic models, shaped particle models and single particle models.

2.6.2.1 Kinetic models

Kinetic models of assembly do not store information about the spatial locations of the species involved. Rather, the populations of each species are recorded, and the flux between the populations is determined by the use of kinetic rate equations. An initial decision therefore arises for the model designer in finding a way to determine (or guess) the rate constants of the various combination and dissociation reactions. A more serious challenge, though, is provided by the extraordinary number of potential intermediate species on the path to the assembly of a complete virus, and the even larger number of possible combination reactions between pairs of them. Thus far researchers have been forced to limit the sets of possible reactions and intermediates, generally allowing only the addition of single capsid proteins or capsomers, and only in the position that will create the most stable structure. These simulations have provided valuable results, but the question has remained whether these restrictions are in fact reasonable. More recent simulations have attempted to tackle this question.

A majority of the important kinetic simulations of virus assembly to date have been carried out by Zlotnick and coworkers. Over the course of three papers[12, 126**?**] (and eight years) the group developed a kinetic model which describes assembly by a series of rate equations. It is assumed that only the most stable species of each size is relevant to assembly, and that growth occurs only by the addition of single subunits. The time-evolution of the system is governed straightforwardly by the equation

$$\frac{d[m]}{dt} = f_m s_m[1][m-1] - f_{m+1} s_{m+1}[1][m] + b_{m+1}\sigma_{m+1}[m+1] - b_m \sigma_m[m] \qquad (2.1)$$

, where $[m]$ is the concentration of the species containing m subunits, f_m and b_m are the forward and backward rate constants respectively for the formation and dissociation of the species with m subunits, and s_m and σ_m are statistical factors giving the number of equivalent ways of forming/dissociating the species. The reactions are assumed to be diffusion limited, and f_m is in the first instance set to be equal in all cases. b_m is then set such that the species have the correct equilbrium constants, which are determined by assuming that the free energy change on adding each subunit is simply proportional to the number of new subunit-subunit contacts formed. Fig. 2.12 shows the progression of species for the assembly of a 12-subunit dodecahedron.

Assembly intermediates and factors describing their assembly

n	Model	Build Up	Build Down	S_n	c	K'_n
1		—	—	—	—	—
2		5/2†	1	5/2	1	$e^{-1\Delta G^\circ_c/RT}$
3		2	3	2/3	2	$e^{-2\Delta G^\circ_c/RT}$
4		3	2	3/2	2	$e^{-2\Delta G^\circ_c/RT}$
5		4	2	4/2	2	$e^{-2\Delta G^\circ_c/RT}$
6		1	5	1/5	3	$e^{-3\Delta G^\circ_c/RT}$

n	Model	Build Up	Build Down	S_n	c	K'_n
7		5	1	5/1	2	$e^{-2\Delta G^\circ_c/RT}$
8		2	4	2/4	3	$e^{-3\Delta G^\circ_c/RT}$
9		2	3	2/3	3	$e^{-3\Delta G^\circ_c/RT}$
10		3	2	3/2	3	$e^{-3\Delta G^\circ_c/RT}$
11		2	5	2/5	4	$e^{-4\Delta G^\circ_c/RT}$
12		1	12	1/12	5	$e^{-5\Delta G^\circ_c/RT}$

FIGURE 2.12: The intermediate species in the assembly of a dodecahedron. The "Build Up" and "Build Down" parameters correspond to s_m and σ_m in Equation 2.1. c gives the number of new contacts created on forming the species by addition of a single subunit to a smaller species, and determines the value of the equilbrium constant K'_n. S_n is not used in the formalism as presented here. †: The unintuitive factor of 1/2 is due to the degeneracy of the dimer. Reproduced from Ref. [**?**].

Fig. 2.14(a) shows a typical trajectory for the model described above. The population of monomers is rapidly depleted as a "production line" of intermediates is set up. There is a short lag before the first complete capsids appear, during which time smaller intermediates are assembled and progressively consumed in the assembly of larger species. A high yield of complete capsid is rapidly produced, which then only gradually approaches the equilbrium yield, with the small remaining concentration of monomers leading to ever slower assembly.

A number of notable points arise from such simulations. Firstly, a lag phase is always observed, whereby no complete capsids are formed at very short times. This used to be considered diagnostic of nucleated processes[135], but in fact has been shown to result from any multistep reaction[136]. Secondly, for successful assembly processes such as that shown in Fig. 2.14(a), only the monomer and complete capsids are observed at appreciable concentrations. The intermediates are present only in trace quantities, and indeed might often not be observable in experiments. Thirdly, the thermodynamics of such a multicomponent reaction is such that, to a first approximation, [capsid] $\propto$ [subunit]^N, where N is the (generally large) number of subunits in the complete capsid. This very strong dependence leads to a sharp transition between the monomer-dominated and capsid-dominated states, such that experimentally the system will tend to appear as if there is a critical concentration, analogous to the critical micelle concentration, above

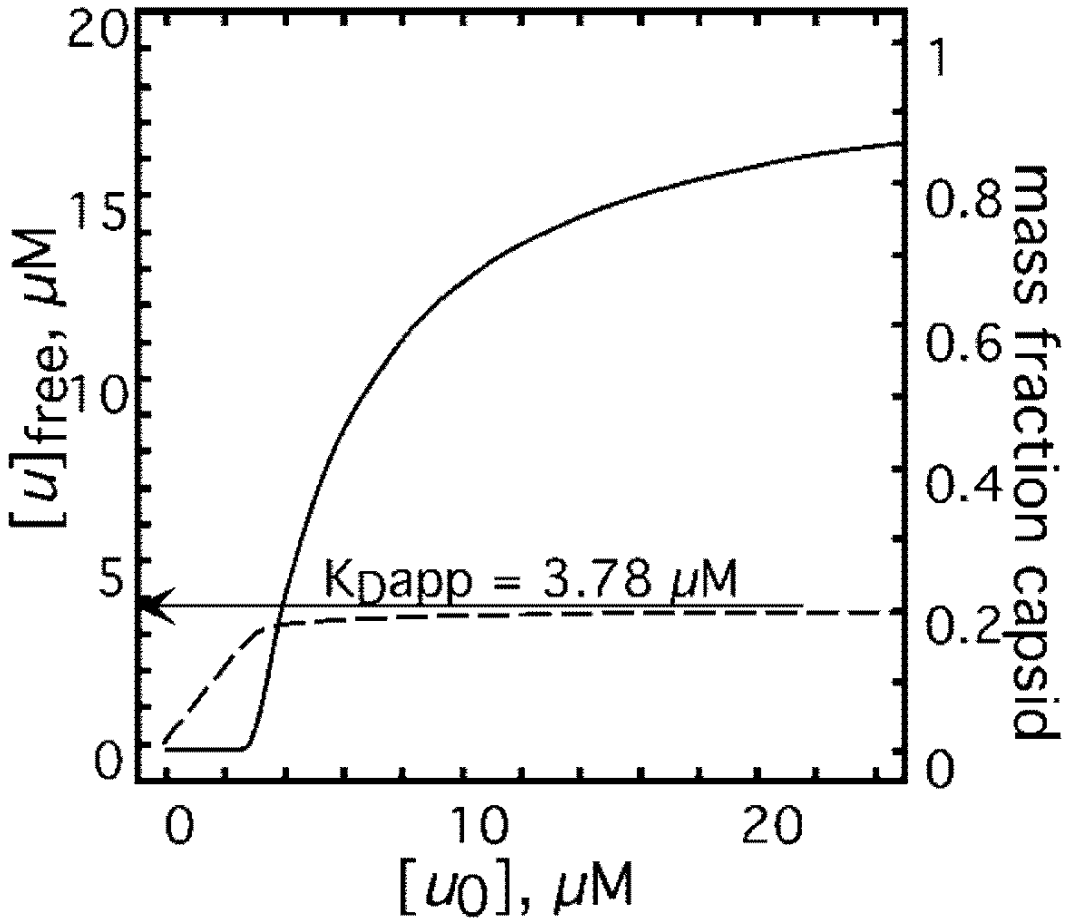

FIGURE 2.13: The mass fraction capsid (solid line, right axis) as a function of the initial subunit concentration, $[u_0]$. Free subunit concentration (dashed line, left axis) varies little when it approaches K_{Dapp}, which is defined as the concentration at which $[u]_{free}$ = [capsid], but it is not constant and will eventually cross K_{Dapp}, i.e. there is no true critical concentration. Reproduced from Ref. **??**.

which any further mass will be added in the form of complete capsids. Fig. 2.13 shows a typical plot for a 30-subunit system.

Fig. 2.14(b) shows the results of a simulation similar to that of Fig. 2.14(a), but with the association constant for the binding of subunits increased by a factor of ten. The results here showcase a kinetic trap which is a recurring theme of capsid assembly. The monomer population is reduced very quickly producing a large population of intermediates. With very few monomers remaining, these intermediates are unable to grow further, and essentially no complete capsids are produced. This phenomenon is referred to elsewhere in the literature as monomer starvation.

Zlotnick et al. suggested the idea that assembly might benefit from being "kinetically limited" by slow nucleation steps in the early stages of assembly. The simulation conditions for Fig. 2.14(c) match those of Fig. 2.14(b), but the forward rate constant of the first two reactions is decreased by a factor of one hundred. Consequently the depletion of

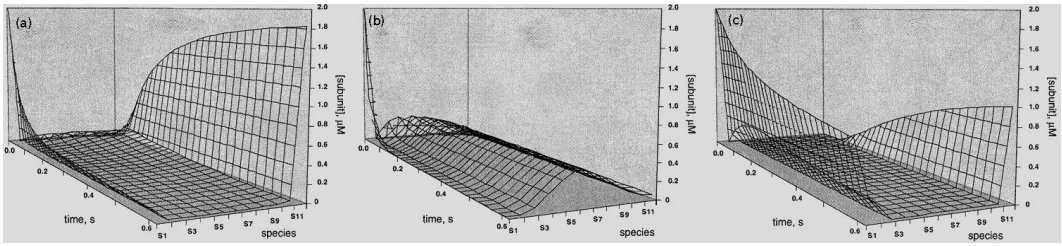

FIGURE 2.14: The concentration of each species, monomer to dodecahedron, for simulations by Zlotnick et al.. (a) Successful assembly. (b) Assembly is stalled by increasing the association constant by a factor of ten as a result of kinetic traps. (c) A reduction in the rate of the assembly steps up to trimer formation under conditions matching those for (b) reduces the nucleation rate and avoids kinetic traps. Adapted from Ref. [12].

the monomer population is greatly retarded. When nuclei are successfully formed they rapidly go on to produce complete capsids, and there is no kinetic trapping. However, the inclusion of nucleation steps does lead to relatively slow assembly.

Zlotnick et al., provide methods for the extraction of the nucleation rate, growth rate, nucleus size and contact free energy from straightforward experimental data, assuming that assembly follows their model. The reliability of any figures obtained using these methods is not clear, given the numerous assumptions in the model. However, their theoretical framework appears to be virtually the only one available for experimentalists at present. The group have themselves used their methods in analysing the results of a number of experimental studies, with apparent success[12**?**].

The key predictions highlighted by Zlotnick et al.[126] are as follows: 1) The inclusion of nucleation steps minimises kinetic trapping. 2) Weak association energy is sufficient for formation of stable capsids[137**?**]. Weak interactions have the additional advantage that capsid assembly and stability are more susceptible to regulation, i.e. small modifications will prevent or reverse assembly. 3) Dissociation of capsids will display a marked hysteresis (as has been observed experimentally[**?** **?**]. This is a result of a kinetic barrier to dissociation: if a complete capsid begins to dissociate, the nearly-complete capsid will be strongly driven to reform a complete capsid rather than dissociate further. This hysteresis is in fact crucial to viral survival - since a single virus particle can in some cases infect a cell, it is necessary that capsids are stable down to infinite dilution.

A number of criticisms of the model are possible. The restriction of the assembly pathway to include only the stablest intermediate of each size, and only monomer addition reactions rather than the combination of intermediates, is an obvious concern. Setting all of the forward reaction rates to be equal (except for nucleation reactions) will obviously introduce some error - even in the case where reaction rates are purely diffusion-limited, the rates would differ because of differing diffusion rates. One of the most serious errors may be the complete exclusion of all misformed clusters. Only one class of kinetic trap is possible - that of monomer starvation - while others such as the formation of oversized

aggregates or simply the misalignment of parts of the capsid are excluded. Finally, the assumption that each contact has an equal free energy of formation is certainly quite inaccurate. Apart from complex phenomena such as cooperative binding, the addition of each subunit has a fixed entropy cost due to the loss of the degrees of freedom of the subunit. As a result, additions creating two contacts will be significantly more than twice as favourable as those creating only one contact.

Zlotnick and coworkers have themselves tackled the question of whether their restriction of the possible set of intermediates to the most stable set is reasonable[138]. For the assembly of a 12-subunit dodecahedron, they found that including 21 of the 73 possible states provided results very close to those obtained from using all 73. However, including only the 12 states shown in Fig. 2.12 led to fairly significant errors. It would seem that at least some diversity of assembly pathways must be included to obtain good statistics.

Zhang and Schwartz performed similar simulations, but lifted the restriction that growth could occur only by monomer addition in order to study the importance of oligomer-oligomer binding[139]. The results were found to be very similar in cases where capsid growth was slow and nucleation-limited, but diverged sharply under conditions where kinetic trapping would generally occur. The systems were able to escape kinetic traps to some extent by combining part-formed capsids, alleviating the monomer starvation issue. Unfortunately the study suffered from a violation of detailed balance - once clusters had formed closed loops they could no longer be broken apart. While the authors made the valid claim that this assumption could be justified under certain experimental conditions, they then extended their analysis to conditions where the assumption was not valid, leading them to the conclusion that the long-time success of assembly improved monotonically with temperaure.

More recently Misra et al.[**?**] (also in the Schwartz group) carried out another study, again to examine the importance of oligomer-oligomer binding. They first studied the assembly of a 12-subunit dodecahedron by a differential equation method similar to that of Zlotnick and coworkers, and then studied assembly of a much larger 60-particle $T = 1$ capsid using a new discrete-event method. They found that over large regions of parameter space the addition of oligomers made an important or even dominant contribution to assembly. For example, in the $T = 1$ model, if the intra-pentamer bonds were stronger than the inter-pentamer bonds, the system would tend to form pentamers which then cluster to form complete capsids. A reasonable argument would be that in that case one could simply use Zlotnick's 12-particle method, and consider each particle to represent a pentamer. However, Misra et al., find that over large regions of parameter space a number of different growth mechanisms are all important, making simplification difficult. Unfortunately there is a lack of experimental data to indicate in what region of parameter space a typical virus assembly (if that can be considered a meaningful concept) would lie.

In summary, kinetic models like those of Zlotnick and coworkers provide a valuable framework for considering virus assembly, and have yielded useful tools for the the interpretation of experimental data. However, they are limited by the need to make a number of sometimes questionable assumptions. This will prove to be a recurring theme in capsid assembly simulations - as discussed earlier, the problem is a difficult one and rather draconian approximations are necessary. More information can be gained, however, by considering the problem from multiple angles, which leads us on to an alternative approach to simulation, in which spatial coordinates for the assembling subunits are tracked. Simulations taking this approach allow consideration of such issues as the formation of misformed capsids or even aggregates, the importance of entropy in assembly and so forth. However, the greater complexity of the models tends to lead to far inferior statistics.

We divide spatial simulations of assembly into two classes, in which the model particles are either shaped or spherical. We consider models of shaped particles first.

2.6.2.2 Shaped particle models

Rapaport et al. were the first to study the dynamics of capsid assembly using simulations with spatial information, assembling icosahedral clusters consisting of 60 triangular particles in molecular dynamics (MD) simulations, in which the particles follow Newton's laws of motion[140]. However, the model included a number of unphysical features in order to promote successful assembly: monomers were only attracted to oligomers which were perfectly formed, so that faulty intermediates could not grow further; oligomers were not attracted to each other; and correctly formed bonds were permanent. Further, particles moved under ballistic motion, including no Brownian motion or drag due to solvent, which of course is very inaccurate for such small particles. The set of rules applied to the interactions made successful assembly essentially inevitable, but the work was nevertheless a useful proof of principle. In a later paper[140] some of the unphysical moves were removed, but a mechanism was added whereby partially completed capsids were eventually arbitrarily broken up to prevent kinetic trapping, again compromising the physical relevance of the simulation.

However, in his most recent work on assembly, Rapaport presents a far more satisfactory model, in which icosahedra are assembled from 20 triangular particles[130]. The subunits are represented by rigid bodies constructed of soft spheres (shown in Fig. 2.15(a)), with multiple interaction sites positioned to stabilise the target structure. The interaction sites are only able to bind to matching sites on adjacent subunits, which limits the stability of misformed structures. Unphysical interaction rules have been removed, and importantly, explicit solvent particles have been included, modelled as individual soft spheres of the same sort that constitute the subunits. The solvent particles are

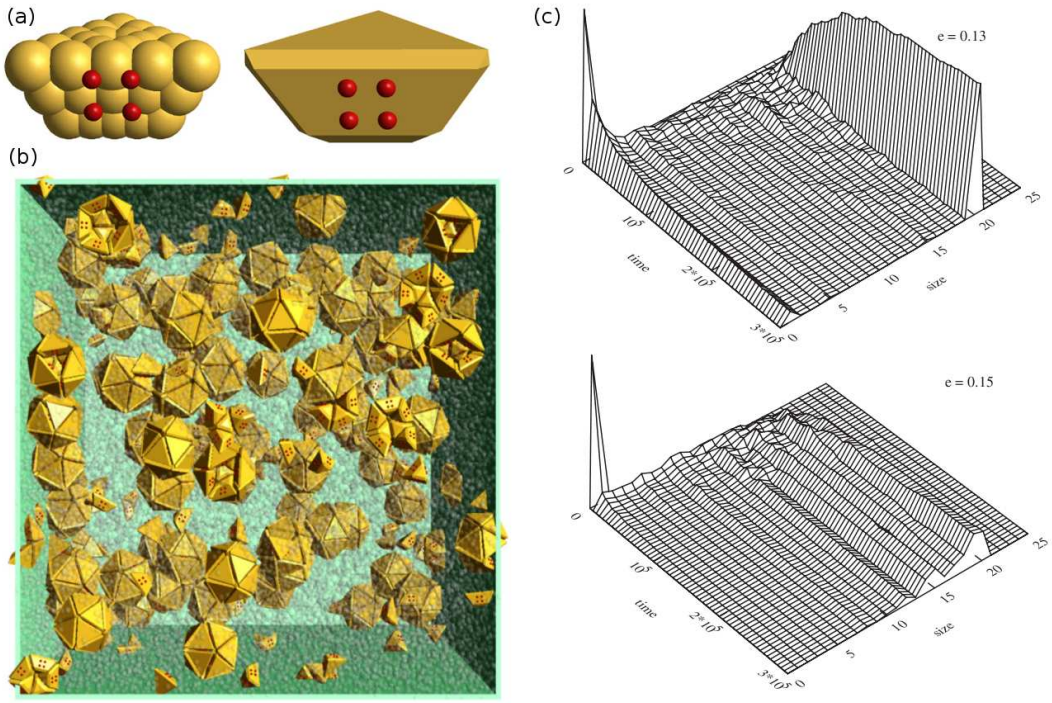

FIGURE 2.15: (a) The model particle used by Rapaport[130] and its effective shape. Large spheres have soft repulsive potentials, while small spheres denote attractive interaction sites. (b) A snapshot of the $e = 0.13$ (optimal) system with 80 complete shells. Solvent particles are shown semi-transparently, and there are visual artifacts due to periodic boundary conditions. (c) Cluster size distributions as functions of time. The distributions are expressed as mass fractions. The $e = 0.13$ simulation showcases optimal assembly, while the $e = 0.15$ simulation suffers from severe kinetic trapping. Adapted from Ref. [130].

far larger than water molecules, and represent a coarse-graining of the solvent, but successfully perform a number of functions: they aid cluster breakup without the need for collisions, they curtail the ballistic nature of particle motion and introduce Brownian motion, and they serve as a heat bath to absorb energy released by bond formation. Fig. 2.15(b) shows a typical simulation box for this system under optimal assembly conditions, displaying a high degree of successful assembly (a maximum yield of 89% was achieved).

Rapaport finds that successful assembly proceeds via a cascade of reversible stages, with a strong preference for the lowest energy intermediate states. This agrees with Zlotnick's findings above[138], indicating that only a relatively small set of intermediates need be considered. Rapaport highlights the importance of the reversibility of the assembly steps: reversible assembly allows for the removal of misattached subunits, the breakup of misformed clusters, and helps in the escape of kinetic traps. Fig. 2.15(c) shows two plots for the populations of clusters as a function of cluster size and time, for different values of the binding energy e. These plots are comparable to those by Zlotnick et al., shown in Fig. 2.14. At the higher value of e, where reactions cease to be significantly reversible, kinetic trapping occurs by monomer starvation and few icosahedra are formed. This

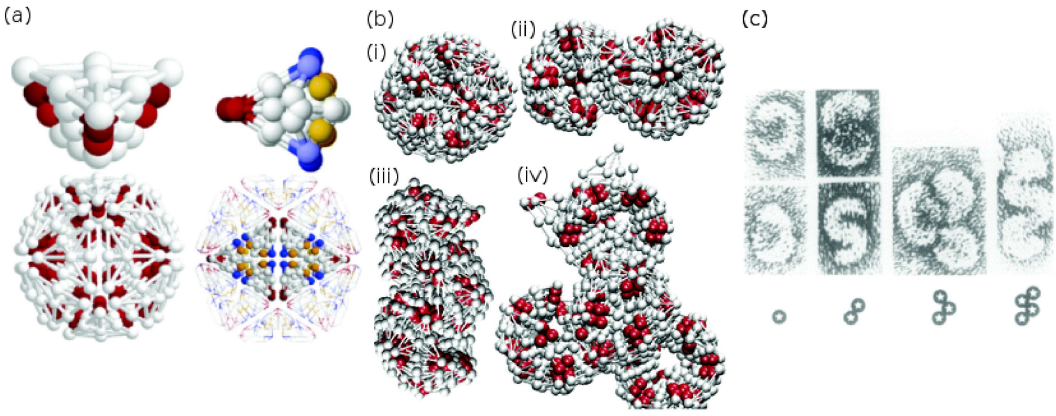

FIGURE 2.16: (a) Geometry of the coarse-grained models used in simulations by Nguyen et al.[131]. Left: a single subunit and a complete capsid for a model forming 20-subunit structures. Right: a single subunit and a complete capsid for a model forming 60-subunit structures, as with a $T = 1$ virus. Rigid bonds are shown between neighbouring psuedoatoms. Hard pseudoatoms are coloured white, while sites of attraction are coloured red, blue or yellow. (b) A gallery of oversized misassembled structures, or monster particles, comprising (i) a large closed aggregate, or two (ii), three (iii) or four (iv) open partial capsids. (c) Experimentally observed structures from electron microscopy on turnip crinkle virus. Adapted from Refs. [131] and [57].

simulation model also displays strong hysteresis, matching the predictions of Zlotnick and coworkers[141] and the observations of experiments[**?**]: when the value of of e was lowered following successful assembly, all partial capsids broke up, but complete capsids remained stable far below the minimum value at which they would assemble.

Nguyen et al. have carried out simulations using similar models[131]. They consider systems designed to form icosahedra from either 20 triangular particles or 60 smaller subunits as shown in Fig. 2.16(a), obtaining surprisingly similar results in both cases. They are able to carry out large simulations with these relatively complex particles by making use of a method called discontinuous molecular dynamics (DMD), in which potentials are required to be discontinuous - in this case either hard sphere potentials or, for interacting sites, square well potentials. The dynamics appear to be effectively ballistic, with no damping or Brownian motion other than imposed by a thermostat (although the extent of the effect of the thermostat in randomising motion is unclear and may mitigate this issue).

The simulations produced good yields of capsids, up to 80% under optimal conditions. For each concentration, there was an optimum temperature - above this temperature nucleation became slow, and capsids became less thermodynamically stable, while below this temperature kinetic trapping similar to that described by Zlotnick and Rapaport was observed. The optimum temperature increased with increasing subunit concentration, as would be expected.

Under optimal conditions the population of intermediates was very small and each individual intermediate was found to exist only very transiently, implying that the assembly was nucleation-limited with nucleated clusters growing to completion very rapidly. Further, the authors state that assembly occurs mainly by the addition of monomers, although the justification for this statement is not clear and so the strength of the evidence is hard to assess. These results all support the approximations made in the kinetic simulations by Zlotnick et al..

Unlike Rapaport's simulations, misformed clusters are reasonably stable - of the four possible interactions between adjacent subunits, two are possible if a subunit is the wrong way up (interaction sites only interact with other sites from the same layer), rather than one of four as in Rapaport's model. As a result "monster particles" are observed, containing substantially more subunits than the target structure. The monster particles tend to be either closed structures with many non-native interactions, or open structures formed from the amalgamation of a number of part-formed capsids. The particles show a resemblance to particles observed in experiments on Turnip crinkle virus[57], as shown in Fig. 2.16(b).

The authors construct an energy landscape for the assembly process, and appear to show that the formation of the complete capsid is an unfavourable step, with the complete capsid having a higher free energy than the capsid with one subunit removed. This is a very surprising result, given that the final subunit forms three contacts on addition, and further their analysis contains an error - the high free energy of the completed capsid is blamed on an unfavourable entropy, but the state is observed to become more stable at higher temperatures, which is the reverse of what would be expected. Most importantly of course, a high free energy for the complete capsid would lead to a majority of capsids possessing holes. We believe that the simulations had simply not sampled the equilibrium well.

The same group have extended this work with two very recent papers, again simulating the assembly of shaped particles using the DMD method, which pay particular attention to the possible misformed structures which can arise during assembly. In the first[132] they consider the assembly of capsids of a very wide range of sizes, $T = 1 \ldots 19$, from pentameric and hexameric capsomers, shown in Fig. 2.17(a). The capsomers are given asymmetric assembly rules based on suggestions by Kerner[**?**] which allow assembly into the correct structures. For example, in $T = 3$ capsids three of the sites in each hexamer are able to bind to pentamers, while in $T = 7$ capsids only one site can bind to pentamers. For $T > 7$ structures more than one kind of hexameric capsomer is needed. They observe and classify a number of misformed structures, some of which are shown in Fig. 2.17(b). The majority of the misformed structures contain pentamers in hexameric sites, known as hexameric dislocations. These dislocations tend to arrange themselves in an ordered manner and the structures can be classified for the most part by the number of such sites. The dislocations arise when part-formed intermediates collapse

together. As a result the rate of errors is greatly increased by conditions which lead to rapid nucleation and the simultaneous presence of multiple intermediates. Under optimal conditions high yields of icosahedral capsids are obtained for capsids of sizes $T = 1 \ldots 7$, while for $T > 7$ yields of less than 10% were obtained. This indicates that while scaffolding proteins and nucleic acids may well play an important role for capsids of $T \leq 7$, they are likely to be essential for larger capsids.

The second recent paper by the Brooks group makes use of a somewhat more detailed model to study the assembly of $T = 1$ and $T = 3$ capsids[142]. For all capsids larger than $T = 1$, the individual subunits occupy different quasiequivalent positions, and are therefore thought to assume different conformations. The rate of interchange is assumed to be much larger than the rate of assembly, with proteins possibly adopting a stable conformation only at the moment that they are added to a growing capsid. Brooks and coworkers model this, somewhat unintuitively, as a "pre-equilibrated" system by simply including the appropriate populations of each of three different conformations for the $T = 3$ structure and forbidding interchange. The three conformations are shown, along with the complete structures, in Fig. **??**(c). The gemoetries of the three conformations are slightly different, and binding rules are assigned so that interactions are only formed between appropriate faces of the subunits. The subunits are initially grouped into permanent capsomers of one, two or three subunits - however, the initial grouping is found to be unimportant, with similar results being obtained in each case. Once again misformed capsids are observed, in which multiple intermediates collapse to produce structures with hexameric dislocations, consisting in this case of six of the normally pentamer-forming "A" conformation subunits coming together to form a strained hexamer. The structures and frequencies of some erroneous structures are shown in Fig. **??**(d). Errors are much more likely in the $T = 3$ structure - the larger structure takes longer to complete, and therefore spends more time as an intermediate vulnerable to disruption by the addition of another intermediate. Once again, errors can be avoided by conditions in which nucleation is slow. Thus slow nucleation has two advantages - firstly the avoidance of monomer starvation, and secondly the prevention of the collapse of multiple intermediates to produce misformed structures.

These studies by the Brooks group collectively make a very important contribution to the field, displaying an impressive combination of a relatively detailed models and fairly good statistics with no unphysical restrictions.The DMD method must be a cause for concern, since the ballistic dynamics will be very different to those actually found for proteins in solution. It is not yet clear how important the details of the dynamics might be. The specificity of the interactions between particular capsid protein conformations in the later papers is also very much a computational convenience; in fact, the experimentally determined structures of different conformations are generally very similar, and the consequent differences in interaction energies are thought to be small.

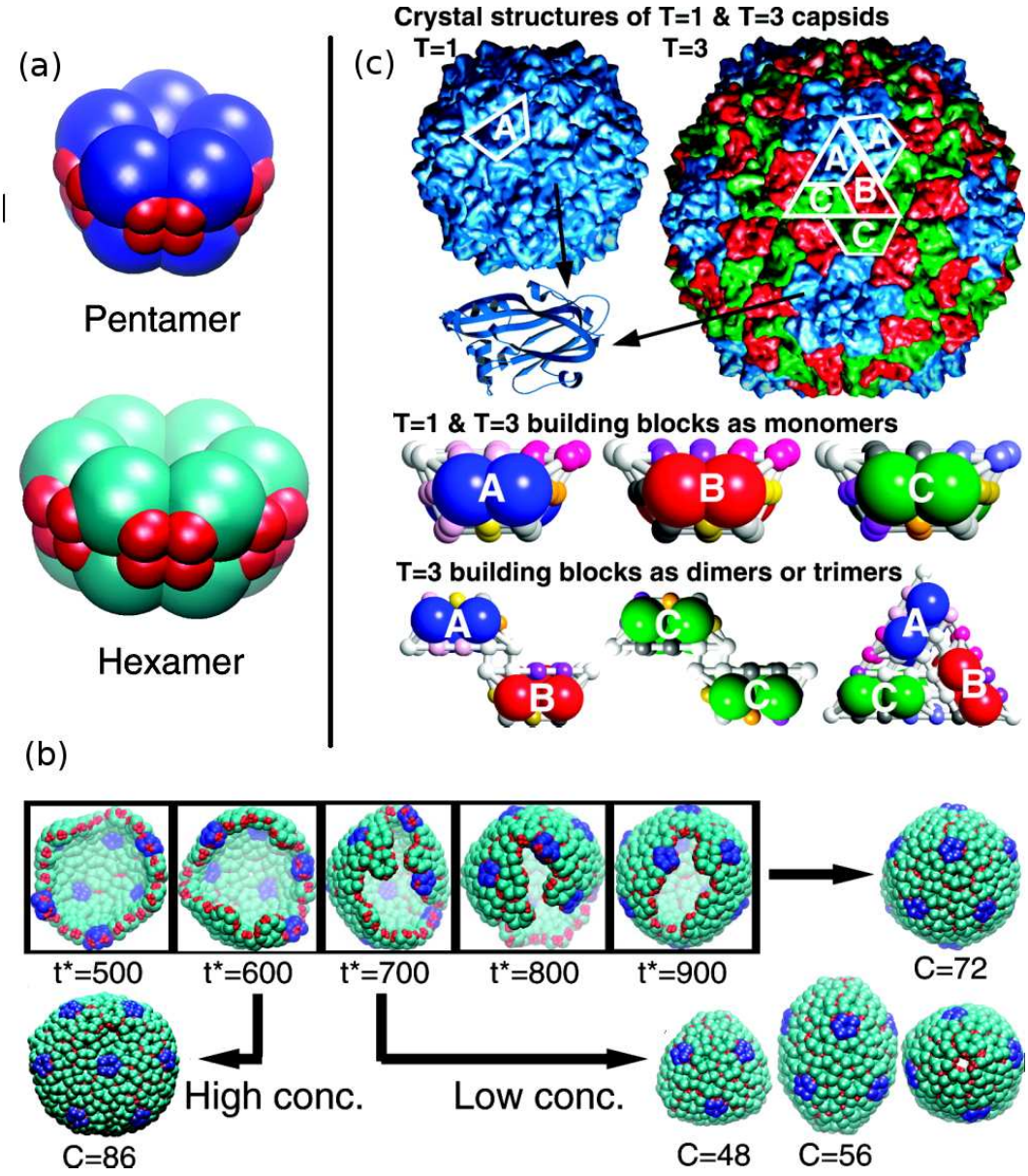

FIGURE 2.17: Models and results from recent work by the Brooks group. (a) Model pentameric and hexameric capsomers used in ref. **??**. (b) The growth process for a $T = 7$ capsid, showing possible misformed capsids, as well as the correct structure in the top right. At low concentrations part-formed capsids tend to close prematuely to form undersized capsids, while at high concentrations multiple intermediates collapse to form oversized capsids with hexameric dislocations. The values of C indicate the numbers of capsomers in the structures. (c) Complete capsids and model subunits used in ref. **??**. The A, B and C subunits represent proteins in different conformations, and have slightly different shapes and different binding rules as appropriate to form the complete capsids. In some simulations the subunits begin in permanent groups of two or three, as shown in the bottom part of the figure. Adapted from Ref. [132] and Ref. [142].

The major findings of these shaped particle simulations tie in well with those of the kinetic simulations - capsid growth appears to proceed through monomer addition via a small set of intermediates, to be nucleation limited, and to be vulnerable to kinetic trapping at low temperatures / high concentrations. The additional feature of misformed and "monster" particles is an interesting one which will prove important in our own work.

2.6.2.3 Spherical particle models

Schwartz et al. carried out an early simulation study of assembly using a simple model whereby the subunits are represented simply as spherical particles[143]. They based their model on a concept termed "local rules"[144], work on which was later described in a helpful summary[**?**]. In the local rules system each protein can exist in one of a set of conformations, and each of the conformations is only able to bind with certain other conformations at certain angles, determined by these "local rules". As each protein is added to the growing capsid, it adopts an appropriate conformation. The realism of this conformational model, in which there is a number of distinct conformations equal to the T number of the capsid with quite distinct binding interactions, is not yet clear. In these particular simulations, however, the target structures were simply $T = 1$ capsids, leading to a relatively simple ruleset. The model included two possible conformations, a stable non-binding form and an unstable form which is able to form bonds. The aspect of assembly on which they concentrated was the formation of malformed clusters. They identified two distinct mechanisms by which malformed clusters could arise. Firstly, a pair of partially completed capsids could come together to form a single oversized structure, similar to those shown in Fig. 2.16(b-c). Secondly, small local errors in the structure of the capsid - such as the formation of a tetrameric ring in place of a pentamer - could lead to significant structural deformities. They highlighted the role of high concentrations, lenient binding interactions (wide sticky patches) and high bonding energies in increasing the rates of malformations by making their formation more likely and their reversal more difficult.

Hagan and Chandler have presented results from simulations with a related model, consisting of spheres patterned with sticky patches[133]. Attractive interactions may only act through the patches, and the particles must also be oriented correctly around the interparticle axis, i.e. the interaction has a dependence on the torsional (dihedral) angle. The patches can be of any of up to five "types", and each type will only interact with a certain subset of other types. This prevents the particles from forming large numbers of undesired bonds, and represents the different interfaces on viral subunits. While the patches are of different types, each particle has the same set of patches and is thus equivalent. Three different models are studied, as shown in Fig. 2.18(a). Each to some extent represents a $T = 1$ virus, in that it contains 60 particles positioned with icosahedral symmetry, but the details of the local environments in each case are different. The "B_5" model probably represents a majority of real virus capsids the best.

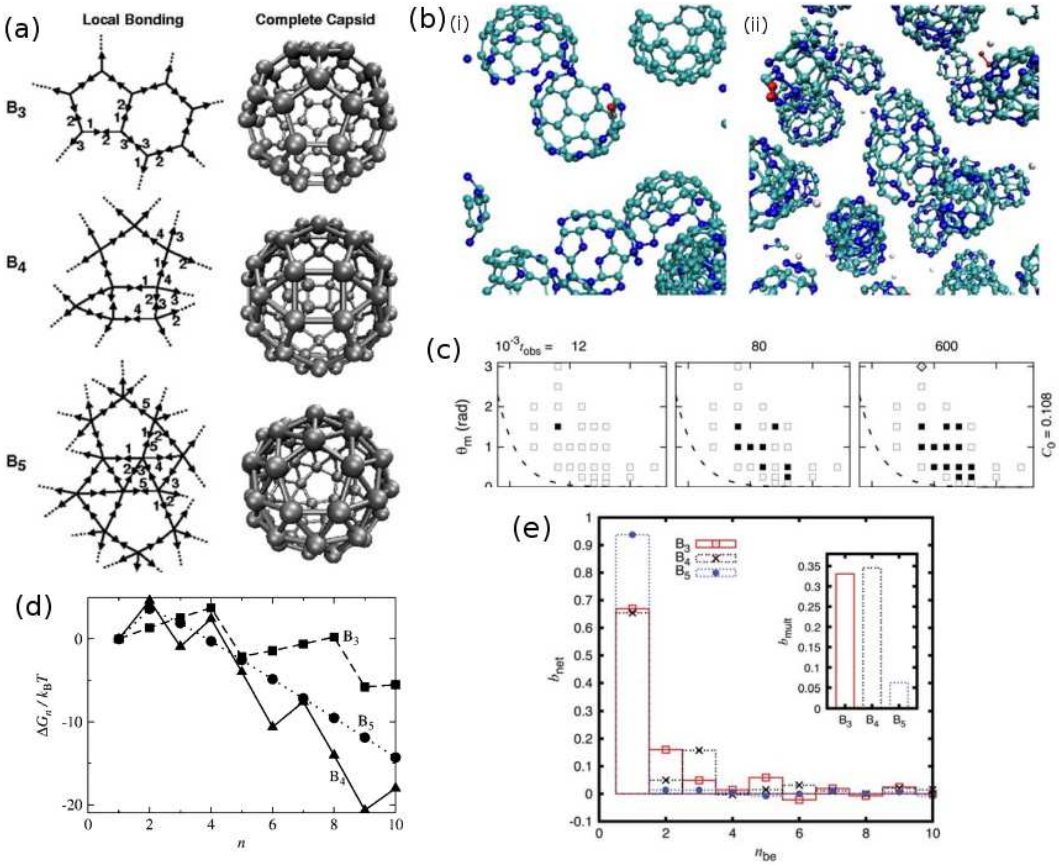

FIGURE 2.18: (a) The three model systems studied by Hagan and Chandler. The left-hand column specifies the model, with the vertices representing subunits and the numbers labelling patches of different types. The patches of each type only form attractive interactions with the appropriate corresponding types. The right-hand column shows complete capsids. (b) Snapshots of kinetic traps: (i) incomplete capsids resulting from monomer starvation, and (ii) misformed capsids resulting from the incorporation of strained bonds. (c) The success of assembly as a function of the binding energy ϵ_b and the patch width θ_m. Solid points denote parameter values for which $30 capsids, while open points denote those for which less than 30$ thermodynamic critical surface, to the left of which capsids are unstable. (d) Free energy profiles for clusters of size n for the three models under optimal conditions, relative to monomers. (e) The fraction of net capsid growth accounted for by the addition of clusters of size n_{be} for the three models. The inset shows the total fractional contribution to capsid growth from the addition of all clusters of size $n_{be} \geq 2$. Adapted from Ref. [133].

The dynamical simulations proceed using molecular dynamics (MD), solving Newton's equations for overdamped motion with an added random buffeting force. By including both strong drag (velocity is proportional only to force, implying very low Reynolds numbers as is appropriate) and Brownian noise, this model has the most appropriate dynamics of those described here.

As with the other models described above, Hagan and Chandler observe hysteresis in capsid formation, and find that optimal values exist for concentration and temperature, and additionally for the width of their patches, for fixed values of the other parameters. At low concentration and patch width and high temperature, monomers dominate and

no large structures are formed, while under the converse conditions assembly is hindered either by kinetic traps - either monomer starvation or the formation of misformed capsids. Fig. 2.18(b) displays both these classes of kinetic trap. The ability to form strained bonds allows the subunits in the model to form a variety of misshapen capsids. The "monster particles" seen here are somewhat less structured than those observed by Nguyen et al. (Fig. 2.16(b)). This is probably because in order to be reasonably stable, misbonded particles in the model of Nguyen et al. are obliged to be rotated by 180° to their native orientation, where as in Hagan and Chandler's model there is more flexibility in the bonding. It is unclear which approach gives more realistic results, although there is probably no physical basis for a preference for rotations of 180°.

Hagan and Chandler attempted to conduct a survey of parameter space, observing the regions in which assembly was successful. Fig. 2.18(c) shows some of their results. These results appear to show that moderate interaction energies and low patch widths are desirable for assembly. No successful assembly occurs close to the calculated thermodynamic line, implying that bonds should be significantly stronger than would be necessary to stabilise the complete capsids. This appears to be somewhat contradictory to Zlotnick's predictions that weak subunit-subunit bonding would be ideal for assembly, and indeed the values of the bonding energy which lead to success, around $15k_BT$, are very high (as compared to values of 5-7kT measured in experiment by Ceres and Zlotnick[**?**]). A possible reason for the difference is that the experimentally measured values are of the free energy of association and hence take into account the unfavorable entropy of association, while the values of the simulation bonding energy do not.

Hagan and Chandler also consider the possible importance of oligomer-oligomer addition steps in assembly. They consider the free energy profile of the early stages of assembly, found by estimating values for the free energies of the most stable intermediate of each size and shown in Fig. 2.18(d). For their "B_5" model, once a triangular nucleus is formed, every subsequent added monomer can form at least two bonds with the structure, and the free energy profile is monotonically downhill beyond the nucleus. For the other two structures, though, many monomer addition steps form only a single bond, leading to an unfavourable free energy of association and a rugged free energy profile. These high free energy states can, however, be bypassed by the combination of oligomers. Fig. 2.18(e) gives information about the degree to which capsid growth occurs by addition of clusters of different sizes. For the B_5 structure growth occurs mainly by monomer addition, but for the other two structures around a third of capsid growth occurs by the addition of oligomers. This is supportive of the work by Schwartz and coworkers[139**?**], indicating that assumptions that growth proceeds only by monomer addition are in some cases unsafe.

More recently Hagan has co-authored a paper with Oren Elrad, making use of a similar patchy spherical particle model, with the spheres in this case representing dimers of brome mosaic virus capsid proteins. However, the particles are able to alternate between

two conformational states, which differ in the angles between the patches, representing the two subunits within the dimer being at different angles to each other. The particles switch between the two states via occasional Monte Carlo-like moves. The particles are able to assemble into either $T = 1$ or $T = 3$ capsids if the particles adopt conformations appropriate to their sites. The structures of differing conformations of a given capsid protein are known in general to be very similar, and consequently the interaction energies between neighbouring subunits are believed to depend only weakly on the conformations. This has raised questions about how ordered arrangements of the various conformations are able to arise in capsids. In this work, however, Elrad and Hagan show that differences in interaction energies as low as the thermal energy are sufficient to produce well-ordered structures.

The work also considers assembly around and encapsidation of large core particles, representing viral cargo. Experimentally purified BMV capsid protein is found to assemble *in vitro* into $T = 3$ capsids, but in the presence of core particles of an appropriate size which attract the capsid proteins, smaller $T = 1$ capsids are formed which encapsulate the core particles. By including large attractive particles to represent the core particles this behaviour is simulated, and in a particular region of parameter space the model system is found to replicate the observed behaviour.

The papers by Hagan and coworkers present perhaps the most rigorous spatial simulations of capsid assembly to date, exploring and testing a number of the issues raised by other workers. The greatest shortcoming, however, lies in the quality of the statistics, with most of their data (not shown here) sampling parameter space rather poorly. This is an inevitable consequence of the design of their model. The work presented in this thesis will share many basic features in common with this work (although the details of our potential and dynamics are rather different and we will not consider the encapsulation of cargo), but will focus on obtaining significantly better statistics and good quality thermodynamic data - while unavoidably sacrificing some detail in the model in the process.

2.6.3 Summary of findings and outstanding questions

The computational studies discussed each approach the problem of capsid assembly from different angles and focus on different points. In considering the body of work together, we find that the following ideas are generally agreed upon:

- Capsid formation is sigmoidal in time
- Low temperatures and high concentrations can lead to kinetic trapping via monomer starvation

- Kinetic trapping can be avoided to some extent if the early stages of assembly are unfavourable
- Hysteresis is expected, with complete capsids stable even under infinite dilution
- Under certain conditions oversized "monster particles" may form (although this possibility is neglected in kinetic simulations).

The following ideas are somewhat controversial, being used as assumptions in some studies but appearing questionable in others:

- Only the most stable intermediate of each size need be considered; in fact it seems that at least a somewhat larger set is likely to be needed
- Capsid growth by addition of oligomers can be ignored since monomer addition dominates; at least under some conditions this is probably incorrect.

Additionally, Schwartz et al.[143], Nguyen et al.[**?**] and Hagan and Chander[133] observe the formation of oversized monster particles, which are either disallowed or not observed in other simulations, but have been seen in experiments.

2.7 The approach and structure of this thesis

As described in Section 2.6.2 there are many possible approaches to simulating viral self-assembly, each of which is obliged to make sacrifices in certain areas while providing clearer insight into others. Our own approach is, firstly, to choose a spatial rather than a kinetic model, so that complicated phenomena can arise emergently without the need for explicit inclusion in the model. Consequently it may be possible to observe misformed capsids, aggregates and a full diversity of assembly pathways which would not be feasible in a kinetic model.

A second choice then lies in the details of the model particles. We choose a minimal model, to allow comprehensive mapping of parameter space and the acquisition of good statistics and equilibrium data. We will model the capsomers simply as spherical particles with attractive "sticky" patches. The disadvantage of this of course is a rather severe loss of biological detail. However, it is our expectation that many of the essential principles of self-assembly will remain the same regardless of the fine details of the system. Protein subunits are obliged to be complex, since they must be formed from chains of amino acids and the interactions in which they are involved inherently require the correct placement of functional groups on the atomic scale. The existence of this complexity of structure should therefore not be taken as an indication that it is necessary for successful self-assembly. If the goal were simply to better understand the assembly of

a single virus it would be wise to include rather more structural detail, insofar as it did not prohibitively slow simulations. However, we feel that in an attempt to further the understanding of the assembly of a range of viruses, as well as potential future synthetic systems, the use of a minimal model is a very valuable approach.

This complexity should not be taken as an indication that complexity is a requirement for successful self-assembly - it is, after all, not easily possible for biological systems to produce patchy spherical particles on the same scale as proteins. There is therefore no indication that their complexity is a result of an evolutionary adapta However, the existence of this complexity does not imply its necessity for self-assembly.

We will describe the details of the model and simulation methods in Chapter **??**, and will then proceed to very carefully examine the assembly of icosahedra from pentavalent particles. If these particles are imagined to represent pentameric capsomers, then this process is analogous to the assembly of $T = 1$ viruses. In Chapter **??** we will extend the study to tetrahedra, octahedra and cubes. These shapes have less relevance to virus assembly, but may be of interest for synthetic self-assembling systems. The differences observed between the different target structures also provide further insight into fundamental self-assembly principles. In Chapter **??** we will modify the model through the inclusion of torsional (dihedral) constraints in the model, bringing the interactions closer to those of capsid subunits (and further from those of likely synthetic patchy particles). Chapter **??** examines the assembly of dodecahedra, which are anomalous among the platonic solids in that they will not assemble in the absence of torsional constraints, for rather complex reasons. Chapter **??** considers methods for the reliable assembly of larger structures by two methods, hierarchical assembly and templating. Finally, in Chapter **??** we study the assembly of $T > 1$ viruses, and the possible effects of scaffolding proteins.

Bibliography

[1] R. Dawkins. *The Selfish Gene*. Oxford University Press, 1976.

[2] CA Suttle. Viruses in the sea. *NATURE*, 437(7057):356–361, SEP 15 2005.

[3] G Rice, L Tang, K Stedman, F Roberto, J Spuhler, E Gillitzer, JE Johnson, T Douglas, and M Young. The structure of a thermophilic archaeal virus shows a double-stranded DNA viral capsid type that spans all domains of life. *PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF THE UNITED STATES OF AMERICA*, 101(20):7716–7720, MAY 18 2004.

[4] D. L. Caspar and A. Klug. Physical principles in the construction of regular viruses. *Cold Spring Harb. Sym.*, 27:1–24, 1962.

[5] X. D. Yan, P. R. Chipman, T. Castberg, G. Bratbak, and T. S. Baker. The marine algal virus ppv01 has an icosahedral capsid with t=219 quasisymmetry. *JOURNAL OF VIROLOGY*, 79(14):9236–9243, JUL 2005.

[6] A. Zlotnick. Theoretical aspects of virus capsid assembly. *J. Mol. Recognit.*, 18:479–490, 2005.

[7] WW Newcomb, FL Homa, and JC Brown. Involvement of the portal at an early step in herpes simplex virus capsid assembly. *JOURNAL OF VIROLOGY*, 79(16):10540–10546, AUG 2005.

[8] PA Thuman-Commike, B Greene, JA Malinski, J King, and W Chiu. Role of the scaffolding protein in p22 procapsid size determination suggested by t = 4 and t = 7 procapsid structures. *Biophys. J.*, 74(1):559–568, JAN 1998.

[9] A Maraver, A Ona, F Abaitua, D Gonzalez, R Clemente, JA Ruiz-Diaz, JR Caston, F Pazos, and JF Rodriguez. The oligomerization domain of vp3, the scaffolding protein of infectious bursal disease virus, plays a critical role in capsid assembly. *J. Virol.*, 77(11):6438–6449, JUN 2003.

[10] Binbin Deng, Christine M. O'Connor, Dean H. Kedes, and Z. Hong Zhou. Cryo-electron tomography of kaposi's sarcoma-associated herpesvirus capsids reveals dynamic scaffolding structures essential to capsid assembly and maturation. *J. Struct. Biol.*, 161(3):419–427, MAR 2008.

[11] A Zlotnick, R Aldrich, JM Johnson, P Ceres, and MJ Young. Mechanism of capsid assembly for an icosahedral plant virus. *Virology*, 277(2):450–456, NOV 25 2000.

[12] Adam Zlotnick, Jennifer M. Johnson, Paul W. Wingfield, Stephen J. Stahl, and Dan Endres. A theoretical model successfully identifies features of hepatitis b virus capsid assembly. *Biochemistry*, 38(44):14644–14652, 1999.

[13] G. L. Casini, D. Graham, D Heine, R. L. Garcea, and D. T. Wu. *In vitro* papillomavirus capsid assembly analysed by light scattering. *Virology*, 325:320–327, 2004.

[14] PE Prevelige. Inhibiting virus-capsid assembly by altering the polymerisation pathway. *Trends Biotechnol.*, 16(2):61–65, FEB 1998.

[15] Adam Zlotnick, Angela Lee, Christina R. Bourne, Jennifer M. Johnson, Paul L. Domanico, and Stephen J. Stray. In vitro screening for molecules that affect virus capsid assembly (and other protein association reactions). *NATURE PROTOCOLS*, 2(3):490–498, 2007.

[16] A Zlotnick, P Ceres, S Singh, and JM Johnson. A small molecule inhibits and misdirects assembly of hepatitis b virus capsids. *J. Virol.*, 76(10):4848–4854, MAY 2002.

[17] HJ Hacker, K Deres, M Mildenberger, and CH Schroder. Antivirals interacting with hepatitis B virus core protein and core mutations may misdirect capsid assembly in a similar fashion. *BIOCHEMICAL PHARMACOLOGY*, 66(12):2273–2279, DEC 15 2003.

[18] S. J. Stray, C. R. Bourne, S. Punna, W. G. Lewis, M. G. Finn, and A. Zlotnick. A heteroaryldihydropyrimidine activates and can misdirect hepatitis b virus capsid assembly. *Proc. Natl. Acad. Sci. USA*, 102:8138–8143, 2005.

[19] SJ Stray, JM Johnson, BG Kopek, and A Zlotnick. An in vitro fluorescence screen to identify antivirals that disrupt hepatitis B virus capsid assembly. *NATURE BIOTECHNOLOGY*, 24(3):358–362, MAR 2006.

[20] L. A. Koutsky, K. A. Ault, C. M. Wheeler, D. R. Brown, E. Barr, F. B. Alvarez, L. M. Chiacchierini, and K. U. Jansen. A controlled trial of a human papillomavirus type 16 vaccine. *New Engl. J. Med.*, 347(21):1645–1651, 2002.

[21] M Shank-Retzlaff, F Wang, T Morley, C Anderson, M Hamm, M Brown, K Rowland, G Pancari, J Zorman, R Lowe, L Schultz, J Teyral, R Capen, C B Oswald, Y Wang, M Washabaugh, K Jansen, and R Sitrin. Correlation between mouse potency and in vitro relative potency for human papillomavirus type 16 virus-like particles and gardasil vaccine samples. *Human Vaccines*, 1(5):191, 2005.

[22] L. Shi, H. L. Sings, J. T. Bryan, B. Wang, Y. Wang, H. Mach, M. Kosinski, M. W. Washabaugh, R. Sitrin, and E. Barr. GARDASIL (R): Prophylactic human papillomavirus vaccine development - From bench top to bed-side. *CLINICAL PHARMACOLOGY & THERAPEUTICS*, 81(2):259–264, FEB 2007.

[23] Mark Young, Debbie Willits, Masaki Uchida, and Trevor Douglas. Plant viruses as biotemplates for materials and their use in nanotechnology. *ANNUAL REVIEW OF PHYTOPATHOLOGY*, 46:361–384, 2008.

[24] P Singh, MJ Gonzalez, and M Manchester. Viruses and their uses in nanotechnology. *DRUG DEVELOPMENT RESEARCH*, 67(1):23–41, JAN 2006.

[25] T Douglas and M Young. Viruses: Making friends with old foes. *Science*, 312(5775):873–875, MAY 12 2006.

[26] ML Flenniken, LO Liepold, BE Crowley, DA Willits, MJ Young, and T Douglas. Selective attachment and release of a chemotherapeutic agent from the interior of a protein cage architecture. *CHEMICAL COMMUNICATIONS*, (4):447–449, JAN 28 2005.

[27] PL Sinn, SL Sauter, and PL McCray. Gene Therapy Progress and Prospects: Development of improved lentiviral and retroviral vectors - design, biosafety, and production. *GENE THERAPY*, 12(14):1089–1098, JUL 2005.

[28] Lars Liepold, Stasia Anderson, Deborah Willits, Luke Oltrogge, Joseph A. Frank, Trevor Douglas, and Mark Young. Viral capsids as mri contrast agents. *Magnet. Reson. Med.*, 58(5):871–879, NOV 2007.

[29] T Douglas and M Young. Host-guest encapsulation of materials by assembled virus protein cages. *Nature*, 393(6681):152–155, MAY 14 1998.

[30] M Knez, AM Bittner, F Boes, C Wege, H Jeske, E Maiss, and K Kern. Biotemplate synthesis of 3-nm nickel and cobalt nanowires. *NANO LETTERS*, 3(8):1079–1082, AUG 2003.

[31] T Douglas, E Strable, D Willits, A Aitouchen, M Libera, and M Young. Protein engineering of a viral cage for constrained nanomaterials synthesis. *ADVANCED MATERIALS*, 14(6):415+, MAR 18 2002.

[32] E Gillitzer, D Willits, M Young, and T Douglas. Chemical modification of a viral cage for multivalent presentation. *CHEMICAL COMMUNICATIONS*, (20):2390–2391, 2002.

[33] JD Lewis, G Destito, A Zijlstra, MJ Gonzalez, JP Quigley, M Manchester, and H Stuhlmann. Viral nanoparticles as tools for intravital vascular imaging. *Nat. Med.*, 12(3):354–360, MAR 2006.

[34] Q Wang, TW Lin, JE Johnson, and MG Finn. Natural supramolecular building blocks: Cysteine-added mutants of cowpea mosaic virus. *CHEMISTRY & BIOLOGY*, 9(7):813–819, JUL 2002.

[35] ML Flenniken, DA Willits, AL Harmsen, LO Liepold, AG Harmsen, MJ Young, and T Douglas. Melanoma and lymphocyte cell-specific targeting incorporated into a heat shock protein cage architecture. *CHEMISTRY & BIOLOGY*, 13(2):161–170, FEB 2006.

[36] A Chatterji, W Ochoa, L Shamieh, SP Salakian, SM Wong, G Clinton, P Ghosh, TW Lin, and JE Johnson. Chemical conjugation of heterologous proteins on the surface of cowpea mosaic virus. *BIOCONJUGATE CHEMISTRY*, 15(4):807–813, JUL-AUG 2004.

[37] E Strable, JE Johnson, and MG Finn. Natural nanochemical building blocks: Icosahedral virus particles organized by attached oligonucleotides. *NANO LETTERS*, 4(8):1385–1389, AUG 2004.

[38] KS Raja, Q Wang, and MG Finn. Icosahedral virus particles as polyvalent carbohydrate display platforms. *CHEMBIOCHEM*, 4(12):1348–1351, DEC 5 2003.

[39] J. A. Speir, S. Munshi, G. J. Wang, T. S. Baker, and J. E. Johnson. Structures of the native and swollen forms of cowpea chlorotic mottle virus determined by x-ray crystallography and cryoelectron microscopy. *Structure*, 3(1):63–78, JAN 15 1995.

[40] M Allen, JWM Bulte, L Liepold, G Basu, HA Zywicke, JA Frank, M Young, and T Douglas. Paramagnetic viral nanoparticles as potential high-relaxivity magnetic resonance contrast agents. *Magnet. Reson. Med.*, 54(4):807–812, OCT 2005.

[41] MT Klem, D Willits, M Young, and T Douglas. 2-D array formation of genetically engineered viral cages on Au surfaces and imaging by atomic force microscopy. *JOURNAL OF THE AMERICAN CHEMICAL SOCIETY*, 125(36):10806–10807, SEP 10 2003.

[42] Y LVOV, H HAAS, G DECHER, H MOHWALD, A MIKHAILOV, B MTCHEDLISHVILY, E MORGUNOVA, and B VAINSHTEIN. SUCCESSIVE DEPOSITION OF ALTERNATE LAYERS OF POLYELECTROLYTES AND A CHARGED VIRUS. *LANGMUIR*, 10(11):4232–4236, NOV 1994.

[43] Peter A. Suci, Michael T. Klem, Fernando T. Arce, Trevor Douglas, and Mark Young. Assembly of multilayer films incorporating a viral protein cage architecture. *LANGMUIR*, 22(21):8891–8896, OCT 10 2006.

[44] LiNa Loo, Richard H. Guenther, Steven A. Lommel, and Stefan Franzen. Encapsidation of nanoparticies by Red Clover Necrotic Mosaic Virus. *JOURNAL OF THE AMERICAN CHEMICAL SOCIETY*, 129(36):11111–11117, SEP 12 2007.

[45] Suraj K. Dixit, Nancy L. Goicochea, Marie-Christine Daniel, Ayaluru Murali, Lyudmila Bronstein, Mrinmoy De, Barry Stein, Vincent M. Rotello, C. Cheng Kao, and Bogdan Dragnea. Quantum dot encapsulation in viral capsids. *NANO LETTERS*, 6(9):1993–1999, SEP 13 2006.

[46] Marta Comellas-Aragones, Hans Engelkamp, Victor I. Claessen, Nico A. J. M. Sommerdijk, Alan E. Rowan, Peter C. M. Christianen, Jan C. Maan, Benedictus J. M. Verduin, Jeroen J. L. M. Cornelissen, and Roeland J. M. Nolte. A virus-based single-enzyme nanoreactor. *NATURE NANOTECHNOLOGY*, 2(10):635–639, 2007.

[47] NG Portney, K Singh, S Chaudhary, G Destito, A Schneemann, M Manchester, and M Ozkan. Organic and inorganic nanoparticle hybrids. *LANGMUIR*, 21(6):2098–2103, MAR 15 2005.

[48] Nathaniel G. Portney, Ricky J. Tseng, Giuseppe Destito, Erica Strable, Yang Yang, Marianne Manchester, M. G. Finn, and Mihrimah Ozkan. Microscale memory characteristics of virus-quantum dot hybrids. *APPLIED PHYSICS LETTERS*, 90(21), MAY 21 2007.

[49] Rebekah A. Miller, Andrew D. Presley, and Matthew B. Francis. Self-assembling light-harvesting systems from synthetically modified tobacco mosaic virus coat proteins. *JOURNAL OF THE AMERICAN CHEMICAL SOCIETY*, 129(11):3104–3109, MAR 21 2007.

[50] Pratik Singh, Giuseppe Destito, Anette Schneemann, and Marianne Manchester. Canine parvovirus-like particles, a novel nanomaterial for tumor targeting. *J. Nanobiotechnol.*, 4:2, Feb 2006.

[51] Robert L. Garcea. Biologic constraints on modelling virus assembly. *COMPUTATIONAL AND MATHEMATICAL METHODS IN MEDICINE*, 9(3-4):257–264, 2008.

[52] T Dokland. Scaffolding proteins and their role in viral assembly. *Cell. Mol. Life Sci.*, 56(7-8):580–603, NOV 15 1999.

[53] BA Fane and PE Prevelige. Mechanism of scaffolding-assisted viral assembly. In *VIRUS STRUCTURE*, volume 64 of *ADVANCES IN PROTEIN CHEMISTRY*, pages 259+. ACADEMIC PRESS INC, 525 B STREET, SUITE 1900, SAN DIEGO, CA 92101-4495 USA, 2003.

[54] R van Driel and E Couture. Assembly of scaffolding core of bacteriophage-t4 preheads. *J. Mol. Biol.*, 123(4):713–719, 1978.

[55] F Arisaka. Assembly and infection process of bacteriophage t4. *Chaos*, 15(4), DEC 2005.

[56] P. X. Guo, S. Erickson, W. Xu, N. Olson, T. S. Baker, and D. Anderson. Regulation of the phage-29 prohead shape and size by the portal vertex. *Virology*, 183(1):366–373, JUL 1991.

[57] P. K. Sorger, P. G. Stockley, and S. C. Harrison. Structure and assembly of turnip crinkle virus .2. mechanism of reassembly in vitro. *J. Mol. Biol.*, 191(4):639–658, OCT 20 1986.

[58] O. M. Elrad and M. F. Hagan. Mechanisms of size control and polymorphism in viral capsid assembly. *Nano Lett.*, 8(11):3850–3857, 2008.

[59] Jingchuan Sun, Chris DuFort, Marie-Christine Daniel, Ayaluru Murali, Chao Chen, Kodetham Gopinath, Barry Stein, Mrinmoy De, Vincent M. Rotello, Andreas Holzenburg, C. Cheng Kao, and Bogdan Dragnea. Core-controlled polymorphism in virus-like particles. *Proc. Natl. Acad. Sci. USA*, 104(4):1354–1359, JAN 23 2007.

[60] O Cayre, VN Paunov, and OD Velev. Fabrication of asymmetrically coated colloid particles by microcontact printing techniques. *JOURNAL OF MATERIALS CHEMISTRY*, 13(10):2445–2450, 2003.

[61] L Manna, EC Scher, and AP Alivisatos. Synthesis of soluble and processable rod-, arrow-, teardrop-, and tetrapod-shaped CdSe nanocrystals. *JOURNAL OF THE AMERICAN CHEMICAL SOCIETY*, 122(51):12700–12706, DEC 27 2000.

[62] YG Sun and YN Xia. Shape-controlled synthesis of gold and silver nanoparticles. *SCIENCE*, 298(5601):2176–2179, DEC 13 2002.

[63] LF Gou and CJ Murphy. Solution-phase synthesis of Cu2O nanocubes. *NANO LETTERS*, 3(2):231–234, FEB 2003.

[64] KS Cho, DV Talapin, W Gaschler, and CB Murray. Designing PbSe nanowires and nanorings through oriented attachment of nanoparticles. *JOURNAL OF THE AMERICAN CHEMICAL SOCIETY*, 127(19):7140–7147, MAY 18 2005.

[65] TS Ahmadi, ZL Wang, TC Green, A Henglein, and MA ElSayed. Shape-controlled synthesis of colloidal platinum nanoparticles. *SCIENCE*, 272(5270):1924–1926, JUN 28 1996.

[66] EC Greyson, JE Barton, and TW Odom. Tetrahedral zinc blende tin sulfide nanoand microcrystals. *SMALL*, 2(3):368–371, MAR 2006.

[67] RC Jin, YW Cao, CA Mirkin, KL Kelly, GC Schatz, and JG Zheng. Photoinduced conversion of silver nanospheres to nanoprisms. *SCIENCE*, 294(5548):1901–1903, NOV 30 2001.

[68] N Malikova, I Pastoriza-Santos, M Schierhorn, NA Kotov, and LM Liz-Marzan. Layer-by-layer assembled mixed spherical and planar gold nanoparticles: Control of interparticle interactions. *LANGMUIR*, 18(9):3694–3697, APR 30 2002.

[69] SS Shankar, A Rai, B Ankamwar, A Singh, A Ahmad, and M Sastry. Biological synthesis of triangular gold nanoprisms. *NATURE MATERIALS*, 3(7):482–488, JUL 2004.

[70] D Dendukuri, DC Pregibon, J Collins, TA Hatton, and PS Doyle. Continuous-flow lithography for high-throughput microparticle synthesis. *NATURE MATERIALS*, 5(5):365–369, MAY 2006.

[71] SH Chen, ZL Wang, J Ballato, SH Foulger, and DL Carroll. Monopod, bipod, tripod, and tetrapod gold nanocrystals. *J. Am. Chem. Soc.*, 125(52):16186–16187, DEC 31 2003.

[72] ZY Tang, Y Wang, S Shanbhag, M Giersig, and NA Kotov. Spontaneous transformation of CdTe nanoparticles into angled Te nanocrystals: From particles and rods to checkmarks, X-marks, and other unusual shapes. *JOURNAL OF THE AMERICAN CHEMICAL SOCIETY*, 128(20):6730–6736, MAY 24 2006.

[73] L Manna, DJ Milliron, A Meisel, EC Scher, and AP Alivisatos. Controlled growth of tetrapod-branched inorganic nanocrystals. *NATURE MATERIALS*, 2(6):382–385, JUN 2003.

[74] SM Lee, YW Jun, SN Cho, and J Cheon. Single-crystalline star-shaped nanocrystals and their evolution: Programming the geometry of nano-building blocks. *JOURNAL OF THE AMERICAN CHEMICAL SOCIETY*, 124(38):11244–11245, SEP 25 2002.

[75] G. A. DeVries, M. Brunnbauer, Y. Hu, A. M. Jackson, B. Long, B. T. Neltner, O. Uzun, B. H. Wunsch, and F. Stellacci. Divalent metal nanoparticles. *Science*, 315:358–361, 2007.

[76] CJ Loweth, WB Caldwell, XG Peng, AP Alivisatos, and PG Schultz. DNA-based assembly of gold nanocrystals. *ANGEWANDTE CHEMIE-INTERNATIONAL EDITION*, 38(12):1808–1812, 1999.

[77] PM Johnson, CM van Kats, and A van Blaaderen. Synthesis of colloidal silica dumbbells. *LANGMUIR*, 21(24):11510–11517, NOV 22 2005.

[78] Y Lu, H Xiong, XC Jiang, YN Xia, M Prentiss, and GM Whitesides. Asymmetric dimers can be formed by dewetting half-shells of gold deposited on the surfaces of spherical oxide colloids. *JOURNAL OF THE AMERICAN CHEMICAL SOCIETY*, 125(42):12724–12725, OCT 22 2003.

[79] V. N. Manoharan, M. T. Elsesser, and D. J. Pine. Dense packing and symmetry in small clusters of microspheres. *Science*, 301:483–487, 2003.

[80] XG Peng, L Manna, WD Yang, J Wickham, E Scher, A Kadavanich, and AP Alivisatos. Shape control of CdSe nanocrystals. *NATURE*, 404(6773):59–61, MAR 2 2000.

[81] YY Yu, SS Chang, CL Lee, and CRC Wang. Gold nanorods: Electrochemical synthesis and optical properties. *JOURNAL OF PHYSICAL CHEMISTRY B*, 101(34):6661–6664, AUG 21 1997.

[82] FM van der Kooij, K Kassapidou, and HNW Lekkerkerker. Liquid crystal phase transitions in suspensions of polydisperse plate-like particles. *NATURE*, 406(6798):868–871, AUG 24 2000.

[83] A Mohraz and MJ Solomon. Direct visualization of colloidal rod assembly by confocal microscopy. *LANGMUIR*, 21(12):5298–5306, JUN 7 2005.

[84] Alicia M. Jackson, Ying Hu, Paula Jacob Silva, and Francesco Stellacci. From homoligand- to mixed-ligand-monolayer-protected metal nanoparticles: A scanning tunneling microscopy investigation. *JOURNAL OF THE AMERICAN CHEMICAL SOCIETY*, 128(34):11135–11149, AUG 30 2006.

[85] K.-H. Roh, D. C. Martin, and J. Lahann. Biphasic janus particles with nanoscale anisotropy. *Nat. Mater.*, 4:759–763, 2005.

[86] G Zhang, DY Wang, and H Mohwald. Decoration of microspheres with gold nanodots-giving colloidal spheres valences. *ANGEWANDTE CHEMIE-INTERNATIONAL EDITION*, 44(47):7767–7770, 2005.

[87] BR Martin, DJ Dermody, BD Reiss, MM Fang, LA Lyon, MJ Natan, and TE Mallouk. Orthogonal self-assembly on colloidal gold-platinum nanorods. *ADVANCED MATERIALS*, 11(12):1021–1025, AUG 2 1999.

[88] Adeline Perro, Fabrice Meunier, Veronique Schmitt, and Serge Ravaine. Production of large quantities of "Janus" nanoparticles using wax-in-water emulsions. *COLLOIDS AND SURFACES A-PHYSICOCHEMICAL AND ENGINEERING ASPECTS*, 332(1):57–62, JAN 5 2009.

[89] Sharon C. Glotzer and Michael J. Solomon. Anisotropy of building blocks and their assembly into complex structures. *NATURE MATERIALS*, 6(8):557–562, AUG 2007.

[90] E YABLONOVITCH, TJ GMITTER, and KM LEUNG. PHOTONIC BAND-STRUCTURE - THE FACE-CENTERED-CUBIC CASE EMPLOYING NON-SPHERICAL ATOMS. *PHYSICAL REVIEW LETTERS*, 67(17):2295–2298, OCT 21 1991.

[91] KM HO, CT CHAN, CM SOUKOULIS, R BISWAS, and M SIGALAS. PHOTONIC BAND-GAPS IN 3-DIMENSIONS - NEW LAYER-BY-LAYER PERIODIC STRUCTURES. *SOLID STATE COMMUNICATIONS*, 89(5):413–416, FEB 1994.

[92] ZL Zhang, AS Keys, T Chen, and SC Glotzer. Self-assembly of patchy particles into diamond structures through molecular mimicry. *LANGMUIR*, 21(25):11547–11551, DEC 6 2005.

[93] Seung-Man Yang, Shin-Hyun Kim, Jong-Min Lim, and Gi-Ra Yi. Synthesis and assembly of structured colloidal particles. *JOURNAL OF MATERIALS CHEMISTRY*, 18(19):2177–2190, 2008.

[94] C CASAGRANDE and M VEYSSIE. JANUS BEADS - REALIZATION AND 1ST OBSERVATION OF INTERFACIAL PROPERTIES. *COMPTES RENDUS DE L ACADEMIE DES SCIENCES SERIE II*, 306(20):1423–1425, MAY 28 1988.

[95] Jing-Qin Cui and Ilona Kretzschmar. Surface-anisotropic polystyrene spheres by electroless deposition. *LANGMUIR*, 22(20):8281–8284, SEP 26 2006.

[96] VN Paunov and OJ Cayre. Supraparticles and "Janus" particles fabricated by replication of particle monolayers at liquid surfaces using a gel trapping technique. *ADVANCED MATERIALS*, 16(9-10):788+, MAY 17 2004.

[97] ZN Bao, L Chen, M Weldon, E Chandross, O Cherniavskaya, Y Dai, and JBH Tok. Toward controllable self-assembly of microstructures: Selective functionalization and fabrication of patterned spheres. *CHEMISTRY OF MATERIALS*, 14(1):24+, JAN 2002.

[98] Liang Hong, Angelo Cacciuto, Erik Luijten, and Steve Granick. Clusters of charged Janus spheres. *NANO LETTERS*, 6(11):2510–2514, NOV 8 2006.

[99] A Perro, S Reculusa, F Pereira, MH Delville, C Mingotaud, E Duguet, E Bourgeat-Lami, and S Ravaine. Towards large amounts of Janus nanoparticles through a protection-deprotection route. *CHEMICAL COMMUNICATIONS*, (44):5542–5543, NOV 28 2005.

[100] Takasi Nisisako and Toru Torii. Formation of biphasic Janus droplets in a microfabricated channel for the synthesis of shape-controlled polymer microparticles. *ADVANCED MATERIALS*, 19(11):1489+, JUN 4 2007.

[101] KH Roh, DC Martin, and J Lahann. Triphasic nanocolloids. *JOURNAL OF THE AMERICAN CHEMICAL SOCIETY*, 128(21):6796–6797, MAY 31 2006.

[102] Y.-S. Cho, G.-R. Yi, J.-M. Lim, S.-H. Kim, V. N. Manoharan, D. J. Pine, and S.-M. Yang. Self-organization of bidisperse colloids in water droplets. *J. Am. Chem. Soc.*, 127:15968–15975, 2005.

[103] Likui Wang, Linhua Xia, Gang Li, Serge Ravaine, and X. S. Zhao. Patterning the surface of colloidal microspheres and fabrication of nonspherical particles. *ANGEWANDTE CHEMIE-INTERNATIONAL EDITION*, 47(25):4725–4728, 2008.

[104] JH CHEN and NC SEEMAN. SYNTHESIS FROM DNA OF A MOLECULE WITH THE CONNECTIVITY OF A CUBE. *NATURE*, 350(6319):631–633, APR 18 1991.

[105] YW ZHANG and NC SEEMAN. CONSTRUCTION OF A DNA-TRUNCATED OCTAHEDRON. *JOURNAL OF THE AMERICAN CHEMICAL SOCIETY*, 116(5):1661–1669, MAR 9 1994.

[106] R. P. Goodman, I. A. T. Schaap, C. F. Tardin, C. M. Erben, R. M. Berry, C. F. Schmidt, and A. J. Turberfield. Rapid chiral assembly of rigid dna building blocks for molecular nanofabrication. *Nature*, 310:1661–1665, 2005.

[107] Russell P. Goodman, Mike Heilemann, Soeren Doose, Christoph M. Erben, Achillefs N. Kapanidis, and Andrew J. Turberfield. Reconfigurable, braced, three-dimensional DNA nanostructures. *NATURE NANOTECHNOLOGY*, 3(2):93–96, FEB 2008.

[108] Christoph M. Erben, Russell P. Goodman, and Andrew J. Turberfield. Single-molecule protein encapsulation in a rigid DNA cage. *ANGEWANDTE CHEMIE-INTERNATIONAL EDITION*, 45(44):7414–7417, 2006.

[109] Yu He, Tao Ye, Min Su, Chuan Zhang, Alexander E. Ribbe, Wen Jiang, and Chengde Mao. Hierarchical self-assembly of DNA into symmetric supramolecular polyhedra. *NATURE*, 452(7184):198–U41, MAR 13 2008.

[110] Chuan Zhang, Min Su, Yu He, Xin Zhao, Ping-an Fang, Alexander E. Ribbe, Wen Jiang, and Chengde Mao. Conformational flexibility facilitates self-assembly of complex DNA nanostructures. *PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF THE UNITED STATES OF AMERICA*, 105(31):10665–10669, AUG 5 2008.

[111] Chuan Zhang, Seung Hyeon Ko, Min Su, Yujun Leng, Alexander E. Ribbe, Wen Jiang, and Chengde Mao. Symmetry Controls the Face Geometry of DNA Polyhedra. *JOURNAL OF THE AMERICAN CHEMICAL SOCIETY*, 131(4):1413+, FEB 4 2009.

[112] Z. Zhang and S. C. Glotzer. Self-assembly of patchy particles. *Nano Lett.*, 4:1407–1413, 2004.

[113] Santanu Mukherjee, Marc V. Thorsteinsson, Lauren Beth Johnston, Pete A. DePhillips, and Adam Zlotnick. A quantitative description of in vitro assembly of human papillomavirus 16 virus-like particles. *J. Mol. Biol.*, 381(1):229–237, AUG 1 2008.

[114] C. M. Teschke, A. McGough, and P. A. Thuman-Commike. Penton release from p22 heat-expanded capsids suggests importance of stabilizing penton-hexon interactions during capsid maturation. *Biophys. J.*, 84:2585–2592, 2003.

[115] D Willits, X Zhao, N Olson, TS Baker, A Zlotnick, JE Johnson, T Douglas, and MJ Young. Effects of the cowpea chlorotic mottle bromovirus beta-hexamer structure on virion assembly. *Virology*, 306(2):280–288, FEB 15 2003.

[116] J. M. Johnson, D. A. Willits, M. J. Young, and A. Zlotnick. Interaction with capsid protein alters rna structure and the pathway for in vitro assembly of cowpea chlorotic mottle virus. *J. Mol. Biol.*, 335:455–464, 2004.

[117] K. N. Parent, S. M. Doyle, E. Anderson, and C. M. Teschke. Electrostatic interactions govern both nucleation and elongation during phage p22 procapsid assembly. *Virology*, 340:33–45, 2005.

[118] K. N. Parent, A. Zlotnick, and C. M. Teschke. Quantitative analysis of multicomponent spherical virus assembly: Scaffolding protein contributes to the global stability of phage p22 procapsids. *J. Mol. Biol.*, 359:1097–1106, 2006.

[119] K. N. Parent, M. M. Suhanovsky, and C. M. Teschke. Phage p22 procapsids equilibrate with free coat protein subunits. *J. Mol. Bio.*, 365:513–522, 2007.

[120] SJ Hanslip, NR Zaccai, APJ Middelberg, and RJ Falconer. Assembly of human papillomavirus type-16 virus-like particles: Multifactorial study of assembly and competing aggregation. *Biotechnol. Progr.*, 22(2):554–560, MAR-APR 2006.

[121] Simon J. Hanslip, Nathan R. Zaccai, Anton P. J. Middelberg, and Robert J. Falconer. Intrinsic fluorescence as an analytical probe of virus-like particle assembly and maturation. *Biochem. Bioph. Res. Co.*, 375(3):351–355, OCT 24 2008.

[122] Roman Tuma, Hiro Tsuruta, Kenneth H. French, and Peter E. Prevelige. Detection of intermediates and kinetic control during assembly of bacteriophage p22 procapsid. *J. Mol. Biol.*, 381(5):1395–1406, SEP 19 2008.

[123] Christina Bourne, Sejin Lee, Bollu Venkataiah, Angela Lee, Brent Korba, M. G. Finn, and Adam Zlotnick. Small-molecule effectors of hepatitis b virus capsid assembly give insight into virus life cycle. *J. Virol.*, 82(20):10262–10270, OCT 2008.

[124] Tao Hu and B. I. Shklovskii. Kinetics of viral self-assembly: Role of the single-stranded RNA antenna. *PHYSICAL REVIEW E*, 75(5, Part 1), MAY 2007.

[125] Ottar Rolfsson, Katerina Toropova, Victoria Morton, Simona Francese, Gabriella Basnak, Gary S. Thompson, Stephen W. Homans, Alison E. Ashcroft, Nicola J. Stonehouse, Neil A. Ranson, and Peter G. Stockley. RNA packing specificity and

folding during assembly of the bacteriophage MS2. *COMPUTATIONAL AND MATHEMATICAL METHODS IN MEDICINE*, 9(3-4):339–349, 2008.

[126] D. Endres and A. Zlotnick. Model-based analysis of assembly kinetics for virus capsids or other spherical polymers. *Biophys. J.*, 83:1217–1230, 2002.

[127] PL Freddolino, AS Arkhipov, SB Larson, A McPherson, and K Schulten. Molecular dynamics simulations of the complete satellite tobacco mosaic virus. *Structure*, 14(3):437–449, MAR 2006.

[128] Anton Arkhipov, Peter L. Freddolino, and Klaus Schulten. Stability and dynamics of virus capsids described by coarse-grained modeling. *Structure*, 14(12):1767–1777, DEC 2006.

[129] D. C. Rapaport. Self-assembly of polyhedral shells: A molecular dynamics study. *Phys. Rev. E*, 70:051905, 2004.

[130] D. C. Rapaport. Role of reversibility in viral capsid growth: A paradigm for self-assembly. *Phys. Rev. Lett.*, 101(18):186101, 2008.

[131] H. D. Nguyen, V. S. Reddy, and C. L. Brooks III. Deciphering the kinetic mechanism of spontaneous self-assembly of icosahedral capsids. *Nano Lett.*, 7:338–344, 2007.

[132] H. D. Nguyen and C. L. Brooks III. Generalized structural polymorphism in self-assembled viral particles. *Nano Lett.*, 8(12):4574–4581, 2008.

[133] M. F. Hagan and D. Chandler. Dynamic pathways for viral capsid assembly. *Biophys. J.*, 91:42–54, 2006.

[134] A. W. Wilber, J. P. K. Doye, A. A. Louis, E. G. Noya, M. A. Miller, and P. Wong. Reversible self-assembly of patchy particles into monodisperse icosahedral clusters. *J. Chem. Phys.*, 127(8):085106, 2007.

[135] F. Oosawa and Asakura S. *Thermodynamics of the Polymerization of Protein*. Academic Press, New York, 1975.

[136] A. R. Fersht. *Structure and Mechanism in Protein Science*. W.H. Freeman & Co., New York, 1999.

[137] A. Zlotnick. Are weak protein-protein interactions the general rule in capsid assembly? *Virology*, 315:269–274, 2003.

[138] D. Endres, M. Miyahara, P. Moisant, and A. Zlotnick. A reaction landscape identifies the intermediates critical for self-assembly of virus capsids and other polyhedral structures. *Protein Sci.*, 14:1518–1525, 2005.

[139] T. Q. Zhang and R. Schwartz. Simulation study of the contribution of oligomer/oligomer binding to capsid assembly kinetics. *Biophys. J.*, 90:57–64, 2006.

[140] DC Rapaport, JE Johnson, and J Skolnick. Supramolecular self-assembly: molecular dynamics modeling of polyhedral shell formation. *COMPUTER PHYSICS COMMUNICATIONS*, 121(Sp. Iss. SI):231–235, SEP-OCT 1999.

[141] S Singh and A Zlotnick. Observed hysteresis of virus capsid disassembly is implicit in kine tic models of assembly. *J. Biol. Chem.*, 278(20):18249–18255, MAY 16 2003.

[142] Hung D. Nguyen, Vijay S. Reddy, and Charles L. Brooks, III. Invariant Polymorphism in Virus Capsid Assembly. *JOURNAL OF THE AMERICAN CHEMICAL SOCIETY*, 131(7):2606–2614, FEB 25 2009.

[143] R Schwartz, PW Shor, PE Prevelige, and B Berger. Local rules simulation of the kinetics of virus capsid self-assembly. *BIOPHYSICAL JOURNAL*, 75(6):2626–2636, DEC 1998.

[144] B BERGER, PW SHOR, L TUCKERKELLOGG, and J KING. LOCAL RULE-BASED THEORY OF VIRUS SHELL ASSEMBLY. *PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF THE UNITED STATES OF AMERICA*, 91(16):7732–7736, AUG 2 1994.